

# Online Appendix: Supporting Information for *The Shadow of Deterrence: Why capable actors engage in contests short of war*

J Andres Gannon      Erik A. Gartzke      Jon R. Lindsay      Peter Schram

2023-03-10

## Contents

<b>1</b>	<b>Main Model (in the Paper) Discussion</b>	<b>2</b>
1.1	Parameters Used in Figures . . . . .	2
1.2	On Only Considering Pure Strategy Equilibria . . . . .	2
1.3	Equilibrium . . . . .	3
1.4	More Details on Why C Sometimes Risks War . . . . .	9
1.5	Proving Observations 1 and 2 . . . . .	13
1.6	Proving Observations 3 and 4 . . . . .	21
1.7	Observations 5-6 . . . . .	27
1.8	Discussion and Proof of Observation 7 . . . . .	27
<b>2</b>	<b>Model Extension 1: Correlated Gray Zone and War Costs</b>	<b>35</b>
2.1	Model Setup (Correlated Gray Zone/War Costs) . . . . .	35
2.2	Model Extension 1: When Improvements in Gray Zone Capabilities Lead to More War . . . .	41
<b>3</b>	<b>Extension 2: Commitment Problems</b>	<b>43</b>
3.1	Extension 2 Model: Game form and utilities . . . . .	44
3.2	Extension 2 Model: Equilibria Types and Existence . . . . .	45
3.3	Proving Proposition 3: Stable Gray Zone Equilibrium . . . . .	47
<b>4</b>	<b>New data</b>	<b>53</b>
4.1	Comparison of current datasets . . . . .	53
4.2	Variable codings . . . . .	53
4.3	Summary statistics . . . . .	55
<b>5</b>	<b>Alternate model specifications</b>	<b>57</b>
5.1	Alternate alliance measure . . . . .	57
5.2	Alternate distance measures . . . . .	59
5.3	Odds ratios . . . . .	60
5.4	OLS regression . . . . .	60
5.5	Ordered logit . . . . .	60
5.6	Multiple imputation . . . . .	60
5.7	Targeted states sample . . . . .	62
<b>6</b>	<b>Case Study: US 2016</b>	<b>65</b>
	<b>References</b>	<b>66</b>

This appendix accompanies the paper “The Shadow of Deterrence: Why capable actors engage in contests short of war”. It provides supplemental information concerning proofs for the formal model, the data set of Russian gray zone campaigns introduced in the paper, and robustness checks and alternate specifications for the statistical model.

## 1 Main Model (in the Paper) Discussion

### 1.1 Parameters Used in Figures

#### 1.1.1 Figure 1

All three panels share the following parameters:  $\rho = 0.1$ ,  $\kappa_C = 0.1$ ,  $\kappa = 0.3$ ,  $\psi = 0.12$ ,  $\beta_D = 5$ ,  $\theta = 3$ , and  $\omega = 0.1$ . In the top panel  $\beta_C = 3$ , in the middle panel  $\beta_C = 5$ , and in the bottom panel  $\beta_C = 8$ .

#### 1.1.2 Figure 2

Both panels share the following parameters:  $\rho = 0.1$ ,  $\kappa_C = 0.1$ ,  $\kappa = 0.2$ ,  $\psi = 0.45$ ,  $\beta_D = 4$ ,  $\theta = 3.5$ , and  $\omega = 0.1$ . In the top panel  $\beta_C = 1.8$  and in the bottom panel  $\beta_C = 2.5$ .

#### 1.1.3 Figure 3

$\rho = 0.1$ ,  $\kappa_C = 0.1$ ,  $\psi = 0.1$ ,  $\beta_D = 1.5$ ,  $\beta_C = 0.9$ ,  $\theta = 1.25$ , and  $\omega = 0.2$ . The parameter  $\kappa$  varies from 0.5 to 0.245.

#### 1.1.4 Figure 4

$\rho = 0.1$ ,  $\kappa_C = 0.15$ ,  $\kappa = 0.4$ ,  $\psi = 0.3$ ,  $\beta_D = 1.8$ ,  $\beta_C = 1$ , and  $\omega = 0.02$ . The parameter  $\theta$  varies from 1 to 2.06.

#### 1.1.5 Figure 5

$\rho = 0.05$ ,  $\kappa_C = 0.3$ ,  $\kappa = 0.4$ ,  $\psi = 0.2$ ,  $\beta_D = 3.7$ ,  $\theta = 2$ , and  $\omega = 0.05$ . The parameter  $\beta_C$  varies from 0.8 to 2.51.

#### 1.1.6 Figure 8

Both panels share the following parameters:  $\rho = 0.1$ ,  $\kappa_C = 0.04$ ,  $\kappa = 0.2$ ,  $\psi = 0.075$ ,  $\omega = 0.3$ ,  $\beta_C = 0.5$ ,  $\theta = 1.6$ . In the top panel  $\beta_D = 5$  and in the bottom panel  $\beta_D = 1.5$ .

### 1.2 On Only Considering Pure Strategy Equilibria

This paper only considers pure strategy Perfect Bayesian Nash Equilibria. Naturally including mixed strategy equilibria does not rule out the equilibrium in the text, but it does introduce the possibility of knife-edge cases where one player may mix over their actions in such a manner that does not "break" the equilibrium. For example, it could be that for a given set of parameters, C is indifferent between playing a  $g_C$  that sometimes risks war and playing a  $g_C$  that induces all types of D to conduct a limited response. In these edge-cases, C could mix over these two  $g_C$  values, resulting in a lower utility for some types of D (how much lower depends on how much mixing occurs). Alternatively, in the game, for some  $g_C$ , D is indifferent between declaring war in the third stage or conducting a limited response and engaging in gray zone conflict. With some small probability, these D's could sometimes go to war without breaking the equilibrium.<sup>1</sup> Ultimately, these cases are not particularly interesting to the analysis; because these mixed strategies are knife-edge cases and add complications with little added theoretical insight, they are excluded.

---

<sup>1</sup>This will function so long that C is also not indifferent between playing the  $g_C$  that always results in gray zone conflict and playing the status quo.

### 1.3 Equilibrium

#### 1.3.1 Defining C's Selected Equilibrium Actions (for example $g_C^{W,EDT}$ ) and Utilities

We abuse notation in a specific way here. For example, we let  $U_C(g_C = g_C^{GZ,IE} | g_C^{GZ,DTB} \geq g_C^{GZ,IE})$  denote C setting  $g_C = g_C^{GZ,IE}$  conditional on the game parameters being such that  $g_C^{GZ,DTB} \geq g_C^{GZ,IE}$ . Recall that  $g_C^{W,EDT} = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $g_C^{W,IE} = \frac{(1-\omega)\theta}{2\beta_C}$ ,  $g_C^{GZ,EDT} = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , and  $g_C^{GZ,IE} = \frac{\theta}{2\beta_C}$ . C's utilities are as follows:

$$U_C(g_C = 0) = \theta\rho - \kappa_C,$$

$$U_C(g_C = g_C^{GZ,IE} | g_C^{GZ,DTB} \geq g_C^{GZ,IE}) = \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C,$$

$$U_C(g_C = g_C^{GZ,DTB} | g_C^{GZ,DTB} < g_C^{GZ,IE}) = \theta \left( \rho - \frac{1}{4\beta_D} + \underline{\kappa}_D \right) - \beta_C \left( g_C^{GZ,EDT} \right)^2,$$

$$U_C(g_C = g_C^{W,IE} | g_C^{W,DTB} \geq g_C^{W,IE}) = \omega(\theta\rho - \kappa_C) + (1-\omega) \left( \theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C \right) - \beta_C \left( g_C^{W,IE} \right)^2, \text{ and}$$

$$U_C(g_C = g_C^{W,DTB} | g_C^{W,DTB} < g_C^{W,IE}) = \omega(\theta\rho - \kappa_C) + (1-\omega)\theta \left( \rho - \frac{1}{4\beta_D} + \bar{\kappa}_D \right) - \beta_C \left( g_C^{W,EDT} \right)^2.$$

#### 1.3.2 Mathematical Background for Equilibrium

#### 1.3.3 Stages 4 & 5

D will optimally avoid war in Stage 5. To do so, D will offer C some  $x^*$  that makes C indifferent between setting  $w_C = 0$  and setting  $w_C = 1$ . This is

$$\theta x^* - \beta_C g_C^2 = \theta P(g_C, g_D) - \kappa_C - \beta_C g_C^2$$

or

$$x^* = P(g_C, g_D) - \frac{\kappa_C}{\theta}.$$

If the game reaches this stage, then C's payoff will be  $U_C = \theta P(g_C, g_D) - \kappa_C - \beta_C g_C^2$ , and D's payoff will be  $U_D = 1 - P(g_C, g_D) + \frac{\kappa_C}{\theta} - \beta_D g_D^2$ .

#### 1.3.4 Stage 3

Suppose it is optimal for D to not go to war. We can solve for D's optimal gray zone response  $g_D^*$ . For a fixed  $g_C$  this will be

$$g_D^* \in \argmax \left\{ 1 - P(g_C, g_D) + \frac{\kappa_C}{\theta} - \beta_D g_D^2 \right\}.$$

Of course, we cannot simply take first order conditions due to the bounds on  $P$ . So, to define  $g_D^*$ , we first characterize value  $\hat{g}_D$ , which is

$$\hat{g}_D \in \argmax \left\{ 1 - (\rho + g_C - g_D) + \frac{\kappa_C}{\theta} - \beta_D g_D^2 \right\},$$

or

$$\hat{g}_D = \frac{1}{2\beta_D}.$$

Clearly D will not select a level of gray zone  $g_D > g_C$  because this would be counterproductive (as it runs into the lower bound at  $\rho$ ). There will also need to be some considerations with respect to the upper bound

that we will ignore for now.<sup>2</sup> Supposing the upper bound does not bind, we can define D's optimal level of gray zone conflict as

$$g_D^*(g_C) = \begin{cases} \frac{1}{2\beta_D} & \text{if } \frac{1}{2\beta_D} \leq g_C \\ g_C & \text{otherwise} \end{cases}$$

### 1.3.5 Stage 2: C's Optimal Choice and the Constraints that Drive It

Given all this, we can characterize how C optimally plays the game. Essentially, C can do one of five things. (1) C can accept the status quo. (2) C can engage in gray zone conflict defined by C's own internal cost-benefit balancing. (3) C can engage in gray zone conflict defined by the external deterrent threat. (4) C can select a level of gray zone conflict that results in war sometimes, where the selected limited response is analogous to C's internal efficiency constraint binding knowing that type  $\underline{\kappa}_D$ 's will go to war. Or (5) C can select a level of gray zone conflict that results in war sometimes, where the selected limited response is analogous to the external deterrent threat binding. We define these cases formally below.

**1.3.5.1 Status Quo** C could select  $g_C = 0$ , which will result in D selecting  $g_D^* = 0$ . This gives C utility

$$U_C(g_C = 0) = \theta\rho - \kappa_C,$$

and both types of D

$$U_D(g_C = 0) = 1 - \rho + \frac{\kappa_C}{\theta}.$$

**1.3.5.2 Gray Zone Conflict** First, consider the set of  $g_C$  that will result in neither type of D going to war (setting  $w_D = 0$ ) and that could be productive for C. On the latter point, whenever  $g_C \leq \frac{1}{2\beta_D}$ , D will select a limited response that fully undercuts C's challenge. Thus, we know that C would never want to select any  $g_C \in (0, \frac{1}{2\beta_D})$ , because C would prefer to set  $g_C = 0$  and not pay the costs from challenging. Next, we consider what values of  $g_C$  will keep D from going to war. This will be values of  $g_C$  that maintain the following constraint:

$$1 - (\max\{\min\{1, \rho + g_C - g_D^*\}, \rho\}) + \frac{\kappa_C}{\theta} - \beta_D (g_D^*)^2 \geq 1 - \rho - \underline{\kappa}_D.$$

So long that  $g_C \geq \frac{1}{2\beta_D}$  and the upper bound does not bind, we can simplify this expression to the following:

$$1 - \left(\rho + g_C - \frac{1}{2\beta_D}\right) + \frac{\kappa_C}{\theta} - \beta_D \left(\frac{1}{2\beta_D}\right)^2 \geq 1 - \rho - \underline{\kappa}_D,$$

or

$$g_C \geq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}.$$

With this, we can set up an optimization.

Suppose C does best selecting a limited challenge that always results in gray zone conflict. Taking into account C's utility function and how the game plays out in equilibrium, C will select a  $g_C^*$  that optimizes the following

$$g_C^* \in \operatorname{argmax} \left\{ \theta \left( \rho + g_C - \frac{1}{2\beta_D} \right) - \kappa_C - \beta_D g_C^2 \right\},$$

such that

$$g_C \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}, \& \tag{1}$$

$$g_C \geq \frac{1}{2\beta_D}. \tag{2}$$

---

<sup>2</sup>Specifically, it must be that  $\rho + g_C - g_D^* \leq 1$ .

Note that we do not need to worry about the upper bound on  $P$  binding for C's selected  $g_C$  because of Assumption 1 in the text. The objective function  $\theta \left( \rho + g_C - \frac{1}{2\beta_D} \right) - \kappa_C - \beta_C g_C^2$  can be thought of as C's internal cost-benefit calculations. As C selects greater levels of  $g_C$ , then C benefits because it's a more aggressive limited challenge that helps in bargaining later in the game. However, C does experience increased costs from the more aggressive limited challenge, as captured in the  $-\beta_C g_C^2$  term. If the constraints do not bind, then C would select  $g_C^* = \frac{\theta}{2\beta_C}$ . In this regard, even though optimizing the objective function and disregarding (1) and (2) is an unconstrained optimization problem, C is constrained by the costs of conducting limited challenges ( $g_C^*$  decreases as C's costs of limited challenges,  $\beta_C$ , increases). While this is not a constraint in a strictest mathematical sense ((1) and (2) are the optimization constraints), C's selected limited challenge is constrained the costs of conducting challenges. We will refer to this as C's internal efficiency constraint binding.

As discussed previously, constraint (1) represents how far C can go before D is unwilling to conduct a gray zone response. This means that if C wants to always avoid war with D, C cannot select a  $g_C$  that exceeds  $\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ . This is a constraint (in the standard mathematical sense of the word): if C has very low costs to conducting limited challenges but still wants to always avoid war (i.e.  $\beta_C$  is very low), then C will select the greatest limited challenge that C can without trigger D to go to war, which is  $g_C^* = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ .

This generates some of our mathematical expressions. Still suppose it is optimal for C to always select into gray zone conflict. When  $\frac{\theta}{2\beta_C} < \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , then C's optimal limited challenge is  $g_C^* = \frac{\theta}{2\beta_C}$ . Here C will select their  $g_C$  based on their internal cost-benefit analysis without needing to take into account D's threat of war. Whenever  $\frac{\theta}{2\beta_C} < \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , we will say that C's internal efficiency constraint binds. Alternatively, when  $\frac{\theta}{2\beta_C} \geq \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , C's optimal limited challenge is  $g_C^* = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ . Here C would like to select a greater limited challenge, but is constrained by the threat of D escalating to war. Whenever  $\frac{\theta}{2\beta_C} \geq \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , we will say that D's external deterrent threat binds. These values are quite important to the game, and are repeated often. We let  $g_C^{GZ,EDT} = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$  and  $g_C^{GZ,IE} = \frac{\theta}{2\beta_C}$ .

Whenever the internal efficiency constraint binds and C sets  $g_C^* = g_C^{GZ,IE}$ , C's utility is

$$U_C(g_C = g_C^{GZ,IE}) = \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C.$$

And D's utility is (for both types) is

$$U_D(g_C = g_C^{GZ,IE}) = 1 - \left( \rho + \frac{\theta}{2\beta_C} - \frac{1}{2\beta_D} \right) + \frac{\kappa_C}{\theta} - \frac{1}{4\beta_D}$$

or

$$U_D(g_C = g_C^{GZ,IE}) = 1 - \rho - \frac{\theta}{2\beta_C} + \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}.$$

And, whenever the external threat constraint binds and C sets  $g_C^* = g_C^{GZ,EDT}$ , C's utility is

$$U_C(g_C = g_C^{GZ,EDT}) = \theta \left( \rho - \frac{1}{4\beta_D} + \kappa_D \right) - \beta_C \left( \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta} \right)^2.$$

And both types of D will receive

$$U_D(g_C = g_C^{GZ,EDT}) = 1 - \rho - \kappa_D.$$

#### **Type $\underline{\kappa}_D$ fights and type $\bar{\kappa}_D$ conducts gray zone conflict.**

Sometimes it is optimal for C to risk war with type  $\underline{\kappa}_D$  but not type  $\bar{\kappa}_D$ . Fighting occurs on the equilibrium path because, while C does not like fighting a war, C may be willing to sometimes fight in order to undertake a more aggressive limited challenges. C will select a limited challenge that C (a) knows will end with war if D is type  $\underline{\kappa}_D$  (the types that have low costs to fighting a war) and (b) knows will not end in war if D is type

$\bar{\kappa}_D$  (the types that have high costs to fighting a war). This is analogous to the standard risk-return tradeoff within crisis bargaining: instead of one actor being uncertain over how aggressive a bargaining position they can take, here C is uncertain over how aggressive a limited challenge they can take before D declares war.

Now we consider the set of  $g_C$  that will result in type  $\underline{\kappa}_D$  going to war, type  $\bar{\kappa}_D$  not going to war, and that is productive for C. Following the discussion above, type  $\underline{\kappa}_D$  will go to war for any  $g_C > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ . And similarly, type  $\bar{\kappa}$  will not go to war whenever  $g_C \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ . When C is selecting a  $g_C$  that risks war with type  $\underline{\kappa}_D$  but not type  $\bar{\kappa}_D$ , C faces the following optimization problem:

$$g_C^* \in \operatorname{argmax} \left\{ \omega(\theta\rho - \kappa_C) + (1-\omega)\theta \left( \rho + g_C - \frac{1}{2\beta_D} - \frac{\kappa_C}{\theta} \right) - \beta_C g_C^2 \right\},$$

such that

$$g_C \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}, \& \quad (3)$$

$$g_C > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}, \& \quad (4)$$

$$g_C \geq \frac{1}{2\beta_D} \quad (5)$$

Note that if this optimization problem runs into open set issues, then C always prefers avoiding war.<sup>3</sup>

As it was above, here C faces analogous internal efficiency and external deterrent threat constraints. The optimum of the objective function (without constraints) is at  $g_C = \frac{(1-\omega)\theta}{2\beta_C}$ , which is the value where C's costs become greater than the additional benefit C gets from the more aggressive challenge. And, if constraint (3) binds, then the optimization solution is  $g_C = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ , which is the greatest  $g_C$  C can select without pushing type  $\bar{\kappa}_D$  to go to war (i.e. without pushing both types of D to go to war). Because these values are repeated often, we let  $g_C^{W,EDT} = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$  and  $g_C^{W,IE} = \frac{(1-\omega)\theta}{2\beta_C}$ .

Whenever the internal efficiency constraint binds and C sets  $g_C^* = g_C^{W,IE}$ , C's utility is

$$U_C(g_C = g_C^{W,IE}) = \omega(\theta\rho - \kappa_C) + (1-\omega) \left( \theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C \right) - \beta_C \left( \frac{(1-\omega)\theta}{2\beta_C} \right)^2.$$

good-at-war types would attain utility

$$U_D(g_C = g_C^{W,IE}, \underline{\kappa}_D) = 1 - \rho - \underline{\kappa}_D,$$

and bad-at-war types would attain utility

$$U_D(g_C = g_C^{W,IE}, \bar{\kappa}_D) = 1 - \rho - \frac{(1-\omega)\theta}{2\beta_C} + \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}.$$

And, whenever the external threat constraint binds and C sets  $g_C^* = g_C^{GZ,EDT}$ , C's utility is

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

And both types of D will receive their respective wartime payoffs, or

$$U_D(g_C = g_C^{W,EDT}, \underline{\kappa}_D) = 1 - \rho - \underline{\kappa}_D$$

and

$$U_D(g_C = g_C^{W,EDT}, \bar{\kappa}_D) = 1 - \rho - \bar{\kappa}_D.$$

---

<sup>3</sup>Suppose the objective function optimized over  $g_C \in [\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}, \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}]$  would have the solution at  $g_C = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ . If this is the case, then C prefers values of  $g_C$  that avoid war to sometimes risking war. Note this may not be C's optimal choice, because C could also select  $g_C = 0$ .

### 1.3.6 Intuition on the Constraints in C's Decision-Making in Stage 2

Here we offer some intuition behind how C behaves in Stage 2. Note that this is not necessary to understanding the equilibrium, but it may be helpful if aspects of C's behavior are unclear. Case in point: in the text, we present a series of inequalities that characterize when C selects one level of challenge over another; we elaborate on each of these inequalities here. Note that we will not go into details on when some  $g_C$  values are greater than others (for example, when  $g_C^{W,IE} > g_C^{W,GZ}$ )—this follows from the discussion above.

When  $\theta\rho - \kappa_C < \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(g_C^{W,EDT}\right)^2$  (first referenced in Sometimes War, Deterrent Threat Binds), then C's payoff from staying in the status quo ( $\theta\rho - \kappa_C$ ) is less than C's payoff from risking war with probability  $\omega$  and selecting the limited challenge that is characterized by type  $\bar{\kappa}_D$ 's deterrent threat ( $\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(g_C^{W,EDT}\right)^2$ ). This is a utility comparison. When the constraint is flipped, then C prefers accepting the status quo to ever risking war when the external deterrent threat binds..

When  $\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(g_C^{W,EDT}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2$  (first referenced in Sometimes War, Deterrent Threat Binds), then C's payoff from selecting into gray zone conflict and never risking war is less than C's payoff from risking war with probability  $\omega$  and selecting the limited challenge that is characterized by type  $\bar{\kappa}_D$ 's deterrent threat. This is a utility comparison. When the constraint is flipped, then C prefers conducting gray zone conflict without resorting to war to sometimes risking war.

When  $\theta\rho - \kappa_C < \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(g_C^{W,IE}\right)^2$  (first referenced in Sometimes War, Internal Efficiency Binds), then C's payoff from staying in the status quo is less than C's payoff from risking war with probability  $\omega$  and selecting the limited challenge that is characterized by C's internal efficiency ( $g_C = \frac{(1-\omega)\theta}{2\beta_C}$ ). This is a utility comparison. When the constraint is flipped, then C prefers accepting the status quo to risking war.

When  $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2 < \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(g_C^{W,IE}\right)^2$  (first referenced in Sometimes War, Internal Efficiency Binds), then C's payoff from conducting gray zone conflict and never risking war is less than C's payoff from risking war with probability  $\omega$  and selecting the limited challenge that is characterized by C's internal efficiency ( $g_C = \frac{(1-\omega)\theta}{2\beta_C}$ ). This is a utility comparison. When the constraint is flipped, then C prefers conducting gray zone conflict without ever risking war to sometimes risking war.

When  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2$  (first referenced in Always Gray Zone, Deterrent Threat Binds), then C's limited challenge as characterized by the deterrent threat binding ( $g_C = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ) produces a payoff for C ( $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(g_C^{GZ,EDT}\right)^2$ ) that is greater than C's payoff from accepting the status quo  $\theta\rho - \kappa_C$ . This is a utility comparison; when it does not hold, C does better accepting the status quo to selecting into gray zone conflict.

When  $\theta\rho - \kappa_C < \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$  (first referenced in Always Gray Zone, Internal Efficiency Binds), then C's payoff from not conducting a limited challenge and ending the game with offer  $x^* = \theta\rho - \kappa_C$  is less than C's payoff from conducting a limited challenge based on their internal efficiency constraint ( $g_C = g_C^{GZ,IE}$ ) and ending the game with gray zone conflict and payoff  $\theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$ . Essentially, this is a utility comparison, and when it does not hold, C does better engaging in gray zone conflict (where the internal efficiency constraint holds) relative to accepting the status quo.

Lastly, below (but not in the main paper!) we will consider the possibility that  $\frac{\theta}{2\beta_C} \leq \frac{1}{2\beta_D}$ . When this holds, it is the case that C's optimal limited challenge (i.e. when the internal efficiency constraint binds) of  $g_C = \frac{\theta}{2\beta_C}$  is less than  $g_D = \frac{1}{2\beta_D}$ , which is how far D is willing to go with their gray zone response. C will be unwilling to ever challenge here, because D's gray zone response will essentially negate C's (costly) limited challenge, thus generating costs without gains for C; C can do better setting  $g_C = 0$ . When this constraint

does not hold, then it is possible for C's limited challenge to overcome D's gray zone response, but this is not to say that gray zone conflict becomes cost effective (this will depend on utility constraints as well).

### 1.3.7 Statement of Full Equilibrium

Below we have the full statement of the equilibrium where we impose no restrictions on the relationship between  $\frac{\theta}{2\beta_C}$  and  $\frac{1}{2\beta_D}$ . Additionally, we elaborate on the cases that are discussed in the equilibrium cases in the appendix—we fully spell out all the conditions rather than include the “or” statements. In proving Observations 1-6, breaking these terms apart makes for a much easier series of proofs.

Also below, we discuss a series of ways C can behave in stage 2. While there are many cases below, it is useful to think about clusters of cases where preliminaries (i.e. conditions like  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ) hold across the cases, and all that differs are the utility conditions. Using this typology, Status Quo 2 is most similar to Gray Zone, Internal Efficiency Binds because all that differs is the  $\theta\rho - \kappa_C \geq \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$  utility condition. Status Quo 3 is most similar to Gray Zone, Deterrent Threat Binds 1. Status Quo 4 is most similar to Gray Zone, Deterrent Threat Binds 2 and War 1. Status Quo 5 is most similar to Gray Zone, Deterrent Threat Binds 3 and War 2. Status Quo 1 does not share any preliminary conditions.

#### In Stage 5:

If  $x \geq P(g_C, g_D) - \frac{\kappa_C}{\theta}$ , then C will set  $w_C = 0$ . Otherwise C will set  $w_C = 1$ . 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  $1 - \omega$  ( $\omega$ ).

#### In Stage 4:

D will always set  $x = P(g_C, g_D) - \frac{\kappa_C}{\theta}$ .

#### In Stage 3:

For all  $\kappa_D \in \{\underline{\kappa}_D, \bar{\kappa}_D\}$ , 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  $g_D = \min\left\{\frac{1}{2\beta_D}, g_C\right\}$ .

#### In Stage 2:

C believes D is type  $\bar{\kappa}_D$  with probability  $1 - \omega$  and type  $\underline{\kappa}_D$  with probability  $\omega$ .

What C does in Stage 2 is conditional on what parameters holds. These are the following.

**Status Quo 1:** If  $\frac{\theta}{2\beta_C} \leq \frac{1}{2\beta_D}$ , then  $g_C = 0$ .

**Status Quo 2:** If  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\theta\rho - \kappa_C \geq \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$ , then  $g_C = 0$ .

**Status Quo 3:** If  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , and  $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , then  $g_C = 0$ .

**Status Quo 4:** If  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , and  $\theta\rho - \kappa_C \geq \omega(\theta\rho - \kappa_C) + (1 - \omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2$ , then  $g_C = 0$ .

**Status Quo 5:** If  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , and  $\theta\rho - \kappa_C \geq \omega(\theta\rho - \kappa_C) + (1 - \omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , then



$g_C = 0$ .

**Gray Zone, Internal Efficiency Binds:** If  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\theta\rho - \kappa_C < \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$ , then  $g_C = \frac{\theta}{2\beta_C}$ .

**Gray Zone, Deterrent Threat Binds 1:** If  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , and  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , then  $g_C = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ .

**Gray Zone, Deterrent Threat Binds 2:** If  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , and  $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2$ , then  $g_C = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ .

**Gray Zone, Deterrent Threat Binds 3:** If  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , and  $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , then  $g_C = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ .

**War, Internal Efficiency Binds (War 1):** if  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 > \theta\rho - \kappa_C$ , and  $\omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , then  $g_C = \frac{(1-\omega)\theta}{2\beta_C}$ .

**War Deterrent Threat Binds (War 2):** if  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 > \theta\rho - \kappa_C$ , and  $\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , then  $g_C = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ .

## 1.4 More Details on Why C Sometimes Risks War

In this game, we have a defender who has a private type. The defender either has relatively lower costs to fighting (type  $\underline{\kappa}_D$ ), or has relatively higher costs to fighting (type  $\bar{\kappa}_D$ ). D knows their own type, but C does not. War occurs on the equilibrium path because, while C does not like fighting a war, C may be willing to sometimes risk a war in order to undertake a more aggressive challenges. Under some parameters, C will select a challenge that C (a) knows will end with war if D is type  $\underline{\kappa}_D$  and (b) knows will not end in war if D is type  $\bar{\kappa}_D$ . C chooses this more aggressive challenge that sometimes results in war over choosing a more limited challenge that never produces war. This is analogous to the standard risk-return tradeoff within crisis bargaining: instead of one actor being uncertain whether or not an aggressive bargaining position will lead to war (as is standard in crisis bargaining), here C is uncertain over how aggressive a limited challenge they can take before D declares war.

To be clear, under no parameters will C ever select a  $g_C$  that would result in both types of D declaring war. C can always do weakly better setting  $g_C = 0$  to setting any  $g_C$  that always results in war (i.e. any  $g_C > g_C^{W,EDT}$ ). In equilibrium, by setting  $g_C = 0$ , C attains utility  $U_C(g_C = 0) = \theta\rho - \kappa_C$ . And, in equilibrium, by setting some  $g_C = g'_C$  where  $g'_C > g_C^{W,EDT}$ , C attains utility  $U_C(g_C = g'_C) = \theta\rho - \kappa_C - \beta_C(g'_C)^2$ . The latter is strictly worse for any  $\beta_C > 0$ , and weakly worse when  $\beta_C = 0$ . This means that war only occurs in the model when C is selecting into this risk-return tradeoff rather than selecting a less aggressive challenge that would never lead to war. For this reason, below, we do not describe the payoffs from C selecting a  $g_C$  that always

results in war. And, while we do plot  $g_C$  values that would result in these regions (i.e. we do display the “Always War” region), the payoffs are regularly off-the-chart bad; we opt not to graph them rather than scale down the parameter spaces where the interesting choices are being made.

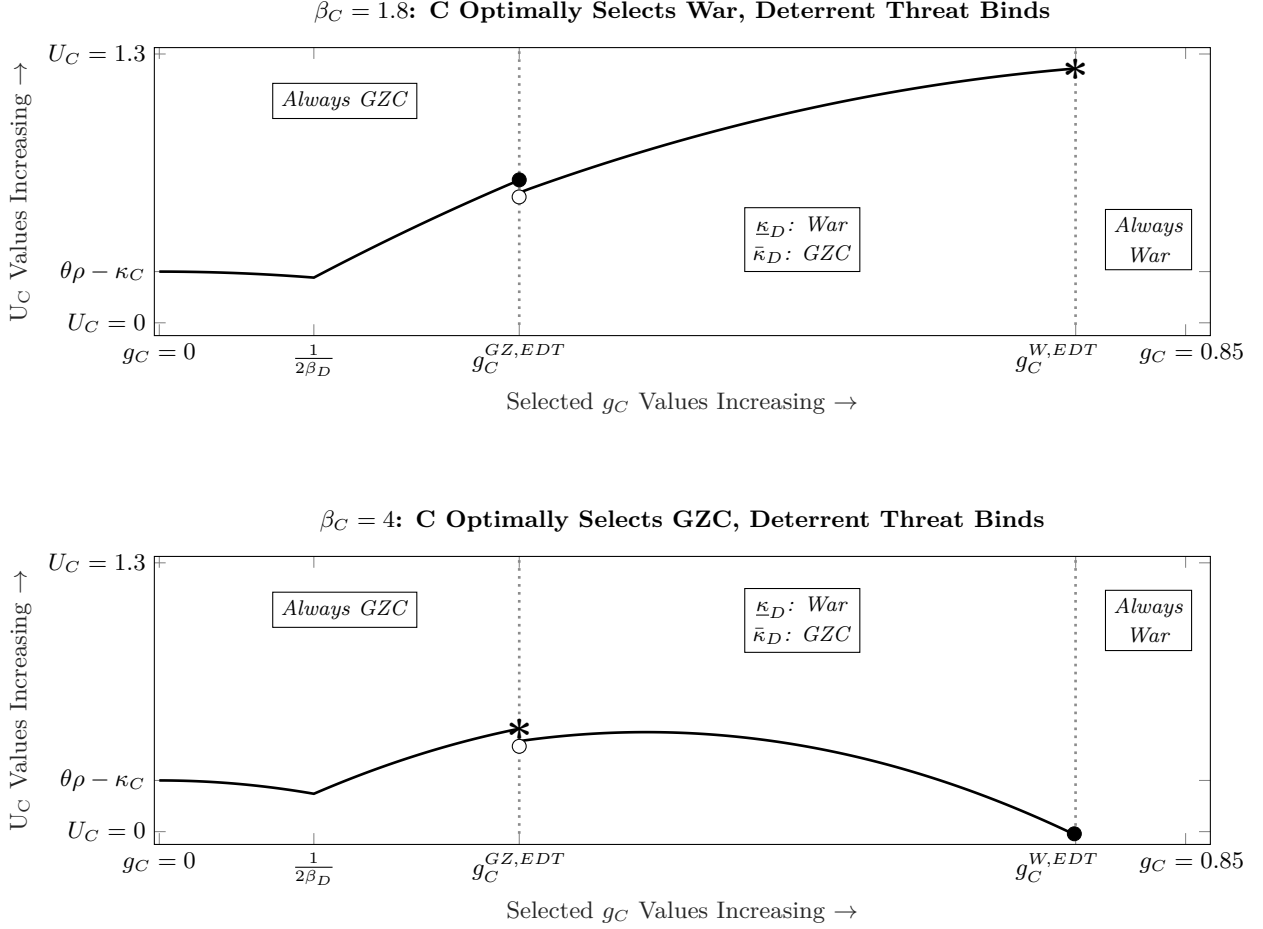


Figure A1: C's Utility Across selected  $g_C$ 's (War Outcomes).

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. In the top panel, the parameters are  $\beta_C = 1.8$ ,  $\rho = 0.1$ ,  $\kappa_C = 0.1$ ,  $\kappa = 0.2$ ,  $\psi = 0.45$ ,  $\beta_D = 4$ ,  $\theta = 3.5$ , and  $\omega = 0.1$ . In the bottom panel, we only manipulate  $\beta_C$ ; all the other parameters are the same, but  $\beta_C = 4$ .

This intuition can be visualized. Consider Figure A1, where the top panel (titled “ $\beta_C = 1.8$ : C Optimally Selects War, Deterrent Threat Binds”) matches the top panel in Figure 2 in the main text. For these parameters,  $g_C^{W,EDT} < g_C^{W,IE}$  and  $g_C^{GZ,EDT} < g_C^{GZ,IE}$ , meaning the external deterrent threat binds in C's selection into gray zone conflict and into sometimes risking war. As a quick summary of why the top panel looks the way it does, first consider the relationship between  $g_C$  and  $U_C$  when  $g_C \in [0, \frac{1}{2\beta_D}]$ ; in this region, C's utility is decreasing as D's optimal  $g_D^* = g_C$ , meaning in equilibrium C's challenge is effectively being cancelled out while C is also paying (increasing) costs for conducting the challenge.<sup>4</sup> In the  $g_C[\frac{1}{2\beta_D}, g_C^{W,EDT}]$  region, C's utility is increasing as  $g_C > g_D^* = \frac{1}{2\beta_D}$ , meaning that C's challenges are productive; while C is paying a cost to these challenges, these costs are fairly low, and only result in a slight concavity in this region.

<sup>4</sup>For these  $g_C$ , D will set  $w_D = 0$ . In the final stages of the game, D offers C  $x^* = P(g_C, g_D) - \frac{\kappa_C}{\theta}$ , where here  $P(g_C, g_D) = \rho + g_C - g_D$ .

$\beta_C = 1.8$ : C's Utility Outcomes					
Outcome	Status Quo	Always GZC	C Risks War	C Risks War, C Gets War	C Risks War, C Gets GZC
C sets $g_C =$	0	$g_C^{GZ,EDT} = 0.291$	$g_C^{W,EDT} = 0.741$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.125	-	-	0.125
$g_C - g_D^*$	0	0.166	-	-	0.6161
C's Payoffs	0.25	0.6787	1.202	-0.739	1.497

Table A1: C's choice in stage 2 and how the game plays out. The parameters here match the parameters in the top panel of A1. These parameter are  $\beta_C = 1.8$ ,  $\rho = 0.1$ ,  $\kappa_C = 0.1$ ,  $\kappa = 0.2$ ,  $\psi = 0.45$ ,  $\beta_D = 4$ ,  $\theta = 3.5$ , and  $\omega = 0.1$ . Note here it is optimal for C to sometimes risk war.

The discontinuity at  $g_C = g_C^{GZ,EDT}$  arises because at  $g_C = g_C^{GZ,EDT}$ , type  $\underline{\kappa}_D$  is indifferent between ending the game with gray zone conflict or war. Any  $g_C$  slightly to the right of  $g_C^{GZ,EDT}$  will result in a  $\omega$  likelihood of war without offering much added benefit should the game end in gray zone conflict. However, in the  $(g_C^{GZ,EDT}, g_C^{W,EDT}]$ , C's utility increases as C conducts more aggressive challenges. These more aggressive challenges still result in an  $\omega$  likelihood of war, but are productive with probability  $1 - \omega$ . This plot is also concave due to the costs of challenges, but only slightly. There is another discontinuity at  $g_C^{W,EDT}$  because here both types of D will go to war for any  $g_C$  to the right of this value. C's payoffs in the "Always War" space are off-the-chart bad and decreasing (starting at  $U_C = -0.739$  and decreasing).

In this top panel, the parameters are such that C optimally sets  $g_C^* = g_C^{W,EDT}$ , meaning C will go to war with probability  $\omega$  and will end the game in gray zone conflict with probability  $1 - \omega$ . This visualization offers some insight into the risk-return tradeoff mentioned above. C could select  $g_C = g_C^{GZ,EDT}$ , thus never risking war. However, by not selecting  $g_C = g_C^{W,EDT}$ , C is forgoing a very aggressive challenge that would only result in war with probability  $\omega = 0.1$ , which is represented by the gap between  $g_C = g_C^{GZ,EDT}$  and  $g_C = g_C^{W,EDT}$ . Ultimately here, C does best risking war sometimes rather than playing it safe and avoiding war altogether.

Table A1 goes into even more detail, and compliments the discussion above of the top panel in Figure @ref(fig:moreonwar\_fig). Moving left-to-right, the first two columns (labeled "Status Quo" and "Always GZC") describe what happens when C selects  $g_C = 0$  and when C selects C's best  $g_C \in (0, g_C^{GZ,EDT}]$  (i.e. C selects their optimal  $g_C$  conditional on the game always ending in gray zone conflict). As the second row reports, for these parameters, the best  $g_C$  that always results in gray zone conflict is  $g_C^{GZ,EDT}$ .<sup>5</sup> These two columns describe D's equilibrium response  $g_D^*$  to the selected  $g_C$ , then describes the difference between  $g_C$  and  $g_D^*$  ( $g_C - g_D^*$ ), which is essentially how much surplus C can get when the game ends in gray zone conflict. Finally, it describes C's expected payoff from the given  $g_C$ . Moving to the right—to the column labeled "C Risks War"—we do something similar, but we do not list  $g_D^*$  and  $g_C - g_D^*$ ; we do not do so because when C risks war, type  $\underline{\kappa}_D$  does not select a  $g_D^*$ ; rather, this type of D will go to war when  $g_C = g_C^{W,EDT}$ . However, we decompose what happens when C risks war in the last two columns. These columns, labeled "C Risks War, C Gets War" and "C Risks War, C Gets GZC" describe what happens when C selects (here)  $g_C = g_C^{W,EDT}$  and C faces a type  $\underline{\kappa}_D$  that declares war and type  $\bar{\kappa}_D$  that conducts gray zone conflict (respectively). It also describes the payoffs for these cases. When D is type  $\underline{\kappa}_D$  and C selects  $g_C = g_C^{W,EDT}$ , C's final (expected) payoff is  $-0.739$ . And, when D is type  $\bar{\kappa}_D$  and C selects  $g_C = g_C^{W,EDT}$ , C's final payoff is  $1.497$ . Together, when C sets  $g_C = g_C^{W,EDT}$ , in expectation, C attains payoff  $1.202$ , as reported in the "C Risks War" column.<sup>6</sup> As can be seen, within this set of parameters, C does best risking war and setting  $g_C^* = g_C^{W,EDT}$ .

In the bottom panel in Figure A1—labeled " $\beta_C = 4$ : C Optimally Selects GZC, Deterrent Threat Binds"—we leave all parameters the same but increase  $\beta_C$ , C's costs of conducting challenges. When  $\beta_C = 4$ , the parameters are such that C optimally sets  $g_C^* = g_C^{GZ,EDT}$ , meaning C does best by never go to war. This panel offers a useful counterpoint to the top panel, where war occurs. Comparing the top to bottom panel, we

<sup>5</sup>For these parameters,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , meaning if C selects into gray zone conflict, the external deterrent threat binds.

<sup>6</sup>When  $\omega = 0.1$  (as it is here),  $1.202 = \omega * (-0.739) + (1 - \omega) * 1.497$ .

$\beta_C = 4$ : C's Utility Outcomes					
Outcome	Status Quo	Always GZC	C Risks War	C Risks War, C Gets War	C Risks War, C Gets GZC
C sets $g_C =$	0	$g_C^{GZ,EDT} = 0.291$	$g_C^{W,IE} = 0.3938$	$g_C^{W,IE}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,IE}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.125	-	-	0.125
$g_C - g_D^*$	0	0.166	-	-	0.2688
C's Payoffs	0.25	0.4924	0.4764	-0.37	0.5705

Table A2: C's choice in stage 2 and how the game plays out. The parameters here match the parameters in the top panel of A1, only here  $\beta_C$  is greater. These parameter are  $\beta_C = 4$ ,  $\rho = 0.1$ ,  $\kappa_C = 0.1$ ,  $\kappa = 0.2$ ,  $\psi = 0.45$ ,  $\beta_D = 4$ ,  $\theta = 3.5$ , and  $\omega = 0.1$ . Note here it is optimal for C to never risk war.

can see three clear changes. First, for all  $g_C > 0$ , C's utility is lower, which is due to the increased costs from challenging. Second, C's utility curves are now more concave, and, in the case of  $g_C \in (g_C^{GZ,EDT}, g_C^{W,EDT}]$ , sometimes decreasing in  $g_C$ . In the bottom panel,  $g_C^{W,IE} \leq g_C^{W,EDT}$ , meaning the internal efficiency constraint binds when C sometimes risks war; this implies for all  $g_C \in [g_C^{W,IE}, g_C^{W,EDT}]$ , C's utility is decreasing in  $g_C$ . This feature accounts for C's utility decreasing in  $g_C$  in this region. Third, the greater  $\beta_C$  now means that C is unwilling to sometimes risk war.

With parameter  $\beta_C = 4$ , C does much worse risking war. Why? Table A2 offers more details on C's choice here. C's best  $g_C$  that will sometimes result in war is  $g_C^{W,IE} = 0.3938$ . This is a less aggressive challenge than C setting  $g_C = g_C^{W,EDT}$ , which means that C's challenge isn't yielding the robust returns in the bargaining in Stages 4 and 5.<sup>7</sup> Previously, when  $\beta_C = 1.8$ , by risking war the  $g_C - g_D^*$  value (which represents C's bargaining surplus down the line) jumps from 0.1667 when  $g_C = g_C^{GZ,EDT}$  to 0.6161 when  $g_C = g_C^{W,EDT}$ . In contrast, when  $\beta_C = 4$ , by risking war the  $g_C - g_D^*$  value jumps from 0.1667 when  $g_C = g_C^{GZ,EDT}$  to 0.2688 when  $g_C = g_C^{W,IE}$ , which is a much smaller difference. C isn't risking more when  $\beta_C = 4$  because then C's costs for challenging overwhelm the benefit from challenging. Together, this means that when  $\beta_C = 4$ , even if C risks war and then (luckily) doesn't have to go to war (i.e. faces a  $\bar{\kappa}_D$ ), C does not do particularly well. When C risks war then gets gray zone conflict under  $\beta_C = 1.8$  C attains utility 1.497; when C risks war then gets gray zone conflict under  $\beta_C = 4$  C attains 0.5705. For this reason, ultimately, C will not risk war when  $\beta_C = 4$ .

Of course, there are other ways to incentivise C to risk war outside of lowering  $\beta_C$ . This can be seen much later in the memo (in Figure A2), where, despite  $\beta_C = 0$ , C sometimes wants to risk war and other times C does not. This in particular is an interesting case because when  $\beta_C = 0$  the internal efficiency constraint never actually binds yet.

Finally, we can also conduct a more mathematically oriented analysis to illustrate when war occurs. Suppose  $g_C^{GZ,EDT} < g_C^{GZ,IE}$  and  $g_C^{W,EDT} < g_C^{W,IE}$ . When these hold, C prefers risking war sometimes when

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

A complete analysis of comparative statics is beyond the scope of this paper, but here is a brief discussion. Recall that  $-\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta} \geq 0$ , which implies that  $(\theta\rho - \kappa_C) \leq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) \leq \theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right)$ . Forgetting about costs (for now), this means by risking war, C is forgoing a middle-ground certain payoff of  $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right)$  for a gamble between a worse outcome  $(\theta\rho - \kappa_C)$  and a better outcome  $\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right)$ . One factor that makes the gamble more appealing is if  $\omega$  is low, which means there is a lower likelihood of C getting the worst payoff of  $(\theta\rho - \kappa_C)$ . Additionally, C is more willing to risk war if  $\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right)$  is much larger than  $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right)$ , which occurs when there is a big difference between  $\bar{\kappa}_D$  and  $\underline{\kappa}_D$ . This results in a big difference between  $g_C^{GZ,EDT}$  (which recall is defined as  $g_C^{GZ,EDT} = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ) and  $g_C^{W,EDT}$

<sup>7</sup>Recall D offers C  $x^* = P(g_C, g_D) - \frac{\kappa_C}{\theta}$ , where, for the values considered here,  $P(g_C, g_D) = \rho + g_C - g_D$ .

(which recall is defined as  $g_C^{W,EDT} = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ), which represents C undertaking much more aggressive challenges by risking war.

One factor that can challenge this logic is when the costs of challenging are high. When  $\beta_C$  is large, then C incurs more disutility from selecting greater challenges. This can disincentivise C from selecting the larger challenge that sometimes results in war (here  $g_C^{W,EDT}$ ), instead encouraging C into select a more restrained challenge.

Now suppose  $g_C^{GZ,EDT} < g_C^{GZ,IE}$  and  $g_C^{W,IE} \leq g_C^{W,EDT}$ . This produces a different mathematical constraint, but, once more, C is willing to risk war when the following holds:

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

While it is a bit harder to see when C will want to risk war, a similar logic still holds. When C risks war, C selects  $g_C^{W,IE} = \frac{(1 - \omega)\theta}{2\beta_C}$ . And, when C never risks war, C selects  $g_C^{GZ,EDT} = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ . When there is a greater difference between these values, so long that  $\beta_C$  remains low, C has a greater incentive to risk war.

## 1.5 Proving Observations 1 and 2

Recall that  $\underline{\kappa}_D = \kappa$  and  $\bar{\kappa} = \kappa + \psi$  where  $\psi > 0$ . Observation 1 is the following.

**Observation 1:** *If increasing the defender's willingness to go to war (decreasing  $\kappa$ ) 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.*

Note: every equilibrium case referenced here matches the discussion of Equilibrium in the appendix, and not the equilibrium in the text. We prove all matters here allowing for the possibility that  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$  does not hold.

To prove these Observations, we show that as  $\kappa$  decreases, we might observe positive, negative, or no changes on  $g_C^*$ . However there is a structure to this relationship; the decrease in  $\kappa$  will always result in a decrease in  $g_C^*$  when the deterrent threat constraint plays a role in C's selection (Observation 1), and the decrease in  $\kappa$  will always either result in an increase or no change in  $g_C^*$  when the internal efficiency constraint plays a role (Observation 2). Part of this proof is straightforward; it is possible to look within each parameter space (for example, referencing C's behavior in stage 2, the "War, Deterrent Threat Binds" parameter space) and identify the effects of  $\kappa$  on  $g_C^*$ . However, because changes in  $\kappa$  can also result in changes to the underlying conditions of what parameter space holds, these jumps must also be considered.

Here is how we will proceed. We start by assuming the parameters are such that we are in one of the equilibrium cases. So below, we first consider that we are in the case where the parameters are such that C will risk war sometimes, but the deterrent threat will bind. We then consider what happens at  $\kappa$  decreases—whether this can result in the game ending in a different equilibrium, and what the final effects on the selected challenge will be. In doing so, we can show that whenever decreasing  $\kappa$  results in  $g_C^*$  decreasing, this occurs because the external deterrent threat is influencing C's final choices.

By proceeding in this way, we can identify the drivers of the final effect of  $\kappa$  on  $g_C^*$ . We will show, for example, whenever  $\kappa$  decreases  $g_C^*$  will only decrease when C's external deterrent threat plays a role. For example, below, within the "War, Deterrent Threat Binds" case, we can see that  $g_C^*$  is decreasing in  $\kappa$ . This happens here C is responding to deterrence. What would disprove Observation 1 would be, for example, if C is choosing a  $g_C^*$  within the "Gray Zone, Deterrent Threat Binds" case, and  $g_C^*$  is increasing in  $\kappa$ . What happens below is a little more complicated than that—because we need to consider jumps across equilibrium cases—but this is the gist of what we are doing.

As a point of clarification, in the discussion surrounding Observations 1 and 2, we discussed the possibility that as  $\kappa$  decreases, it could push C into moving from gray zone conflict into war. This means that it is

possible to start in Gray Zone, Deterrent Threat Binds 3 and, as  $\kappa$  decreases, C will switch into war (with an increased  $g_C^*$ ). However, we characterize this as the “internal efficiency constraint” binding, because C is optimizing based on their benefits of limited challenges and their costs of sometimes going to war. We do not consider this as a case of the “deterrent threat” influencing C’s b because C here is willing to risk war; rather, we conjecture this is better characterized as the

## War, Deterrent Threat Bind

First, because  $g_C = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $g_C$  is decreasing as  $\kappa$  decreases. When C remains in this equilibrium case, here the deterrent threat is entirely shaping C’s selected  $g_C^*$ , and  $g_C^*$  is decreasing in  $\kappa$ .

Consider what happens when  $\kappa$  decreases conditional on the initial parameters being in the **”War, Deterrent Threat Bind”** parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ , and  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ , are unchanged or continue to hold. We can show constraint  $\omega(\theta\rho - \kappa_C) + (1 - \omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  will continue to hold. We define placeholder function  $A = \omega(\theta\rho - \kappa_C) + (1 - \omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 - \theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) + \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ . We can then take first order conditions, showing

$$\begin{aligned}\frac{dA}{d\kappa} &= (1 - \omega)\theta - 2\beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right) - \theta + 2\beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right) \\ &= -\omega\theta - 2\beta_C\psi\end{aligned}$$

This is always negative. Therefore, as  $\kappa$  decreases, the the inequality is strengthened. This means the equilibrium will remain in the War, Deterrent Threat Bind case, where  $g_C^*$  is decreasing. Thus here, where C’s choices are influenced by the external deterrent threat,  $g_C^*$  is decreasing in  $\kappa$ .

Based on the above, we only need to consider the effect of  $\kappa$  decreasing on constraint  $\omega(\theta\rho - \kappa_C) + (1 - \omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 > \theta\rho - \kappa_C$ . We define placeholder  $B = \omega(\theta\rho - \kappa_C) + (1 - \omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 - (\theta\rho - \kappa_C)$ . We can then take first order conditions, showing

$$\frac{dB}{d\kappa} = (1 - \omega)\theta - 2\beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)$$

We can divide through by  $2\beta_C$  giving

$$\text{sign}\left\{\frac{dB}{d\kappa}\right\} = \frac{(1 - \omega)\theta}{2\beta_C} - \left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)$$

By the conditions of the case (specifically  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ), the first order condition with respect to  $\kappa$  is positive, meaning as  $\kappa$  decreases, the inequality could grow weaker or break. If the inequality weakens and does not break, then  $g_C = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$  is decreasing in  $\kappa$ —deterrence shapes  $g_C^*$  here. If the inequality does break and  $\omega(\theta\rho - \kappa_C) + (1 - \omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 \leq \theta\rho - \kappa_C$ , then the conditions match Status Quo 5, where  $g_C^* = 0$ . By switching to the status quo, this would mean that C is completely deterred from selecting any challenge—thus,  $g_C^*$  has decreased due to the deterrent threat.

Within this equilibrium space, because C is selecting the  $g_C^*$  tailored to make a type  $\bar{\kappa}_D$  indifferent between war and gray zone conflict (as to avoid war with type  $\bar{\kappa}_D$  D’s), here the deterrent threat constraint binds and the limited challenge is decreasing as  $\kappa$  decreases. Alternatively, it could be the case that as  $\kappa$  decreases, C is fully deterred and selects into the status quo (setting  $g_c^* = 0$ ).

## War, Internal Efficiency Binds

First, because  $g_C = \frac{(1-\omega)\theta}{2\beta_C}$ , the limited challenge is unchanging in  $\kappa$ . When C remains in this equilibrium case, here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is unchanging in  $\kappa$ .

Consider what happens when  $\kappa$  decreases conditional on the initial parameters being in the "War, Internal Efficiency Binds" parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}$  and  $\omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 > \theta\rho - \kappa_C$  are unchanged or continue to hold. We can show constraint  $\omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right) - \beta_C\left(\frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right)^2$  will continue to hold. We define placeholder  $E = \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 - \theta\left(\rho - \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right) + \beta_C\left(\frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right)^2$ . We can then take first order conditions, showing

$$\frac{dE}{d\kappa} = -\theta + 2\beta_C\left(\frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right)$$

We can divide through by  $2\beta_C$  giving

$$\text{sign}\left\{\frac{dE}{d\kappa}\right\} = -\frac{\theta}{2\beta_C} + \left(\frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right)$$

By the conditions of the case (specifically  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}$ ), the first order condition with respect to  $\kappa$  is negative. Therefore, as  $\kappa$  decreases, the inequality is strengthened. This means the equilibrium will remain in the War, Internal Efficiency Binds case, where  $g_C^*$  is unchanging. Thus here, where C's choices are influenced by C's internal efficiency,  $g_C^*$  is unchanging in  $\kappa$ .

Based on the above, We only need to consider the effect of  $\kappa$  decreasing on the constraint  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}$ . As  $\kappa$  decreases, this inequality will either be weakened or break. If the inequality is weakened, then the level of limited challenge is unchanging in  $\kappa$ . If decreasing  $\kappa$  breaks the inequality, then  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}$ . Consider what occurs when  $\kappa$  decreases to the point where  $\frac{(1-\omega)\theta}{2\beta_C} = \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}$ . Based on the calculus immediately above with placeholder  $E$  (i.e. decreasing  $\kappa$  preserves in inequality), we can substitute in using the expression  $\frac{(1-\omega)\theta}{2\beta_C} = \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}$  and say

$$\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right) - \beta_C\left(\frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right) - \beta_C\left(\frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right)^2$$

Based on work surrounding placeholder  $A$ , we know that decreasing  $\kappa$  will strengthen this inequality. Therefore, if decreasing  $\kappa$  results in  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}$ , then the conditions match War, Deterrent Threat binds. Because there  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}$ , C will be selecting a lower limited challenge following the change in  $\kappa$ . Additionally, as discussed above, once within the War, Deterrent Threat Binds space, the level of limited challenge is decreasing as  $\kappa$  decreases, or could decrease to the point where  $\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right) - \beta_C\left(\frac{1}{4\beta_D} + \frac{\kappa_C}{\theta}\right)^2 \leq \theta\rho - \kappa_C$ , which will result in  $g_C^* = 0$ . Together, this means that when  $g_C^*$  is decreasing in  $\kappa$ , it is because deterrence shapes  $g_C^*$  here. And, changes in  $\kappa$  could result in dynamics illustrated above in the discussion of the War, Deterrent Threat Binds Case (i.e.  $g_C^*$  could be further decreasing in  $\kappa$ ). When this is the case, it is because deterrence has come to shape C's choices.

Within the equilibrium space, because C is selecting the  $g_C^*$  based on their own internal cost benefit balance (while risking war), here the internal efficiency constraint binds and  $g_C^*$  is unchanging in  $\kappa$ . It could also be

that  $\kappa$  decreases to the point where C still risks war and the external deterrent threat starts to bind; when this is the case, C's selected  $g_C^*$  will decrease. Finally, as  $\kappa$  decreases, it could be the case that C is fully deterrence and selects into the status quo (setting  $g_C^* = 0$ ).

### 1.5.1 Gray Zone, Deterrent Threat Binds 3

First, because  $g_C = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ ,  $g_C$  is decreasing as  $\kappa$  decreases. When C remains in this equilibrium case, here the deterrent threat is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is decreasing in  $\kappa$ .

Consider what happens when  $\kappa$  decreases conditional on the initial parameters being in the “Gray Zone, Deterrent Threat Binds 3” parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , and  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ , are unchanged or continue to hold. This means the equilibrium will remain in the Gray Zone, Deterrent Threat Binds 3 case, where  $g_C^*$  is decreasing. Thus here, where C's choices are influenced by the external deterrent threat,  $g_C^*$  is decreasing in  $\kappa$ .

We need to consider the effect of  $\kappa$  decreasing on constraint  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$ . We define placeholder  $F = \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2 - (\theta\rho - \kappa_C)$ . We can take first order conditions, showing

$$\frac{dF}{d\kappa} = \theta - 2\beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)$$

We can divide through by  $2\beta_C$  giving

$$\text{sign}\left\{\frac{dF}{d\kappa}\right\} = \frac{\theta}{2\beta_C} - \left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)$$

By the conditions of the case ( $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ ), the first order condition with respect to  $\kappa$  is positive, meaning as  $\kappa$  decreases, the inequality could grow weaker or break.

Additionally, in the analysis of placeholder function  $A$ , we know that  $\theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2 > \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  will also grow weaker or break as  $\kappa$  decreases. If  $\kappa$  decreases and no inequality breaks, then  $g_C^* = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$  will be decreasing. If  $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$  and the second inequality still holds, then by transitivity the case will be Status Quo 5 and C will select  $g_C^* = 0$ . By switching to the status quo, this would mean that C is completely deterred from selecting any challenge—thus,  $g_C^*$  has decreased due to the deterrent threat.

If  $\theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2 \leq \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  and the first inequality still holds, then the case will be War 2 and C will select  $g_C = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ . This is the special case described in the discussion before Observation 1 and above (where we describe verbally what we do here); in this case, D's deterrent threat isn't truly “holding” because C is willing to switch into war. If neither inequality holds, then depending upon what is larger  $(\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  or  $\theta\rho - \kappa_C$ ), the case will be Status Quo 5 or War, Deterrent Threat Binds. In all cases, either  $g_C^*$  is decreasing (when the external deterrent threat influences C's actions), or C switches from gray zone conflict to some outcome with war (when internal efficiency influences C's decision).

Within this equilibrium space, because C is selecting the  $g_C^*$  tailored to make a type  $\kappa_D$  indifferent between war and gray zone conflict (as to avoid war all the time), here the deterrent threat constraint binds and



the limited challenge is decreasing as  $\kappa$  decreases. Also, it could be the case that as  $\kappa$  decreases, C is fully deterred and selects into the status quo (setting  $g_C^* = 0$ ). It could be the case that as  $\kappa$  decreases, C will switch to war and select a greater  $g_C^*$ ; when this is happening, the deterrent threat constraint is not binding, and rather C's internal efficiency (that depends on C's willingness to risk war) binds.

Note that it is possible that C could switch to war and this could result in a lower selected  $g_C^*$ ; under some parameters, if  $\kappa$  decreased enough, it could be that the external deterrent from type  $\bar{\kappa}_D$  D's will result in C selecting a  $g_C^*$  below the initial selected limited challenge within the Gray Zone, Deterrent Threat Binds 3 space, but that still results in war. While this case is unusual, it is also clear here that type  $\bar{\kappa}_D$ 's external deterrent threat is binding here (thus resulting in the lower  $g_C^*$ ).

### 1.5.2 Gray Zone, Deterrent Threat Binds 2

First, because  $g_C = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $g_C$  is decreasing as  $\kappa$  decreases. When C remains in this equilibrium case, here the deterrent threat is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is decreasing in  $\kappa$ .

Consider what happens when  $\kappa$  decreases conditional on the initial parameters being in the "Gray Zone, Deterrent Threat Binds 2" parameter space. Constraints  $\frac{\theta}{2\beta_D} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , and  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$  are unchanged or continue to hold. Constraint  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$  will grow weaker or break as  $\kappa$  decreases. Based on the analysis of placeholder function F,  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  will grow weaker or break as  $\kappa$  increases. Also based on the analysis in placeholder function F,  $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2$  will grow weaker or break as  $\kappa$  decreases.<sup>8</sup>

If no constraint breaks, then  $g_C^* = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$  is decreasing in  $\kappa$ —here the deterrent threat is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is decreasing in  $\kappa$ . If only the second constraint breaks—i.e.  $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ —then the case will be Status Quo 4, and the selected  $g_C = 0$ . By switching to the status quo, this would mean that C is completely deterred from selecting any challenge—thus,  $g_C^*$  has decreased due to the deterrent threat. If only the third constraint breaks—i.e.  $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 < \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2$ —then the case will be War, Internal Efficiency Binds. One of two things could happen here. Either  $g_C^*$  has decreased—which is because (while war is happening more) the external deterrent threat has succeeded in shrinking C's selected  $g_C^*$ . Otherwise,  $g_C^*$  could have increased, meaning this is a case where internal efficiency shapes C's choices. If both the second and third constraint break, then depending upon what is larger— $\theta\rho - \kappa_C$  or  $\omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2$ —the case will be Status Quo 4 or War, Internal Efficiency Binds. These switches will result in changes to  $g_C$  consistent with Observations 1 and 2.

Finally, consider the constraint  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ . As  $\kappa$  decreases, this inequality will either be weakened or break. If the inequality breaks, we cannot use the existing parameters to say much here, and how the game plays out is based on an alternate set of conditions.<sup>9</sup> Ultimately, the game will end in either War, External Deterrent Threat Binds, Status Quo 4, or Gray Zone, Deterrent Threat Binds 3. If the case switches to Gray Zone, Deterrent Threat Binds 3, the the level of limited challenge takes on the same form  $g_C^* = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$  which is decreasing in  $\kappa$ . When C switches to a different gray zone equilibrium or the status quo,  $g_C^*$  will be decreasing in  $\kappa$ —this is happening because C's choices are shaped by deterrence. When C switches to sometimes risking war,  $g_C^*$  may be increasing in  $\kappa$ , and this is driven by C's internal

<sup>8</sup>The analysis of placeholder F shows that  $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  is decreasing as  $\kappa$  decreases; the right hand side of this inequality is unchanging in  $\kappa$ .

<sup>9</sup>Specifically, how, with the new  $\kappa$ ,  $\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ ,  $\theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , and  $\theta\rho - \kappa_C$  relate to one another.

cost-benefit analysis.

Within this equilibrium space, because C is selecting the  $g_C^*$  tailored to make a type  $\kappa_D$  indifferent between war and gray zone conflict (as to avoid war all the time), here the deterrent threat constraint binds and the limited challenge is decreasing as  $\kappa$  decreases. It could also be the that  $\kappa$  decreases to the point where C the external deterrent threat binds then becomes too constraining, resulting in C being fully deterred and switching into the status quo. It could also be that C is constrained by the shift in  $\kappa$  and switches to war; when  $g_C^*$  is increasing following the shift, it is because C's internal efficiency constraint binds.<sup>10</sup> When  $g_C^*$  is decreasing following the shift, it is because the external deterrent threat from  $\bar{\kappa}_D$  binds.

### 1.5.3 Gray Zone, Deterrent Threat Binds 1

First, because  $g_C = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ ,  $g_C$  is decreasing as  $\kappa$  decreases. When C remains in this equilibrium case, here the deterrent threat is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is decreasing in  $\kappa$ .

Consider what happens when  $\kappa$  decreases conditional on the initial parameters being in the “Gray Zone, Deterrent Threat Binds 1” parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$  and  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$  are unchanged or continue to hold. Constraints  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$  and  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$  will grow weaker or break. If no constraint breaks, then  $g_C^* = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$  is decreasing in  $\kappa$ . If only the second constraint breaks—i.e.  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$ —then the case will be Status Quo 3. By switching to the status quo, this would mean that C is completely deterred from selecting any challenge—thus,  $g_C^*$  has decreased due to the deterrent threat.

Finally, we need to consider the effect of  $\kappa$  decreasing on the constraint  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ . As  $\kappa$  decreases, this inequality will either be weakened or break. If the inequality breaks, we cannot use the existing parameters to say much here, and, based on a new set of conditions,<sup>11</sup> the game will end in either War 2 or 1, Status Quo 5 or 4, or Gray Zone, Deterrent Threat Binds 2 or 3. When C switches to a different gray zone equilibrium or the status quo,  $g_C^*$  will be decreasing in  $\kappa$ —this is happening because C's choices are shaped by deterrence. When C switches to sometimes risking war,  $g_C^*$  may be increasing in  $\kappa$ , and this is driven by C's internal cost-benefit analysis.

Within this equilibrium space, because C is selecting the  $g_C^*$  tailored to make a type  $\kappa_D$  indifferent between war and gray zone conflict (as to avoid war all the time), here the deterrent threat constraint binds and the limited challenge is decreasing as  $\kappa$  decreases. It could also be the that  $\kappa$  decreases to the point where C the external deterrent threat binds then becomes too constraining, resulting in C being fully deterred and switching into the status quo. It could also be that C is constrained by the shift in  $\kappa$  and switches to war; when  $g_C^*$  is increasing following the shift, it is because C's internal efficiency constraint binds. When  $g_C^*$  is decreasing following the shift, it is because the external deterrent threat from  $\bar{\kappa}_D$  binds.

### 1.5.4 Gray Zone, Internal Efficiency Binds:

First, because  $g_C = \frac{\theta}{2\beta_C}$ ,  $g_C$  is unchanging as  $\kappa$  decreases. When C remains in this equilibrium case, here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is unchanging in  $\kappa$ .

Consider what happens when  $\kappa$  decreases conditional on the initial parameters being in the “Gray Zone, Internal Constraint Binds” parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$  and  $\theta\rho - \kappa_C < \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$  are unchanged. The constraint  $\frac{\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$  will grow weaker or break. If it does not break, then  $g_C^* = \frac{\theta}{2\beta_C}$  is unchanging in  $\kappa$ . If it breaks—i.e.  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ —then we cannot use the existing parameters to say much here, and, based on a new set of conditions, the game will end in either War 2 or 1, Status Quo 5, 4, or 3 or Gray Zone, Deterrent Threat Binds 3, 2, or 1. Note that within the gray zone equilibria, the level of limited challenge is decreasing as  $g_C^* = \frac{\theta}{2\beta_C}$  switches to  $g_C^* = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , which is

<sup>10</sup>This is discussed more in the Gray Zone, Deterrent Threat Binds 3 discussion.

<sup>11</sup>Specifically, how, with the new  $\kappa$ ,  $\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ ,  $\theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$ , and  $\theta\rho - \kappa_C$  relate to one another.

shrinking based on the broken inequality. Ultimately, when C switches to a different gray zone equilibrium or the status quo,  $g_C^*$  will be decreasing in  $\kappa$ —this is happening because C's choices are shaped by deterrence. When C switches to sometimes risking war,  $g_C^*$  may be increasing in  $\kappa$ , and this is driven by C's internal cost-benefit analysis.

Within the equilibrium space, because C is selecting the  $g_C^*$  based on their own internal cost benefit balance, here the internal efficiency constraint binds and  $g_C^*$  is unchanging in  $\kappa$ . It could also be the case that as  $\kappa$  decreases, the external deterrent threat could start to bind; if this is the case, C will decrease their selected  $g_C^*$ . It could also be the that  $\kappa$  decreases to the point where C the external deterrent threat binds then becomes too constraining, resulting in C being fully deterred and switching into the status quo. It could also be that C is constrained by the shift in  $\kappa$  and switches to war; when  $g_C^*$  is increasing following the shift, it is because C's internal efficiency constraint binds. When  $g_C^*$  is decreasing following the shift, it is because the external deterrent threat from  $\bar{\kappa}_D$  binds.

### 1.5.5 Status Quo 5

First, because  $g_C^* = 0$ ,  $g_C^*$  is unchanging as  $\kappa$  decreases. Here C's costs from selecting a greater  $g_C^*$  are too great, despite the deterrent threat not binding at  $g_C = 0$ —when C remains in this equilibrium case, here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is unchanging in  $\kappa$ .

Consider what happens when  $\kappa$  decreases conditional on the initial parameters being in the “Status Quo 5” parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , and  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ , are unchanged or continue to hold. Based on analysis of Function F and Function B the constraints  $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  and  $\theta\rho - \kappa_C \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  will continue to hold.

Altogether, this implies that as  $\kappa$  decreases, C will not switch out of selecting  $g_C^* = 0$ . Here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is unchanging in  $\kappa$ .

### 1.5.6 Status Quo 4

First, because  $g_C^* = 0$ ,  $g_C^*$  is unchanging as  $\kappa$  decreases. Here C's costs from selecting a greater  $g_C^*$  are too great, despite the deterrent threat not binding at  $g_C = 0$ —when C remains in this equilibrium case, here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is unchanging in  $\kappa$ .

Consider what happens when  $\kappa$  decreases conditional on the initial parameters being in the “Status Quo 4” parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , and  $\theta\rho - \kappa_C \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2$  are unchanged or continue to hold. Additionally, based on analysis of Function B, the constraints  $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  will continue to hold.

The constraint  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ , will grow weaker or break. When the constraint breaks, if C wanted to go to war, C will switch over to playing  $\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ . Consider when  $\kappa$  has decreased to the point where  $\frac{(1-\omega)\theta}{2\beta_C} = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ . Because all the inequalities are strengthening or unchanging as  $\kappa$  decreases, we can say

$$\theta\rho - \kappa_C \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$$

Based on analysis of the F function, this will continue to hold or will be strengthened as  $\kappa$  decreases, meaning  $\kappa$  decreasing could make the case Status Quo 5, where  $g_C = 0$  still.

Altogether, this implies that as  $\kappa$  decreases, C will not switch out of selecting  $g_C^* = 0$ . Here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is unchanging in  $\kappa$ .

### 1.5.7 Status Quo 3

First, because  $g_C^* = 0$ ,  $g_C^*$  is unchanging as  $\kappa$  decreases. Here C's costs from selecting a greater  $g_C^*$  are too great, despite the deterrent threat not binding at  $g_C = 0$ —when C remains in this equilibrium case, here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is unchanging in  $\kappa$ .

Consider what happens when  $\kappa$  decreases conditional on the initial parameters being in the “Status Quo 3” parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$  and  $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  are unchanged or continue to hold. The constraint  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , will grow weaker or could break. However, before  $\kappa$  decreases, we are able to say that because  $\theta\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  and  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , then C selecting some  $g_C < \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$  will result in a lower utility, or

$$\theta\rho - \kappa_C > \theta\left(\rho + \left(\frac{(1-\omega)\theta}{2\beta_C}\right) - \frac{1}{2\beta_D}\right) - \kappa_C - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2$$

Then, if decreasing  $\kappa$  results in  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , it will be the case that selecting  $g_C = \frac{(1-\omega)\theta}{2\beta_C}$  results in war. However, due to the inequality immediately above, we can say that

$$\theta\rho - \kappa_C \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2,$$

which implies that decreasing  $\kappa$  will result in C still preferring to select into the status quo (i.e. C will select into Status Quo 4). Finally, the above assumed that  $\frac{(1-\omega)\theta}{2\beta_C} < \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ; if this inequality broke, then it follows C would also do worse selecting  $g_C = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$  relative to  $g_C = 0$  (i.e. it will be the Status Quo 5 case).

Altogether, this implies that as  $\kappa$  decreases, C will not switch out of selecting  $g_C^* = 0$ . Here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is unchanging in  $\kappa$ .

### 1.5.8 Status Quo 2

First, because  $g_C^* = 0$ ,  $g_C^*$  is unchanging as  $\kappa$  decreases. Here C's costs from selecting a greater  $g_C^*$  are too great, despite the deterrent threat not binding at  $g_C = 0$ —when C remains in this equilibrium case, here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is unchanging in  $\kappa$ .

Consider what happens when  $\kappa$  decreases conditional on the initial parameters being in the “Status Quo 2” parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$  and  $\theta\rho - \kappa_C \geq \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$  are strengthened or unchanged. Constraint  $\frac{\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$  could break as  $\kappa$  decreases. Consider what occurs when  $\kappa$  decreases to the point where  $\frac{(1-\omega)\theta}{2\beta_C} = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ . We can use inequality  $\theta\rho - \kappa_C \geq \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$  to say that

$$\rho - \kappa_C \geq \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2.$$

The right hand side is decreasing as  $\kappa$  decrease, based on the analysis of placeholder function F. This implies that the case could shift to Status Quo 3, but not to any case where C increases they selected limited challenge.

Here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is unchanging in  $\kappa$ .

### 1.5.9 Status Quo 1

The constraints are unchanging in  $\kappa$ .

## 1.6 Proving Observations 3 and 4

**Observation 3:** *If increasing the challenger's resolve ( $\theta$ ) 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.*

To prove these, we will analyze the effects of changing  $\theta$  on C's selected challenge. We will show that whenever there is a positive relationship between C's resolve and C's challenge, then internal efficiency is shaping C's choices. Alternative, whenever there is a negative relationship between C's resolve and C's challenge, then the external deterrent threat is shaping C's challenge. Part of this is analyzing how changes in  $\theta$  affect play within a given equilibrium type. However, there can also be jumps, as changes in  $\theta$  break constraints resulting in a new equilibrium type.

It is worthwhile going into a little more detail on what Observations 3 and 4 are showing. Note that because the deterrent threat and internal efficiency generate opposite effects on C's selected  $g_C^*$ , it can be the case that the external deterrent threat binds at some point in the shift from  $\theta$  to  $\theta'$ , but that the challenge increases. However, importantly, what we can say is that whenever the challenge is increasing in  $\theta$ , this means that C's internal efficiency constraints played a role.

Note that similar to the proof of Observation 1, we use placeholder functions below; all placeholder function labels (function A, function B, etc) are describing functions within Observations 3 and 4, not referencing previous functions in the proof of Observations 1 and 2.

### 1.6.1 War, Deterrent Threat Binds

First, because  $g_C = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $g_C$  is decreasing as  $\theta$  increases. When C remains in this equilibrium case, here the deterrent threat is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is decreasing in  $\theta$ .

Consider what happens when  $\theta$  increases conditional on the initial parameters being in the "War, Deterrent Threat Binds" parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ , and  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$  are unchanged or continue to hold. We can show constraint  $\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 > \theta\rho - \kappa_C$  will continue to hold. We define placeholder function  $A = \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 - (\theta\rho - \kappa_C)$ . We can then take first order conditions, showing

$$\frac{dA}{d\theta} = (1-\omega)\left(-\frac{1}{4\beta_D} + \bar{\kappa}_D\right) + 2\beta_C \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \bar{\kappa}_D\right) + \frac{\kappa_C}{\theta}\right)}{\theta^2}$$

We can divide through by  $2\beta_C$  giving

$$\text{sign}\left\{\frac{dA}{d\theta}\right\} = \frac{(1-\omega)\theta}{2\beta_C} \left(-\frac{1}{4\beta_D} + \bar{\kappa}_D\right) + \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \bar{\kappa}_D\right) + \frac{\kappa_C}{\theta}\right)}{\theta}$$

If  $-\frac{1}{4\beta_D} + \bar{\kappa}_D$  is positive, this always holds. Suppose that it is negative. Because  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$  (by the conditions of the case), we can say

$$\text{sign}\left\{\frac{dA}{d\theta}\right\} \geq \left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right) \left(-\frac{1}{4\beta_D} + \bar{\kappa}_D\right) + \frac{\kappa_C}{\theta} \left(\left(\frac{1}{4\beta_D} + \bar{\kappa}_D\right) + \frac{\kappa_C}{\theta}\right)$$

Or re-writing as

$$\text{sign}\left\{\frac{dA}{d\theta}\right\} \geq \left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right) \left(-\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right).$$

Note that  $-\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta} \geq 0$  by the initial assumptions in the text, making the full expression positive. Thus, the constraint  $\omega(\theta\rho - \kappa_C) + (1 - \omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 > \theta\rho - \kappa_C$  is strengthening in  $\theta$ . This means the equilibrium will remain in the case, where  $g_C^*$  is decreasing. Thus here, where C's choices remain influenced by the external deterrent threat,  $g_C^*$  is decreasing  $\theta$ .

Next, we can also show that constraint  $\omega(\theta\rho - \kappa_C) + (1 - \omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$  will continue to hold. Let  $B = \omega(\theta\rho - \kappa_C) + (1 - \omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 - \left(\theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2\right)$ . We can then take first order conditions, showing

$$\frac{dB}{d\theta} = \rho + (1 - \omega)\left(-\frac{1}{4\beta_D} + \bar{\kappa}_D\right) + 2\beta_C \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \bar{\kappa}_D\right)\theta + \kappa_C\right)}{\theta^3} - \rho - \left(-\frac{1}{4\beta_D} + \kappa_D\right) - 2\beta_C \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \kappa_D\right)\theta + \kappa_C\right)}{\theta^3}$$

or

$$\frac{dB}{d\theta} = (1 - \omega)\left(-\frac{1}{4\beta_D} + \bar{\kappa}_D\right) + 2\beta_C \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \bar{\kappa}_D\right)\theta + \kappa_C\right)}{\theta^3} - \left(-\frac{1}{4\beta_D} + \kappa_D\right) - 2\beta_C \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \kappa_D\right)\theta + \kappa_C\right)}{\theta^3}$$

We can divide through by  $2\beta_C$  giving

$$\text{sign}\left\{\frac{dB}{d\theta}\right\} = \frac{\theta}{2\beta_C} \left( (1 - \omega)\left(-\frac{1}{4\beta_D} + \bar{\kappa}_D\right) + \frac{1}{4\beta_D} - \kappa_D \right) + \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \bar{\kappa}_D\right)\theta + \kappa_C\right)}{\theta} - \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \kappa_D\right)\theta + \kappa_C\right)}{\theta}$$

The right hand side will be positive whenever  $(1 - \omega)\left(-\frac{1}{4\beta_D} + \bar{\kappa}_D\right) + \frac{1}{4\beta_D} - \kappa_D \geq 0$ , which is a starting assumption. Therefore, the constraint will be strengthened. This means the equilibrium will remain in the case where the deterrent threat binds, and where  $g_C^*$  is decreasing.

Within this equilibrium space, because C is selecting the  $g_C^*$  tailored to make a type  $\bar{\kappa}_D$  indifferent between war and gray zone conflict (as to avoid going to war all the time), here the deterrent threat constraint binds and the limited challenge is decreasing as  $\theta$  increases. Also, as  $\theta$  increases, C will remain in this equilibrium space.

### 1.6.2 War, Internal Efficiency Binds

First, because  $g_C = \frac{(1-\omega)\theta}{2\beta_C}$ ,  $g_C$  is increasing as  $\theta$  increases. When C remains in this equilibrium case, here C's internal efficiency is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is increasing in  $\theta$ .

Consider what happens when  $\theta$  increases conditional on the initial parameters being in the "War, Internal Efficiency Binds" parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , and  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , are unchanged or continue to hold. We can show constraint  $\omega(\theta\rho - \kappa_C) + (1 - \omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 > \theta\rho - \kappa_C$  will continue to hold. We define placeholder  $E = \omega(\theta\rho - \kappa_C) + (1 - \omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 - \theta\rho + \kappa_C$ . We can then take first order conditions, showing

$$\frac{dE}{d\theta} = (1 - \omega)\left(\frac{(1 - \omega)\theta}{\beta_C} - \frac{1}{2\beta_D}\right) - \theta\left(\frac{(1 - \omega)^2}{2\beta_C}\right)$$

or

$$\frac{dE}{d\theta} = (1 - \omega)\left(\frac{(1 - \omega)\theta}{2\beta_C} - \frac{1}{2\beta_D}\right)$$

Because  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , we can say

$$\frac{dE}{d\theta} \geq (1 - \omega)\left(-\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)$$

By the initial assumptions in the text, the right hand side is positive. This implies that constraint  $\omega(\theta\rho - \kappa_C) + (1 - \omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 > \theta\rho - \kappa_C$  will continue to hold as  $\theta$  increases. This means the equilibrium will remain in the case, where  $g_C^*$  is increasing. Thus here, where C's choices remain influenced by C's internal efficiency,  $g_C^*$  is increasing in  $\theta$ .

We can also show constraint  $\omega(\theta\rho - \kappa_C) + (1 - \omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$  will continue to hold. We define function  $F = \omega(\theta\rho - \kappa_C) + (1 - \omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 - \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) + \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$ . We can then take first order conditions, showing

$$\frac{dF}{d\theta} = \frac{(1-\omega)^2\theta}{\beta_C} - \frac{(1-\omega)}{2\beta_D} - \frac{(1-\omega)^2\theta}{2\beta_C} + \frac{1}{4\beta_D} - \kappa_D - 2\beta_C \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \kappa_D\right)\theta + \kappa_C\right)}{\theta^3}$$

or

$$\frac{dF}{d\theta} = \frac{(1-\omega)^2\theta}{2\beta_C} - \frac{(1-\omega)}{2\beta_D} - \left(-\frac{1}{4\beta_D} + \kappa_D\right) - 2\beta_C \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \kappa_D\right) + \frac{\kappa_C}{\theta}\right)}{\theta^2}$$

As the next, step, we use that  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ .

$$\frac{dF}{d\theta} \geq \frac{(1-\omega)^2\theta}{\beta_C} - \frac{(1-\omega)}{2\beta_D} - \left(-\frac{1}{4\beta_D} + \kappa_D\right) - 2\beta_C \frac{\kappa_C \left(\frac{(1-\omega)\theta}{2\beta_C}\right)}{\theta^2}$$

which can be simplified to

$$\frac{dF}{d\theta} \geq (1-\omega) \left( \frac{(1-\omega)\theta}{\beta_C} - \frac{1}{2\beta_D} - \frac{\kappa_C}{\theta} \right) + \frac{1}{4\beta_D} - \kappa_D$$

From here, we can use the initial conditions of the case. The case assumes  $\omega(\theta\rho - \kappa_C) + (1 - \omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$ . Thus, it must be that  $(1 - \omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) > \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right)$ ,<sup>12</sup> or (re-writing) that  $(1 - \omega)\left(\frac{(1-\omega)\theta}{2\beta_C} - \frac{1}{2\beta_D} - \frac{\kappa_C}{\theta}\right) > -\frac{1}{4\beta_D} + \kappa_D$ . Therefore, the above inequality holds, which means the constraint grows stronger as  $\theta$  increases. This means the equilibrium will remain in the case, where  $g_C^*$  is increasing. Thus here, where C's choices remain influenced by C's internal efficiency,  $g_C^*$  is increasing in  $\theta$ .

The constraint  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$  will grow weaker or could break as  $\theta$  increases. If the constraint does not break, then because  $g_C^* = \frac{(1-\omega)\theta}{2\beta_C}$ , C is selecting a greater level of limited challenge as  $\theta$  increases. If the constraint breaks, then we can still show that C would never revert to the status quo or to gray zone conflict, instead playing as defined in the War, Deterrent Threat Binds space. Based on the analysis of placeholder function E, because the inequality  $\omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 > \theta\rho - \kappa_C$  is strengthened as  $\theta$  increases, at the point where  $\frac{(1-\omega)\theta}{2\beta_C} = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ , we can say

$$\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2 > \theta\rho - \kappa_C$$

As  $\theta$  increases, the constraint flips and  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ . This becomes just like the analysis of function A, which suggests the war becomes more desirable as  $\theta$  increases relative to peace. When C switches

<sup>12</sup>This is based on  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$ , which is due to  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ .

to the War, Deterrent Threat Binds case, the final challenge could be increasing or decreasing in  $\theta$ —but ultimately, the reason that this could go either way is because the internal efficiency and external deterrent threat both influence the final selected  $g_C^*$  in ways that Observations 3 and 4 predict.

We can run a similar analysis of inequality  $\omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2 > \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$ , again starting at the point where  $\frac{(1-\omega)\theta}{2\beta_C} = \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ . Based on the analysis of function B, war becomes more desirable to gray zone conflict as  $\theta$  increases.

Within the equilibrium space, because C is influenced by their own internal cost benefit balance, we can observe  $g_C^*$  increasing in  $\theta$ . It could also be the case that as  $\theta$  increases, C becomes more ambitious in their desired level of gray zone conflict, and the external deterrent threat could start to bind; if this is the case, C will start decreasing their selected  $g_C^*$ . Note that while these effects are in competition as  $\theta$  increases, this is because both effects are influencing the final selected  $g_C^*$ .

### 1.6.3 Gray Zone, Deterrent Threat Binds 3

First, because  $g_C = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ ,  $g_C$  is decreasing as  $\theta$  increases. When C remains in this equilibrium case, here the deterrent threat is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is decreasing in  $\theta$ .

Consider what happens when  $\theta$  increases conditional on the initial parameters being in the “Gray Zone, Deterrent Threat Binds 3” parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ ,  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ , will continue to hold. We can show constraint  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$  will continue to hold. We define placeholder  $G = \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2 - \theta\rho + \kappa_C$ . We can then take first order conditions, showing

$$\frac{dG}{d\theta} = \left(-\frac{1}{4\beta_D} + \kappa_D\right) + 2\beta_C \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \kappa_D\right) + \frac{\kappa_C}{\theta}\right)}{\theta^2}$$

We multiply through by  $\frac{\theta}{2\beta_C}$

$$\text{sign} \left\{ \frac{dG}{d\theta} \right\} = \frac{\theta}{2\beta_C} \left(-\frac{1}{4\beta_D} + \kappa_D\right) + \frac{\kappa_C \left(\left(\frac{1}{4\beta_D} + \kappa_D\right) + \frac{\kappa_C}{\theta}\right)}{\theta}$$

Because  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , we can say

$$\text{sign} \left\{ \frac{dG}{d\theta} \right\} \geq \left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right) \left(-\frac{1}{4\beta_D} + \kappa_D\right) + \frac{\kappa_C}{\theta} \left(\left(\frac{1}{4\beta_D} + \kappa_D\right) + \frac{\kappa_C}{\theta}\right) \geq \left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right) \left(-\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)$$

Which is always positive by the initial assumptions. Thus, the inequality is strengthened. This means the equilibrium will remain in the case, where  $g_C^*$  is decreasing. Thus here, where C's choices remain influenced by the external deterrent threat,  $g_C^*$  is decreasing  $\theta$ .

The constraint  $\theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2 \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  could be strengthened or weakened. Consider the case when  $(1-\omega)\left(\frac{(1-\omega)\theta}{2\beta_C} - \frac{1}{2\beta_D} - \frac{\kappa_C}{\theta}\right) \leq -\frac{1}{4\beta_D} + \kappa_D$ . When this holds, based on the analysis of function F, the initial constraint will continue to hold as  $\theta$  increases. When this happens,  $g_C^* = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , which implies the level of gray zone conflict will be decreasing. When  $(1-\omega)\left(\frac{(1-\omega)\theta}{2\beta_C} - \frac{1}{2\beta_D} - \frac{\kappa_C}{\theta}\right) > -\frac{1}{4\beta_D} + \kappa_D$ , then the constraint will be weakened via the



logic on the discussion of Function B. This implies that with higher resolve, C could switch to sometimes going to war. Thus, the final challenge could be increasing or decreasing in  $\theta$ —the reason that this could go either way is because the internal efficiency and external deterrent threat both influence the final selected  $g_C^*$  in ways that Observations 3 and 4 predict.

Within this equilibrium space, because C is selecting the  $g_C^*$  tailored to make a type  $\kappa_D$  indifferent between war and gray zone conflict (as to avoid war all the time), here the deterrent threat constraint binds and the limited challenge is decreasing in  $\theta$ . Also, it could be the case that as  $\theta$  increases, C internal efficiency binds and C becomes willing to risk war. When this happens, it could be that C's selected  $g_C^*$  is increasing in  $\theta$  (when the internal efficiency constraint binds). It could also be the case that  $\theta$  increases enough that the external deterrent threat (from type  $\bar{\kappa}_D$ ) binds and  $g_C^*$  begins decreasing. Ultimately, the reason that the effect on  $g_C^*$  this could go either way is because the internal efficiency and external deterrent threat both influence the final selected  $g_C^*$  in ways that Observations 3 and 4 predict.

### 1.6.4 Gray Zone, Deterrent Threat Binds 2

First, because  $g_C = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ ,  $g_C$  is decreasing as  $\theta$  increases. When C remains in this equilibrium case, here the deterrent threat is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is decreasing in  $\theta$ .

Consider what happens when  $\theta$  increases conditional on the initial parameters being in the “Gray Zone, Deterrent Threat Binds 2” parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ ,  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , and  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , will continue to hold. The constraint  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2$  will continue to hold by the same logic as the discussion on function  $G$ . This means the equilibrium will remain in the case, where  $g_C^*$  is decreasing. Thus here, where C's choices remain influenced by the external deterrent threat,  $g_C^*$  is decreasing  $\theta$ .

Either constraint  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$ ,  $\theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2 \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2$ , or both could break as  $\theta$  increases. First, if neither constraint breaks,  $g_C^* = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ , implying that the level of limited challenge is decreasing. If only constraint  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}$  breaks, then C's utility from going to war will become  $\omega(\theta\rho - \kappa_C) + (1-\omega)\theta\left(\rho - \frac{1}{4\beta_D} + \bar{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$ , meaning the second constraint is no longer relevant. Thus, the new relationship that matters is between C's new utility from going to war and C's utility from gray zone conflict where  $g_C = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ . Ultimately, this means C will either keep selecting into gray zone conflict (with  $g_C^*$  decreasing) or risk going to war (with  $g_C^*$  increasing). Similarly, if both constraints break, then once again C is choosing between gray zone conflict and war. Finally, if only  $\theta\left(\rho - \frac{1}{4\beta_D} + \kappa_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}\right)^2 \geq \omega(\theta\rho - \kappa_C) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2$  breaks, then C will risk war. Together, these effects are ambiguous, but ultimately, the reason that this could go either way is because the internal efficiency and external deterrent threat both influence the final selected  $g_C^*$  in ways that Observations 3 and 4 predict.

Within this equilibrium space, because C is selecting the  $g_C^*$  tailored to make a type  $\kappa_D$  indifferent between war and gray zone conflict (as to avoid war all the time), here the deterrent threat constraint binds and the limited challenge is decreasing in  $\theta$ . Also, it could be the case that as  $\theta$  increases, C internal efficiency binds and C becomes willing to risk war. When this happens, it could be that C's selected  $g_C^*$  is increasing in  $\theta$  (when the internal efficiency constraint binds). It could also be the case that  $\theta$  increases enough that the external deterrent threat (from type  $\bar{\kappa}_D$ ) binds and  $g_C^*$  begins decreasing.

### 1.6.5 Gray Zone, Deterrent Threat Binds 1

First, because  $g_C = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ ,  $g_C^*$  is decreasing as  $\theta$  increases. When C remains in this equilibrium case, here the deterrent threat is entirely shaping C's selected  $g_C^*$ , and  $g_C^*$  is decreasing in  $\theta$ .

Consider what happens when  $\theta$  increases conditional on the initial parameters being in the “Gray Zone, Deterrent Threat Binds 1’’ parameter space. Constraints  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$  and  $\frac{\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ , will continue to hold. The constraint  $\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$  will continue to hold by the same logic as the discussion on function  $G$ .

The constraint  $\frac{(1-\omega)\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$  could break as  $\theta$  increases. For the range where the constraint does not break,  $g_C^* = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$  implies that the level of limited challenge is decreasing in  $\theta$ . When this constraint does break, C may start being willing to go to war; ultimately, through the logic discussed above, if  $\theta$  grows to the point where  $\frac{(1-\omega)\theta}{2\beta_C} > \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$  and  $(1-\omega)\left(\frac{(1-\omega)\theta}{2\beta_C} - \frac{1}{2\beta_D} - \frac{\kappa_C}{\theta}\right) > -\frac{1}{4\beta_D} + \underline{\kappa}_D$ , then based on the discussion for function  $F$ , eventually war will be preferred. Thus, across all cases, as  $\theta$  increases, either gray zone conflict will continue (with  $g_C^*$  decreasing), or war will occur (with  $g_C^*$  increasing). Ultimately, the reason that this could go either way is because the internal efficiency and external deterrent threat both influence the final selected  $g_C^*$  in ways that Observations 3 and 4 predict.

Within this equilibrium space, because C is selecting the  $g_C^*$  tailored to make a type  $\underline{\kappa}_D$  indifferent between war and gray zone conflict (as to avoid war all the time), here the deterrent threat constraint binds and the limited challenge is decreasing in  $\theta$ . Also, it could be the case that as  $\theta$  increases, C internal efficiency binds and C becomes willing to risk war. When this happens, it could be that C’s selected  $g_C^*$  is increasing in  $\theta$  (when the internal efficiency constraint binds). It could also be the case that  $\theta$  increases enough that the external deterrent threat (from type  $\bar{\kappa}_D$ ) binds and  $g_C^*$  begins decreasing.

### 1.6.6 Gray Zone, Internal Constraint Binds

First, because  $g_C = \frac{\theta}{2\beta_C}$ ,  $g_C$  is increasing as  $\theta$  increases. When C remains in the case, C’s internal efficiency is entirely shaping C’s selected  $g_C^*$ , and  $g_C^*$  is increasing in  $\theta$ .

Consider what happens when  $\theta$  increases conditional on the initial parameters being in the “Gray Zone, Internal Constraint Binds’’ parameter space. Constraint  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$  will continue to hold. We can show constraint  $\theta\rho - \kappa_C < \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C$  will continue to hold. We define placeholder  $H = \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C - (\theta\rho - \kappa_C)$ . We can then take first order conditions, showing

$$\frac{dH}{d\theta} = \frac{\theta}{2\beta_C} - \frac{1}{2\beta_D}$$

Because  $\frac{\theta}{2\beta_C} > \frac{1}{2\beta_D}$ , this is always positive. This means the equilibrium will remain in the case, where  $g_C^*$  is increasing. Thus here, where C’s choices remain influenced by C’s internal efficiency,  $g_C^*$  is increasing in  $\theta$ .

The constraint  $\frac{\theta}{2\beta_C} \leq \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$  could break as  $\theta$  increases. For the range where it does not break  $g_C^* = \frac{\theta}{2\beta_C}$ , meaning the level of limited challenge is increasing in  $\theta$ . When this constraint does break, then to avoid war, the external deterrent constraint will bind. Consider  $\theta$  reducing to the point where  $\frac{\theta}{2\beta_C} = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ . From here, we are able to use the initial conditions and say

$$\theta\rho - \kappa_C < \theta\left(\rho - \frac{1}{4\beta_D} + \underline{\kappa}_D\right) - \beta_C\left(\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}\right)^2$$

Based on the discussion on function  $G$ , this constraint will be strengthened. Thus, as  $\theta$  increases, either gray zone conflict will continue, or war will occur. This means  $g_C^*$  could be increasing or decreasing in  $\theta$ —but ultimately, the reason that this could go either way is because the internal efficiency and external deterrent threat both influence the final selected  $g_C^*$  in ways that Observations 3 and 4 predict.

Within the equilibrium space, because C is selecting the  $g_C^*$  based on their own internal cost benefit balance, here the internal efficiency constraint binds and  $g_C^*$  is increasing in  $\theta$ . Also, it could be the case that as  $\theta$  increases, this constraint reaches the external deterrent threat; when this happens,  $g_C^*$  is decreasing in  $\theta$ . Alternatively,  $\theta$  could increase to the point where C’s internal efficiency binds and C becomes willing to risk war. When this happens, it could be that C’s selected  $g_C^*$  is increasing in  $\theta$  (when the internal efficiency constraint binds). It could also be the case that  $\theta$  increases enough that the external deterrent threat (from type  $\bar{\kappa}_D$ ) binds and  $g_C^*$  begins decreasing.

### 1.6.7 Status Quo 1-5

Based on the discussion above, as  $\theta$  increases, it becomes possible that C will switch out of selecting into the status quo and into gray zone conflict or war (with some greater level of limited challenge). Across any set of parameters, C becoming willing to engage in gray zone conflict or war is consistent with C's internal efficiency constraint binding. Because C cannot select some limited challenge strictly less than zero, increasing  $\theta$  will result in weakly greater limited challenges.

## 1.7 Observations 5-6

**Observation 5:** *If decreasing the challenger's challenge costs ( $\beta_C$ ) 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.*

The proof of Observations 5 and 6 takes a nearly identical form to the proof of Observation 3-4, only here within the external deterrent threat constraint the selected  $g_C^*$  is unchanging in  $\beta_C$  rather than within the external deterrent threat constraint where  $g_C^*$  is decreasing in  $\theta$ .

## 1.8 Discussion and Proof of Observation 7

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

To prove Observation 7, it suffices to give a numerical example, then offer a formal analysis of the role of  $\theta$ . Both are below.

### 1.8.1 General intuition

As intuition for this result, we can offer a short answer as to why improvements in a defender's limited response capabilities (D possessing a lower  $\beta_D$ ) will sometimes result in C taking actions that result in war. As  $\beta_D$  decreases, D fights harder within gray zone conflict. This makes gray zone conflict worse for the challenger. Now consider C who, essentially, is choosing between not challenging (accepting the status quo), a challenge that ends in gray zone conflict, and a challenge that risks war. Assume that for a high  $\beta_D$ , C selects a challenge that ends in gray zone conflict. If lowering  $\beta_D$  means C does worse in gray zone conflict, C may switch out of a challenge that always ends in gray zone conflict to either (a) a challenge that risks war or (b) not challenging. Essentially, D becoming better at gray zone conflict will incentivise C to avoid having the game always end in gray zone conflict and seek other options, where one of these options is risking war. And, when C is resolved enough, C will select into war.

We can offer a little more detail here. As  $\beta_D$  decreases, D fights harder within gray zone conflict. This constrains how productive gray zone conflict is for the challenger. Additionally, as  $\beta_D$  is lowered, C's wartime utility remains the same (we discuss how this result still goes through with relaxing this assumption in Model Extension 1). Together, this can incentivise C to gamble and risk war. Consider a setting where accepting the status quo is sufficiently bad for C.<sup>13</sup> In terms of the risk-return trade-off, C selects between two possible challenges, which generate three possible outcomes.<sup>14</sup> First, C can play it safe and choose a conservative challenge that always results in gray zone conflict and never in war (no risk). We refer to this as "Outcome 1." Second, C can select a more aggressive challenge (the riskier choice) that will either result in gray zone conflict ("Outcome 2") or result in war ("Outcome 3"). When C engages in a more aggressive challenge and D does not go to war (i.e. Outcome 2), C does better relative to C's payoff from the conservative challenge (high return). When C engages in a more aggressive challenge and D does go to war (Outcome 3), C does worse relative to C's payoff from the conservative challenge (the downside of the risk). As  $\beta_D$  decreases, all gray zone conflict outcomes become worse for C (i.e. Outcomes 1 and 2 become worse) while

<sup>13</sup>We can offer many parameter values where this is the case (and do so below).

<sup>14</sup>For the sake of parsimony, we are not discussing C selecting into the status quo here. We consider this in all the numerical discussions below.

$\beta_D = 1.5$ : C's Utility Outcomes					
Outcome	Status Quo	Always GZC	C Risks War	C Risks War, C Gets War	C Risks War, C Gets GZC
<b>C sets <math>g_C =</math></b>	0	$g_C^{GZ,EDT} = 0.4317$	$g_C^{W,EDT} = 0.4817$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.3333	-	-	0.3333
$g_C - g_D^*$	0	0.0983	-	-	0.1483
<b>C's Payoffs</b>	0.06	0.1583	0.149	0.06	0.2083

Table A3: C's choice in stage 2 and how the game plays out.

C's war payoffs remain the same (i.e. Outcome 3 remains the same). This means that the downside of a more aggressive challenge—the game ending in war—does not look as bad to C given C is doing worse playing it safe (i.e. Outcome 1 is worse).

To offer a simple numeric analogy: a risk-neutral actor would prefer a sure-thing payoff of \$6 to a 50%-50% gamble between \$0 and \$10. However, if all non-zero payoffs were reduced by \$4, then a risk-neutral actor would prefer a 50%-50% gamble between \$0 and \$6 to a sure-thing payoff of \$2. Returning to our game, as  $\beta_D$  decreases, C's gray zone payoffs are all decreasing (Outcomes 1 and 2 decrease, which is analogous to the non-zero payoffs decreasing by \$4), while the war payoff is not changing (Outcome 3 does not change, analogous to the zero payoff not changing).

### 1.8.2 The result for a set of parameters

We will now discuss how the game plays out for a fixed set of parameters. Then, we will lower  $\beta_D$  (while keeping all other parameters the same) and discuss how this change results in a greater likelihood of war. We offer more mathematical intuition below, but as  $\beta_D$  decreases, C is constrained and attains lower utilities from conducting gray zone conflict. And, as  $\beta_D$  decreases, C's wartime payoff in stage 3 ( $\theta\rho - \kappa_C$ ) does not decrease. This means that C becomes more willing to gamble on an outcome that risks war because the downside risk of war is relatively less bad.

Consider how the game plays out for parameters  $\rho = 0.1$ ,  $\kappa_C = 0.04$ ,  $\underline{\kappa}_D = 0.225$ ,  $\bar{\kappa}_D = 0.275$ ,  $\beta_D = 1.5$ ,  $\beta_C = 0$ ,  $\theta = 1$ , and  $\omega = 0.4$ . It is worthwhile highlighting one parameter in particular:  $\beta_C = 0$ . When this holds, then C faces no costs for conducting challenges, which means that the internal efficiency constraint (i.e. where C balances the costs and benefits of challenging) never binds.<sup>15</sup> Our choice to analyze a case where  $\beta_C = 0$  is done purely to help illustrate the the dynamics of how decreasing  $\beta_D$  can lead to a greater likelihood of war, and we relax this later. Because  $\beta_C = 0$ , under these parameters, C chooses between setting  $g_C = 0$ ,  $g_C = g_C^{GZ,EDT}$ , or  $g_C^{W,EDT}$  in stage 2 in the game. Table A3 summarizes, for these selections, various outcomes to the game.

Table A3 summarizes C's choice in stage 2 in the game. Moving left-to-right, the first two (non-bolded) columns (labeled "Status Quo" and "Always GZC") describe what happens when C selects  $g_C = 0$  and when C selects C's best  $g_C \in (0, g_C^{GZ,EDT}]$  (in other words, in the latter case, C selects their optimal  $g_C$  conditional on the game always ending in gray zone conflict). These two columns also describe D's equilibrium response  $g_D^*$  to the selected  $g_C$ , then reports the difference between  $g_C$  and  $g_D^*$  ( $g_C - g_D^*$ ), which is essentially how much additional (un-normalized) utility or surplus<sup>16</sup> C can get when the game ends in gray zone conflict. Finally, it describes C's expected payoff from the given  $g_C$ . Moving to the right—to the column labeled "C Risks War"—we do something similar, but we do not list  $g_D^*$  and  $g_C - g_D^*$ ; we do not do so because when C risks war, type  $\underline{\kappa}_D$  does not select a  $g_D^*$ ; rather, this type of D will go to war. However, we decompose what happens in the last two columns. These columns, labeled "C Risks War, C Gets War" and "C Risks War, C Gets GZC" describe what happens when C selects  $g_C = g_C^{W,EDT}$  and faces a type  $\underline{\kappa}_D$  that declares war, and when C selects  $g_C = g_C^{W,EDT}$  and faces a type  $\bar{\kappa}_D$  that conducts gray zone conflict (respectively). It

<sup>15</sup>Put another way, when  $\beta_C = 0$ , the only thing that constrains C is the threat of further war—in other words, the external deterrent threat constraint.

<sup>16</sup>When C selects  $g_C = g_C^{GZ,EDT}$ , the game ends in gray zone conflict, C is made offer  $x^* = \rho + g_C - g_D^* - \frac{\kappa_C}{\theta}$ . The  $g_C - g_D$  therefore represents C getting more than C would if C had selected  $g_C = 0$ .

$\beta_D = 1$ : C's Utility Outcomes					
Outcome	Status Quo	Always GZC	C Risks War	C Risks War, C Gets War	C Risks War, C Gets GZC
C sets $g_C =$	$g_C = 0$	$g_C^{GZ,EDT} = 0.515$	$g_C^{W,EDT} = 0.565$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.5	-	-	0.5
$g_C - g_D^*$	0	0.015	-	-	0.065
C's Payoffs	0.06	0.075	0.099	0.06	0.125

Table A4: C's choice in stage 2 and how the game plays out. Parameters are the same as Table A3 except here  $\beta_D = 1$ .

also describes the end-of-game payoffs for these cases. When D is type  $\underline{\kappa}_D$  and C selects  $g_C = g_C^{W,EDT}$ , C's final payoff is 0.06. And, when D is type  $\bar{\kappa}_D$  and C selects  $g_C = g_C^{W,EDT}$ , C's final payoff is 0.2083. Together, when C sets  $g_C = g_C^{W,EDT}$ , in expectation, C attains payoff 0.149.<sup>17</sup> Within this set of parameters, C does best setting  $g_C^* = g_C^{GZ,EDT}$ , meaning C never risks war here.

Next, we lower  $\beta_D$  by shifting  $\beta_D = 1.5$  to  $\beta_D = 1$ , while keeping all other parameters identical. Once again, C chooses between  $g_C = 0$ ,  $g_C = g_C^{GZ,EDT}$ , or  $g_C^{W,EDT}$ . Table A4, displays the new set of values for the new  $\beta_D$ .

Under the (lower)  $\beta_D = 1$ , C's preferred outcome is setting  $g_C = g_C^{W,EDT}$  and sometimes risking war. Why? In short, as D becomes better at conducting limited responses (i.e. as D has a lower  $\beta_D$ ), C does worse when the game ends in gray zone conflict. Because as  $\beta_D$  decreases D is willing to take on more costs to conducting gray zone conflict, this hinders C's ability to use limited challenges to gain an advantage.<sup>18</sup> This makes the gray zone conflict outcomes worse for C, which in turn makes risking war a (relatively) less-bad payout, thus incentivising C to gamble and risk war.

Consider the tables. When  $\beta_D = 1.5$ , if C sets  $g_C = g_C^{GZ,EDT}$  or  $g_C = g_C^{W,EDT}$ , then whenever D responds with a limited response, that limited response will be 0.3333. And, C does well when the game ends in gray zone conflict:  $U_C(g_C = g_C^{GZ,EDT}) = 0.1583$  and  $U_C(g_C = g_C^{W,EDT} | \kappa_D = \bar{\kappa}_D) = 0.2083$ . But, by setting  $g_C = g_C^{W,EDT}$ , C also risks war, which is (relatively) quite bad for C: when war occurs,  $U_C(g_C = g_C^{W,EDT} | \kappa_D = \underline{\kappa}_D) = 0.06$ . Thus here, when  $\beta_D$  is high, C does not find it worthwhile to risk going to war, and prefers getting the sure-thing payoff of  $U_C(g_C = g_C^{GZ,EDT}) = 0.1583$  to the expected payoff of  $U_C(g_C = g_C^{W,EDT}) = 0.149$ .

But, when  $\beta_D = 1$ , things change. If C sets  $g_C = g_C^{GZ,EDT}$  or  $g_C = g_C^{W,EDT}$ , then whenever D responds with a limited response, that limited response will be 0.5. Thus,  $g_C^{GZ,EDT}$  and  $g_C^{W,EDT}$  now do not go far beyond  $g_D^*$ . This is reflected in C's gray zone conflict payoffs:  $U_C(g_C = g_C^{GZ,EDT}) = 0.075$  and  $U_C(g_C = g_C^{W,EDT} | \kappa_D = \bar{\kappa}_D) = 0.125$ . With these payoffs, risking war is not nearly as (relatively) bad for C: under this new  $\beta_D$ , it still gives C  $U_C(g_C = g_C^{W,EDT} | \kappa_D = \underline{\kappa}_D) = 0.06$ . Thus here, because the downside to risking a more aggressive challenge—the downside of war—has gotten much closer to gray zone conflict payoffs, meaning the downside is not nearly as much of a deterrent. This can incentivise risk-taking.

This logic can also be expressed visually. For these parameters, Figure A2 displays C's expected utility on the y-axis and possible  $g_C$ 's on the x-axis. The top panel illustrates outcomes when  $\beta_D = 1.5$ , and the bottom panel illustrates this for  $\beta_D = 1$ . As a quick summary of why the top panel looks the way it does (we keep this brief because a similar figure was discussed above), first consider the relationship between  $g_C$  and  $U_C$  when  $g_C \leq \frac{1}{2\beta_D}$ . In this region, C's utility is flat because D's optimal  $g_D^* = g_C$ . This means that C's challenge is effectively being canceled out by D's limited response.<sup>19</sup> Note that C's utility is flat here because C faces no costs for conducting the challenge. In the  $g_C \in [\frac{1}{2\beta_D}, g_C^{GZ,EDT}]$  region, C's utility is increasing as

<sup>17</sup>When  $\omega = 0.4$  (as it is here),  $0.149 = \omega * 0.06 + (1 - \omega) * 0.2083$ .

<sup>18</sup>For the interested, we go into more detail on this point in the "In the Weeds" subsection below.

<sup>19</sup>While we do not display utilities for  $g_C \in [0, 0.3]$ , C's utility is static across the entire  $g_C \in [0, \frac{1}{2\beta_D}]$  region.

$g_C > g_D^* = \frac{1}{2\beta_D}$ . Here C's challenges are productive for future bargaining, and, because  $\beta_C = 0$ , C faces no costs for conducting these challenges. Together, in this region, greater  $g_C$ 's only add utility for C—this line is increasing, and the slope is constant. The discontinuity at  $g_C = g_C^{GZ,EDT}$  arises because at  $g_C = g_C^{GZ,EDT}$ , type  $\underline{\kappa}_D$  is indifferent between ending the game with gray zone conflict or war. Any  $g_C$  slightly to the right of  $g_C^{GZ,EDT}$  will result in a  $\omega$  likelihood of war without offering much added benefit should the game end in gray zone conflict, as it does with likelihood  $1 - \omega$ . However, in the  $(g_C^{GZ,EDT}, g_C^{W,EDT}]$  region, C's utility increases as C conducts more aggressive challenges. These more aggressive challenges still result in an  $\omega$  likelihood of war, but end in gray zone conflict with probability  $1 - \omega$ . Because there are no costs to C's challenges, the slope of this line is constant. Finally, there is another discontinuity at  $g_C^{W,EDT}$  because here both types of D will go to war for any  $g_C$  to the right of this value. C's payoffs in the "Always War" space is equal to C's Stage 3 wartime payoff  $(\theta\rho - \kappa_C)$  minus C's costs for challenging, which here are zero; thus, C's utility in this region is flat.

As we indicated in the numerical analysis above (discussed in Table A3), for this set of parameters, C does best never risking war and setting  $g_C = g_C^{GZ,EDT}$ . The return from risking war is not high enough here.

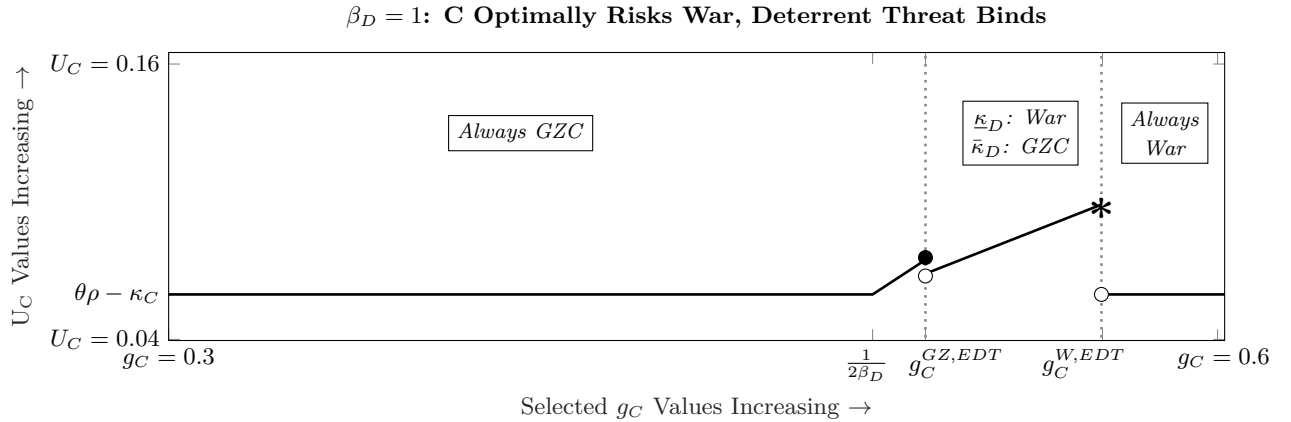
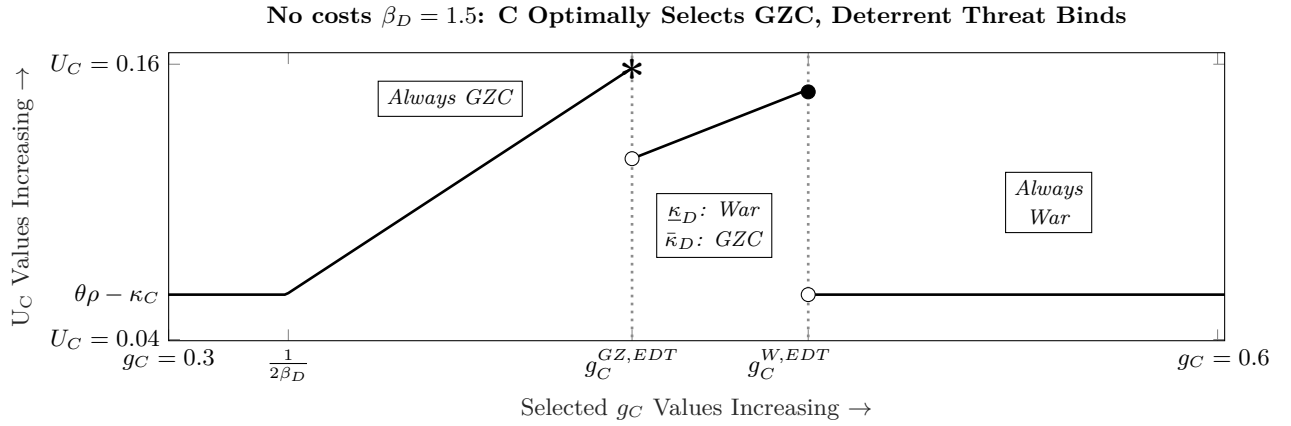


Figure A2: C's Utility Across selected  $g_C$ 's (War Outcomes).

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. The parameters are as follows:  $\rho = 0.1$ ,  $\kappa_C = 0.04$ ,  $\underline{\kappa}_D = 0.225$ ,  $\bar{\kappa}_D = 0.275$ ,  $\beta_C = 0$ ,  $\theta = 1$ , and  $\omega = 0.4$ .  $\beta_D$  varies across the top and bottom panels (as labeled).

The bottom panel similarly displays C's expected utility for the selected  $g_C$  values, only now  $\beta_D = 1$  (all

other parameters remain the same, as it was in Table A4). Under the new  $\beta_D$ , it's now optimal for C select  $g_C^{W,EDT}$ , meaning it's optimal for C to gamble on selecting a more aggressive challenge, knowing that it risks war when D is type  $\underline{\kappa}_D$ . What has changed? Essentially, gray zone conflict outcomes have gotten much worse for C. Under the new  $\beta_D$ , C is constrained in how far they can go beyond  $\frac{1}{2\beta_D}$  before triggering D to declare war. This makes gray zone conflict so much worse, in fact, that the “downside” of a more aggressive challenge—the game ending in war—does not look nearly as bad to the challenger given that the challenger is now doing much worse playing it safe and ending the game in gray zone conflict. This incentivises the challenger to undertake risks.

### 1.8.3 In the weeds

If questions still remain about the analysis above, we can go further still. This is not necessary to understanding the results, but it could be useful for readers very interested in the mechanics of the above. Why is gray zone conflict worse for C? As  $\beta_D$  decreases, essentially, C becomes more constrained in how aggressive their challenges can be. As mathematical intuition, consider a scenario where, after C selects a given  $g_C \geq \frac{1}{2\beta_D}$ , D does not want to go to war but does want to implement a limited response. Assuming everything is interior (i.e. the final  $P$  is not  $\rho$  or 1), D's optimization problem solves

$$g_D^* \in \argmax_{g_D \geq 0} \{1 - (\rho - g_D + g_C) + \kappa_C/\theta - \beta_D * g_D^2\}.$$

The solution here is  $g_D^* = \frac{1}{2\beta_D}$ . Thus, as  $\beta_D$  decreases (moving from  $\beta_D = 1.5$  to  $\beta_D = 1$ ), the selected  $g_D^*$  will increase (moving from  $g_D^*(\beta_D = 1.5) = 0.3333$  to  $g_D^*(\beta_D = 1) = 0.5$ ). And, as  $\beta_D$  decreases, the increase in  $g_D^*$  means that D will take on more costs to conducting limited responses. Through some quick calculations, when  $\beta_D = 1.5$  the expression  $-\beta_D * (g_D^*)^2 = -0.1667$ , and when  $\beta_D = 1$  the expression  $-\beta_D * (g_D^*)^2 = -0.25$ .

These added costs under the lower  $\beta_D$  matter due to how  $g_C^{GZ,EDT}$  was derived. Recall from above that  $g_C^{GZ,EDT}$  was the  $g_C$  that made type  $\underline{\kappa}_D$  indifferent between going to war in stage 3 ( $w_D = 1$ ), and selecting into gray zone conflict (setting  $w_D = 0$  and  $g_D = \frac{1}{2\beta_D}$ ) and letting the game proceed to stages 4 and 5. Formally,  $g_C^{GZ,EDT} = g_C$  when

$$1 - \left(\rho + g_C - \frac{1}{2\beta_D}\right) + \frac{\kappa_C}{\theta} - \beta_D \left(\frac{1}{2\beta_D}\right)^2 = 1 - \rho - \underline{\kappa}_D$$

holds. The left-hand-side is type  $\underline{\kappa}_D$ 's payoff from a gray zone conflict outcome, and the right-hand side is type  $\underline{\kappa}_D$ 's payoff from a war outcome. The  $g_C$  that makes the above hold with equality will be greater than or equal to  $\frac{1}{2\beta_D}$ .<sup>20</sup> This means that when C sets  $g_C = g_C^{GZ,EDT}$ , the selected  $g_C$  surpasses D's limited responses  $g_D^* = \frac{1}{2\beta_D}$  (effectively canceling out the benefit it gives D) and then further lowers the left-hand-side of the equality expression until it equals the right-hand side. How far C can go within gray zone conflict depends on the costs that D expects to accrue, which is the  $-\beta_D \left(\frac{1}{2\beta_D}\right)^2$  term. If  $\beta_D$  decreases (from 1.5 to 1) and C still selects  $g_C = g_C^{GZ,EDT}$  (now with the updated  $\beta_D$  value), C is still going to negate D's limited response  $g_D^* = \frac{1}{2\beta_D}$ , but because D has more costs from the lower  $\beta_D$  (because the  $-\beta_D \left(\frac{1}{2\beta_D}\right)^2$  term is a larger cost), C cannot go as far before hitting the  $1 - \rho - \underline{\kappa}_D$  war payoff. This constraints how aggressive C's challenges can be in going beyond D's limited response ( $g_D^*$ ), which results in C attaining a lower utility. And, this same logic holds for the  $g_C = g_C^{W,EDT}$  expression.

That C is more constrained in their challenges as  $\beta_D$  increases can be seen in two different ways. First, compare the  $g_C^* - g_D^*$  terms in Tables A3 and A4 where  $\beta_D = 1.5$  and  $\beta_D = 1$ . As  $\beta_D$  decreases, this difference shrinks in both the  $g_C = g_C^{GZ,EDT}$  column (going from 0.0983 to 0.015) and the  $g_C = g_C^{W,EDT} | \kappa_D = \bar{\kappa}_D$  column (going from 0.1483 to 0.065). Under  $\beta_D = 1$ , D has selected a greater limited response, which means that D has already accrued high costs for conducting these actions, which curtails C from engaging in more aggressive challenges, hence the smaller difference terms (smaller  $g_C^* - g_D^*$ ). D's accrued gray zone conflict costs in the  $\beta_D = 1$  case are higher than under  $\beta_D = 1.5$ , which means that C can get away with

<sup>20</sup>This is due to the  $-\frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta} \geq 0$  assumption. To see why, the equality can be simplified to  $-g_C + \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta} = 0$ .

more aggressive challenges under  $\beta_D = 1.5$ , which makes risking war a relatively worse prospect for C when  $\beta_D = 1.5$ . Second, this can be seen in the expression  $g_C = g_C^{GZ,EDT} = \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta}$ . As  $\beta_D$  decreases,  $g_C$  does increase (note the  $\frac{1}{4\beta_D}$  term), but not as fast as D's response  $g_D^* = \frac{1}{2\beta_D}$  does. When calculating the final  $P(\cdot)$ , we have  $P(g_D, g_C) = \rho + g_C - g_D$ , meaning the rates of change matter. And, a similar logic holds for  $g_C = g_C^{W,EDT}$ . As  $\beta_D$  decreases, gray zone conflict just isn't as productive for C, as  $g_C^{GZ,EDT}$  and  $g_C^{W,EDT}$  both get closer to  $g_D^*$ . This means the downside risk to moving from  $g_C^{GZ,EDT}$  to  $g_C^{W,EDT}$ —where the downside risk is war—becomes relatively less painful for C.

How does this all tie into Figure A2? Fix  $\beta_D = 1.5$ . When taking the weighted average of C's war payoff and C's gray zone conflict payoff when C sets  $g_C = g_C^{W,EDT}$ , C takes a big hit by risking war. This can, in part, be seen in the discontinuity at  $g_C = g_C^{GZ,EDT}$ , where there is a considerable gap between utilities. While C can gain back a lot of surplus moving up to  $g_C = g_C^{W,EDT}$ , this ultimately isn't enough to offset the initial discontinuity. In contrast, fix  $\beta_D = 1$ . When taking the weighted average of C's war payoff and C's gray zone conflict payoff when C sets  $g_C = g_C^{W,EDT}$ , C is taking a much smaller hit by risking war. The discontinuity at  $g_C = g_C^{GZ,EDT}$  has shrunk considerably because there just isn't that much of a difference between the gray zone conflict payoff and war. And, C can still gain back a lot of surplus moving to  $g_C = g_C^{W,EDT}$  (in fact, the same amount as under  $\beta_D = 1.5$  due to the third point above). As a result, C's utility from setting  $g_C = g_C^{W,EDT}$  and risking war surpasses C's utility from setting  $g_C = g_C^{GZ,EDT}$ . Overall, C is more willing to risk war because the gray zone payoffs have all gotten worse.

#### 1.8.4 General Mathematical Intuition

Above, we've provided a case where, as  $\beta_D$  decreases, C may be more willing to risk war. Here, we can go into more depth on the mathematical logic here (including considering what happens when  $\beta_C \neq 0$ ). It is the most straightforward to see how this works by comparing C's expected payoffs from setting  $g_C = g_C^{GZ,EDT}$  to  $g_C = g_C^{W,EDT}$ . Naturally other conditions must also hold—for example, it cannot be that  $U_C(g_C = 0) > U_C(g_C = g_C^{GZ,EDT})$  and  $U_C(g_C = 0) > U_C(g_C = g_C^{W,EDT})$ —but we focus on the logic behind the choice between never risking war ( $g_C = g_C^{GZ,EDT}$ ) and sometimes risking war ( $g_C = g_C^{W,EDT}$ ). Mathematically, consider the function  $Q$ , which is this difference.

$$Q = \theta \left( \rho - \frac{1}{4\beta_D} + \kappa_D \right) - \beta_C \left( \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta} \right)^2 - \left( (1-\omega)\theta \left( \rho - \frac{1}{4\beta_D} + \bar{\kappa}_D \right) + \omega(\theta\rho - \kappa_C) - \beta_C \left( \frac{1}{4\beta_D} + \bar{\kappa}_D + \frac{\kappa_C}{\theta} \right)^2 \right)$$

To see how parameter  $\beta_D$  affects C choice between  $g_C = g_C^{GZ,EDT}$  to  $g_C = g_C^{W,EDT}$ , we take first order conditions of  $Q$  with respect to  $\beta_D$ . Note, in doing so, we assume that changes in  $\beta_D$  do not result in a shift to C's internal efficiency binding. This gives

$$\frac{\partial Q}{\partial \beta_D} = \frac{\theta}{4\beta_D^2} + \frac{\beta_C(4\kappa_C\beta_D + 4\kappa_D\theta\beta_D + \theta)}{8\theta\beta_D^3} - \left( \frac{(1-\omega)\theta}{4\beta_D^2} + \frac{\beta_C(4\kappa_C\beta_D + 4\bar{\kappa}_D\theta\beta_D + \theta)}{8\theta\beta_D^3} \right)$$

which can be simplified to

$$\frac{\partial Q}{\partial \beta_D} = \frac{\theta}{4\beta_D^2}(1-\omega) + \beta_C \left( (\kappa_D - \bar{\kappa}_D) \frac{1}{2\theta\beta_D} \right)$$

The first part of this expression ( $\frac{\theta}{4\beta_D^2}(1-\omega)$ ) is always positive. The second part of this expression ( $\frac{\beta_C(\kappa_D - \bar{\kappa}_D)}{2\theta\beta_D}$ ) is always negative. Together, the joint effects are ambiguous. When the expression is positive, decreasing  $\beta_D$



could result in C switching from never risking war (setting  $g_C = g_C^{GZ, EDT}$ ) to sometimes risking war (setting  $g_C = g_C^{W, EDT}$ ), but never the reverse. And, when the expression is negative, decreasing  $\beta_D$  could result in C switching from sometimes risking war to never risking war, but never the reverse. Additionally, notice what happens when  $\theta$  is large; the first part of the expression is larger while the second part of the expression is smaller. This means the derivative is more positive (or closer to being positive), suggesting the effect of interest—that as  $\beta_D$  decreases C switches over to fighting more—will come about for a greater set of values for the other parameters when  $\theta$  is high.

It is also worthwhile noting: by no means does this result establish that decreasing  $\beta_D$  will always result in a greater likelihood of war (or vice-versa). As is common in games with incomplete information, comparative statics can often be counterintuitive when not considering the risk-return gamble that the uninformed party is undertaking.

We must also consider comparative statics for when the internal efficiency constraint binds when C sometimes risks war. Note that if the internal efficiency constraint binds when C never risks war, then it is never optimal for C to risk war.

$$Z = \theta \left( \rho - \frac{1}{4\beta_D} + \kappa_D \right) - \beta_C \left( \frac{1}{4\beta_D} + \kappa_D + \frac{\kappa_C}{\theta} \right)^2 - \left( \omega(\theta\rho - \kappa_C) + (1 - \omega) \left( \theta\rho + \frac{(1 - \omega)\theta^2}{2\beta_C} - \frac{\theta}{2\beta_D} - \kappa_C \right) - \beta_C \left( \frac{(1 - \omega)\theta}{2\beta_C} \right)^2 \right)$$

Taking first order conditions gives

$$\frac{\partial Z}{\partial \beta_D} = \frac{\theta}{4\beta_D^2} + \frac{\beta_C(4\kappa_C\beta_D + 4\kappa_D\theta\beta_D + \theta)}{8\theta\beta_D^3} - \left( (1 - \omega) \frac{\theta}{4\beta_D^2} \right)$$

Or simplified

$$\frac{\partial Z}{\partial \beta_D} = \omega \frac{\theta}{4\beta_D^2} + \frac{\beta_C(4\kappa_C)}{8\theta\beta_D^2} + \frac{\beta_C(4\kappa_D)}{8\beta_D^2} + \frac{\beta_C}{8\beta_D^3}$$

This value is always positive. This means whenever the internal efficiency constraint binds when C sometimes risks war, then decreases in  $\beta_D$  always results wanting to risk war.

We are unable to derive as clean as comparative statics on this condition as it relates to  $\theta$  because  $\frac{\partial Z}{\partial \beta_D}$  could be more or less positive as  $\theta$  increases. However, when  $\theta$  is high, then the internal efficiency constraint will not bind; rather, C will be deterred by the external deterrent threat. As a result, when  $\theta$  is high, the relevant analysis is on the  $Q$  expression rather than on the  $Z$  expression.

### 1.8.5 A Second Numerical Example

Readers might take objection to the above set of parameters where  $\beta_C = 0$ . Here we consider a second pair of parameters where decreasing  $\beta_D$  still results in C switching from never risking war to C sometimes risking war. Consider parameters  $\rho = 0.1$ ,  $\kappa_C = 0.04$ ,  $\kappa_D = 0.2$ ,  $\bar{\kappa}_D = 0.275$ ,  $\beta_D = 5$ ,  $\beta_C = 0.5$ ,  $\theta = 1.6$ , and  $\omega = 0.3$ . This is Figure A3.

In the top panel  $\beta_D = 5$  and C attains the highest expected payoff setting  $g_C = g_C^{GZ, EDT}$ , and in the bottom panel  $\beta_D = 1.5$  and C attains the highest expected payoff setting  $g_C = g_C^{W, EDT}$ . Once more, note the

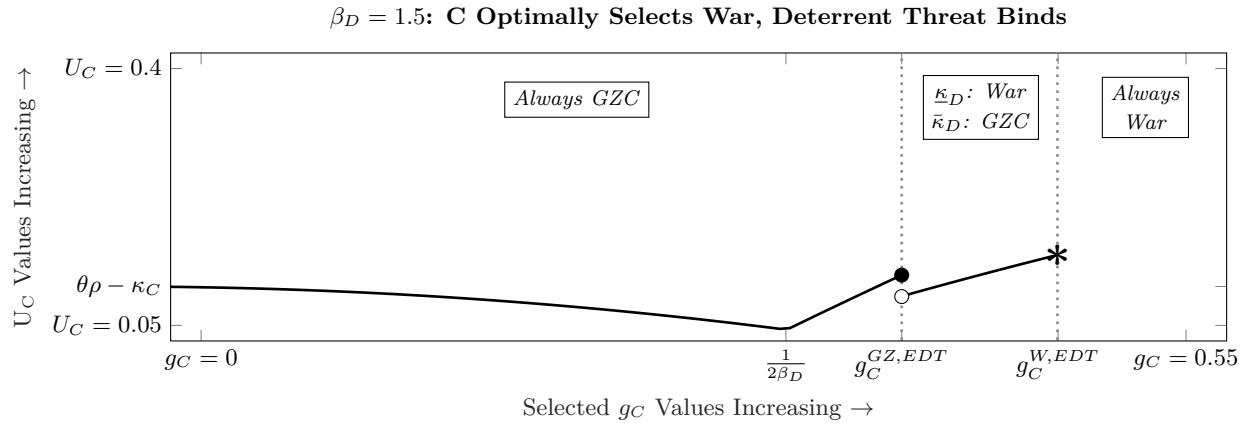
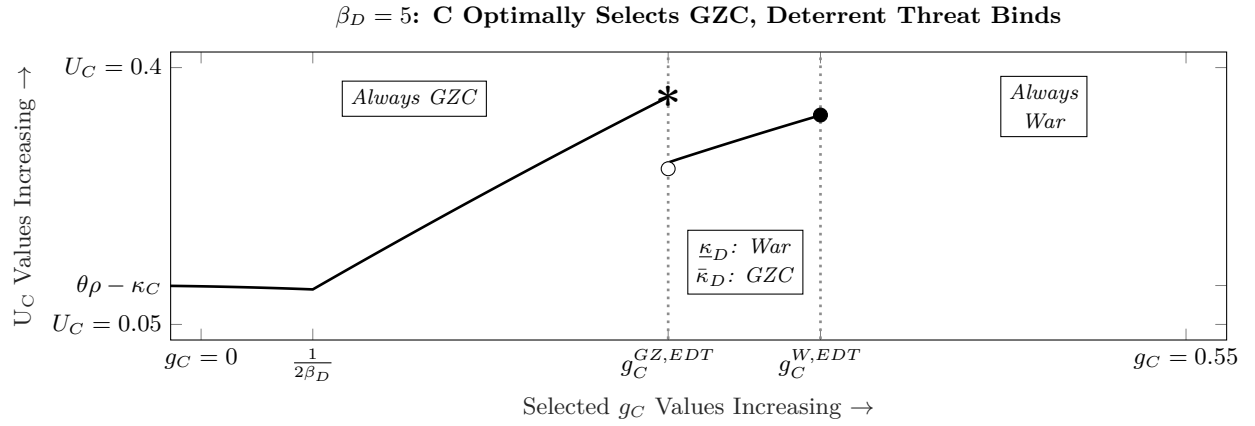


Figure A3: C's Utility Across selected  $g_C$ 's (War Outcomes).

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  $\beta_D$  which varies across sub-figures.

$\beta_D = 5$ : C's Utility Outcomes					
Outcome	Status Quo	Always GZC	C Risks War	C Risks War, C gets war	C Risks War, C gets GZC
<b>C Sets</b> $g_C =$	0	$g_C^{GZ,EDT} = 0.275$	$g_C^{W,EDT} = 0.35$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.1	-	-	0.1
$g_C - g_D^* =$	0	0.175	-	-	0.25
<b>C's Payoffs</b>	0.12	0.3622	0.3388	0.0588	0.4588

Table A5: C's choice in stage 2 and how the game plays out.

$\beta_D = 1.5$ : C's Utility Outcomes					
Outcome	Status Quo	Always GZC	C Risks War	C Risks War, C gets war	C Risks War, C gets GZC
<b>C Sets</b> $g_C =$	0	$g_C^{GZ,EDT} = 0.3917$	$g_C^{W,EDT} = 0.4667$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.333	-	-	0.333
$g_C - g_D^* =$	0	0.0583	-	-	0.1333
<b>C's Payoffs</b>	0.12	0.1366	0.1604	0.0111	0.2244

Table A6: C's choice in stage 2 and how the game plays out. Parameters are the same as they were in Table A6, only  $\beta_D$  is lower here.

difference in the discontinuity size at  $g_C = g_C^{W,EDT}$  across the two  $\beta_D$ 's. The shrinking discontinuity implies that C is not doing much by worse by gambling and risking war when  $\beta_D = 1.5$ .

And, for the sake of completeness, Tables A5 and A6 describe the gameplay as C chooses between  $g_C = 0$ ,  $g_C = g_C^{GZ,EDT}$ , and  $g_C = g_C^{W,EDT}$ .

## 2 Model Extension 1: Correlated Gray Zone and War Costs

### 2.1 Model Setup (Correlated Gray Zone/War Costs)

#### 2.1.1 Motivation

This is the model where D's gray zone conflict and war costs are correlated. As intuition, one might think that the technologies that can be used in war can also be used in gray zone conflict. This matters for when we analyze changes in D's gray zone conflict capabilities—after all, these might also influence D's ability to fight a war. For example, developing a relationship with a rebel group that operates in a rival's territory could be useful for both gray zone conflict or in war (more on this below).

We will allow for a flexibility in how D's gray zone and war payoffs are related—in other words, we will not assume a correlation of 1, as the technologies prioritized in war and gray zone conflict are often different. In the example of the rebel group, the time, effort, and costs put into that relationship may be more effective for gray zone conflict—where there is a premium on attacks that are not easily traced back to the backer state—than war. After all, deploying proxies lends itself to agency problems to an extent that deploying one's own army may not.

The aim of this section is to demonstrate that Observation 7 (or something like it—see Observation 8 below) can still hold when improvements in D's gray zone conflict capabilities also improves D's war capabilities. Because in our model we analyzed changes in D's gray zone conflict costs, here we will consider that D's gray zone conflict costs and D's costs of war are interrelated. Specifically, we will show that decreasing D's gray zone costs (decreasing  $\beta_D$ ) can still result in a greater likelihood of war, even when D's costs from war ( $\underline{\kappa}_D$  or  $\bar{\kappa}_D$ ) also decreases and C's costs of war ( $\kappa_C$ ) increases.

### 2.1.2 Why are Gray Zone Capabilities and War Capabilities Related but Not the Same?

From early 2014 until February 23, 2022, Russia conducted a series of actions in Ukraine that fell below the threshold of outright war. Russian activities in Ukraine at the time were (and still are) commonly referred to as “gray zone conflict,” which is consistent with how we are using the term. Then, on February 24, 2022, Russia invaded Ukraine and started a war.

Analyzing these two periods—the ‘gray zone conflict’ period (pre-invasion) and ‘war’ period (post-invasion)—gives us the chance to check the validity of one of the assumptions that we are making. Here we assume that the costs of gray zone conflict and the costs of war are related, but are not perfectly correlated. As one justification for this, we can look and see how the tools and capabilities used in gray zone conflict differ from the tools and capabilities used in war. If the costs of war were perfectly correlated with the costs of gray zone conflict, we’d expect gray zone conflict to look just like war, but a smaller, less intense version of war. On the other hand, if countries faced related but not perfectly correlated costs, then gray zone conflict would look different (in terms of capabilities and tools) than war. Put another way, if the costs of gray zone conflict and war were related (but not perfectly correlated), we would expect states to sometimes use different capabilities and face different challenges for conducting these two forms of conflict.

Below, we offer several data points on how Russia used different weapons systems, deployed different tactics, or faced distinct challenges across the two types of conflict (war and gray zone conflict). For example, we will first describe how Russian gray zone activities in the Donbas pre-invasion were essentially a proxy war that relied heavily on Ukrainian separatists. This is quite different from the war, which largely relied on Russian troops. Because predominantly using proxies is quite different than predominately using one’s own military forces, this suggests that Russia faced different structural costs to conducting gray zone conflict relative to conducting war. As another example, after invading Ukraine, Russia faced severe sanctions that were much worse than anything they faced during the period of gray zone conflict. Russia was able to withstand these sanctions, in part, because their central bank operated so quickly and so effectively. This means that having a competent central bank lowered Russia’s costs to conducting the war. In contrast, best we can tell, having a competent central bank did not play as significant a role in lowering the costs of gray zone conflict. And, while we focus on Russia below, there were also difference in Ukraine; the weapons systems given to Ukraine during the war—for example Ukraine now has networked anti-aircraft missile systems and long-range (like the NASAMS batteries), high-precision rocket artillery systems (like the HIMARS), which are very different than what NATO sent Ukraine during the period of gray zone conflict.<sup>21</sup>

The examples below should not be viewed as comprehensive—these are just several key examples that highlight that gray zone conflict and war have several distinctions, thus motivating the different cost structures.

**2.1.2.1 Proxy fighters/Ukrainian Separatists** In summary, while Russian-backed Ukrainian proxy forces in the Donbas played a significant role during the period of gray zone conflict, in the war, the Russian military ran (best we can tell) the major operations. Running a proxy war is quite different than using Russian military forces. As examples of how these generate different costs, we would anticipate different forms of agency problems across the techniques, getting proxies to operate properly is difficult, and soldier casualties are viewed differently by the Russian public across the two forms of conflict.

**Gray Zone Conflict Period:** Following the annexation of Crimea in February 2014, one of the most active locations of gray zone conflict between Russia and Ukraine was in the eastern Donbas region encompassing Donetsk and Luhansk. Local Russian-backed Ukrainian separatists acted as the primary fighting force against Ukrainian military forces in the region (“Conflict in Ukraine’s Donbas: A Visual Explainer” 2021). Although Russia intermittently denied doing so, there is evidence that Russia provided military weaponry, their own troops, and financial support to both separatist fighting forces and separatist governments, the Donetsk People’s Republic (DPR) and the Luhansk People’s Republic (LPR) (Vorobyov 2022b). For instance, investigators concluded that Ukrainian separatists were responsible for shooting down Malaysia Airlines Flight MH17 in July 2014 using a Buk mobile surface-to-air missile battery that Russian forces provided and that separatists fired Russian-supplied tanks, heavy weapons, and rocket launchers into Ukraine (Neuman 2020; Wroughton 2014). In July 2014, the estimated manpower of Ukrainian separatists was 10,000 (Foster

<sup>21</sup>For example, Ukraine was only given Javelin warheads near the end of the gray zone period.

and Goryachova 2014). In July 2015, the Ukrainian government claimed that there were 42,500 separatist fighters deployed, including 9,000 Russian soldiers (Vorobyov 2022a). These numbers remained relatively stable into 2017, when Ukraine claimed there were 46,000 separatist fighters deployed, including 6,000 Russian soldiers (Petrenko 2017). Accounts differ about the nature of Russian troops brought into separatist regions during gray zone conflict. Some stated that the troops consisted of Russian volunteers who chose to go aid and contribute manpower to Ukrainian separatist forces (“Ukraine Crisis: ‘Thousands of Russians’ Fighting in East” 2014), while other stated that Russian officers commanded the separatist units (“Russia and the Separatists in Eastern Ukraine” 2016). Additionally, there is some evidence of infighting among separatists, or even that Moscow was involved in the assassination of separatist leadership (Strzelecki, Iwanski, and Zochowski 2018).

**War Period:** In the 2022 conflict, Ukrainian separatists became a secondary fighting force. In June 2022, Russia declared that taking all of Donbas was its main objective in the invasion (Leicester and Arhinova 2022). Ukrainian separatists have been involved in this offense on the eastern front and fought alongside a full battalion of Russian troops to capture the city of Sieverodonetsk (Dress 2022). However, the fighting force is now predominantly the Russian military, with 150,000 troops deployed at the beginning of the invasion and up to 300,000 additional reservists likely to as part of Russia’s partial mobilization strategy (Ilyushina and Timsit 2022). The contribution of Ukrainian separatists seems to be largely limited to the borders of the Donbas region; they did not made up a significant fighting force in Russia’s early push toward Kyiv nor in fighting on the southern front.

**2.1.2.2 Manpower Shortages** In summary, during the pre-invasion period, we do not have evidence of Russian operations suffering from manpower shortages. In contrast, today, evidence suggests that an inability to conscript soldiers is inhibiting Russia’s war efforts and driving up costs. In terms of our model, Russia experienced costs by going to war that it did not experience in gray zone conflict; the political ramifications of raising and deploying a large fighting force.

**Gray Zone Conflict Period:** As discussed above on the use of proxy forces, during the gray zone conflict period, Russia primarily supported separatists in the Donbas region of Ukraine, which furthered their goals of broad disruption. Essentially, Russia deployed the manpower that was already willing to fight by providing arms, training, and direct military support to Ukrainian separatists. Ultimately, we find little evidence to suggest that Russia suffered from a lack of manpower during the 2014-2022 gray zone conflict period, given their heavy use of Ukrainian separatists.

**War Period:** Conscription has historically been Russia’s main source of manpower, but reporting suggests the 2022 spring conscription fell short by approximately 45,500 men. Russian defense minister Sergei Shoigu revealed in a statement that the spring conscription was a “de facto failure”. (Luzin 2022) To address the manpower shortage without formally declaring war on Ukraine or initiating a widespread draft, mobile enlistment offices have been popping up in several Russian cities and Russia is offering incentives to sign on to serve, especially reservists: those who have already completed their military service. According to an advertisement in Tula, a city 200 kilometers south of Moscow, those who opt-in for three-month contracts could earn over 170,000 rubles a month, which is four times the average local salary (Beardsworth 2022). However, even these incentives may not be enough – U.S. intelligence suggests that the Russian military is, aside from seeking to recruit contract service members, “may even draw in convicted criminals” (Holland 2022). In addition to recruitment, Russia is calling in reservists via a draft to supplement their forces. The approximate number of reservists drafted ranges from 200,000 to 300,000. It has been reported that “tens of thousands” of men have fled the country to avoid the draft, which continues even despite (now banned) anti-war protests (Reuters 2022) These reservists and conscripts that make up the bulk of the Russian force are not well trained, nor are they motivated to fight. Studies claim that “The rapid employment of relatively untrained reservists is unlikely to materially increase Russia’s combat power in Ukraine” (Stepanenko, Kagan, and Babcock-Lumish 2022). Of course, importantly, all of this is happening in the face of Russian losses in Ukraine.

Government policies on conscription, increased recruitment, and calling of reservists all suggest that Russia is facing manpower shortages. These difficulties only get exasperated by a lack of training and morale. This is not to mention the massive losses the Russian military has sustained after seven months of the war in

Ukraine, likely due to the low quality of their conscript-heavy troops. Estimates range from 70,000 to 80,000 casualties, which, when considering the struggles Russia has had with manpower, constitutes a major concern for the Russian military.

A study from the Institute for the Study of War outlines that Russia, though making a big push to increase men on the ground, "efforts to mobilize more manpower can bring more people into Russian combat units, but those people are unlikely to be well-enough trained or motivated to generate large amounts of new combat power" (Stepanenko, Kagan, and Babcock-Lumish 2022). The Department of Defense sustains that "adding 300,000 more troops to the mix is not going to make the situation better if you are already having significant challenges and haven't addressed some of those systemic strategic issues that make any large military force capable, there's nothing to indicate that it's going to get any easier by adding more variables to the equation" (Garamone 2022).

**2.1.2.3 Naval Forces** In summary, during the gray zone conflict period, Russia implemented new customs procedures for European and Ukrainian naval vessels traveling through the Kerch Strait, and, very infrequently, harassed vessels from these areas. After the invasion, Russia quickly implemented a blockade, which required considerable effort by Russia's Black Sea Fleet. Once more, this is an area where costs to gray zone conflict and war are related, but not perfectly correlated. Some harassment of European vessels is possible with limited capabilities; implementing a blockade required significant effort on the part of Russia, and, in implementing the blockade, Russia has suffered losses of major naval vessels.

**Gray Zone Conflict Period:** With the annexation of Crimea in 2014, Russia began exerting influence in the Kerch Strait, which is the only access point for Ukrainian ships attempting to access the major cities of Berdiansk and Mariupol along the Sea of Azov. In 2014, after reports of harassment from Russia, the European Union banned any vessels from the European Union from paying fees to now Russian controlled ports in Crimea. Ukrainian ports followed this up by prosecuting ship masters who had records of doing this as well as impounding the vessels. In 2017, Russia began construction of a bridge over the Kerch Strait which Ukraine claimed would block certain larger vessels from reaching its ports. While this claim appears valid, we have been unable to find recorded instances of ships being unable to reach ports in the Sea of Azov. Later, in May 2018, Russia implemented customs proceedings in the Strait which have reportedly increased travel times to Ukraine by more than 24 hours. During that summer, Berdiansk reported a 33% decrease in freight while Mariupol reported a 15% decrease in freight. Following that summer, in November 2018, three Ukrainian ships passing through the Kerch strait were rammed, fired upon, and later seized by Russian ships. However, it is unclear if the Ukrainian ships had followed customs proceedings (Schatz and Koval 2019). As of 2018, over 100 ships had been detained at the strait and by 2019 this number had increased to more than 150 (Atland and Kabanenko 2020).

**War Period:** At the breakout of the war, Russia implemented a blockade on the Kerch Strait allowing no traffic to and from Ukrainian ports. This blockade has been enforced by the Russian Black Sea Fleet. However, Ukraine has been able to keep this fleet at bay and sink the largest ship in the fleet, thus hurting the Russian navy's ability to support land troops (Miller and McLeary 2022). Additionally, Ukraine has not been able to freely export grain since the start of the war. Since the initiation of the blockade, the UN backed a deal called the "Black Sea Grain Initiative" to get grain out of Ukraine. The first ship was able to leave on the 1st of August and up until September 22nd 180 ships have been able to leave carrying 4.2 million metric tons of grain (Macias and Cortes 2022).

**2.1.2.4 Sanctions** In summary, while Russia was exposed to some sanctions after the 2014 invasion of Crimea and subsequent 2016 U.S. election interference, after invading Ukraine, Russia faced sanctions that were much worse than anything they faced during the period of gray zone conflict. However, Russia was able to withstand these severe sanctions, partly due to their role as an oil exporter and partly due to their quite effective central bank. This means that having a competent central bank and an economy that leans heavily on exporting oil lowered Russia's costs to conducting the war. In contrast, best we can tell, having a competent central bank did not play as significant a role in lowering the costs of gray zone conflict.

**Gray Zone Conflict Period:** When Russia annexed Crimea, the United States and European Union members along with other western countries implemented economic sanctions on Russia. The sanctions

include blocking state owned enterprises from accessing financial markets and exporting oil and military equipment to Russia. In 2014, Russia responded with banning all food imports from Western countries. After a year, NATO assessed that these sanctions had caused the devaluing of the Rouble as well as caused an exodus of capital. These sanctions at this time also caused increased food prices, led to inflation, and a decrease in GDP (Christie 2015). Then, in 2016, the United States implemented new measures after alleged election interference. These new sanctions blocked the sale of dual-use products as well as restricted more financial transactions. The European Union followed this with more sanctions against individuals and legal entities in 2015. Interestingly, Russian Public opinion hold that the sanctioning countries are more affected than the Russian economy (78%). Despite this, poverty has increased in Russia by 3% resulting in 4 million more Russians entering poverty. Furthermore another poll found that in regard to sanctions, “34% of the respondents noted a positive effect, 30% noted a negative effect, and 20% think that there are no consequences” (Klinova and Sidorova 2019).

As of 2020, the European Union imported nearly 40% of its natural gas from Russia. In 2019, a political analyst said that despite geopolitical situations there is an interdependent relationship that remains relatively stable. While during this period discussions took place about decreasing dependence on Russian gas, no substantial actions were taken (Moravcsik 2019).

**War Period:** With the invasion of Ukraine in February 2022, the United States and the European Union imposed new sanctions on Russia. These new sanctions arrived much faster than those in 2014 which took multiple months to agree upon. The new sanctions include blocking Russia from the system of SWIFT payments, blocking Russian access to bank assets in other countries which compromise a majority of Russian Capital, blocking the foreign assets of Duma members, decreasing exports to Russia of items vital to the war effort. Furthermore, Russia was blocked from receiving financing from the IMF, World Bank, and WTO. In regards to airspace, the U.S. and the E.U. have banned Russian planes and the U.S. has taken this a step further to ban Russian ships from its waters. On the private side, companies like Visa and MasterCard have left Russia along with other international tech companies who provide critical services. The United States ended importing Russian energy, however, European countries have not followed suit. Putin responded to these sanctions calling all of them “akin to a declaration of war” (Eichensehr 2022).

In response, Russia has focused their efforts differently and sold oil and gas to other countries and with rising prices actually earned 50% more than in 2021. Additionally, Russia has threatened to seize property of western businesses and infringe on patents and trademarks. In response to sanctions on government officials Russia responded with sanctions on U.S. government officials (Eichensehr 2022). As of September 2022, E.U. officials said they are prepared for a winter with limited Russian gas. While they highlighted they hope Russia will abide by contracts they are prepared for a scenario where Putin cuts the supply. However, this comes with energy saving measures being implemented in many countries such as Germany only heating public buildings to 66° Fahrenheit (Schuetze 2022).

Additionally, faced with the unprecedented sanctions and the disruption to Russian financial markets caused by Western efforts, the Russian Central Bank increased its interest rate from 9.5% to 20%. This move bolstered the ruble and, by April 2022, the Central Bank began to decrease interest rates. This trend continued and by July, the rate went down to 8%. Inflation in Russia remains fairly controlled: in May the rate was 17%, and in July the rate was 15.9%. Russia was so successful in this period with managing economic fallout that the rouble hit a seven year high against the U.S. Dollar (Nelson 2022).<sup>22</sup> However, Putin acknowledged that he was worried about the future of imports with specific regard to microchips and microelectronics. Despite some marks of strength and resiliency the economy has shrunk by 4% compared to the same period in 2021 and the bank estimates this could increase to 6%. However this is significantly less than originally predicted (Nelson 2022).

Looking forward, the Russian Central Bank has changed its outlook for inflation saying it now predicts prices will rise 11-13% in the year compared to the preceding expectation of 12-15%. In regards to interest rates, the Governor of the Central Bank noted that the most recent decrease in September 2022 would be followed

---

<sup>22</sup>This strength is likely due to a couple of reasons, increased energy prices and demand which brought an extra \$6 billion to the Russian Government and decreased spending from Russian consumers in possible anticipation of needed savings and also gives time to Russian companies to set up new methods for importing goods (Troianovski and Nechepurenko 2022).

by most likely a holding of the rate or an increase however she added a rate cut would not be impossible (Ostroukh and Marrow 2022). That said, overall, Russia seems to have weathered the sanctions quite well, at least in the short-term. Russian Central Bank officials also stated how the full effects of Sanctions may not be felt yet and effects could be long lasting. One specific example is the airline industry which is almost entirely dependent on parts from western nations and will likely feel the effects as planes are unable to be repaired and must be grounded (“Western Sanctions Will Eventually Impair Russia’s Economy” 2022).

**2.1.2.5 Rockets and cruise missiles** Before the March invasion, pro-Russian forces used unguided BM-21 “Grad” rockets in conflict zones in Eastern Ukraine (Lokshina 2017). These rockets, which are launched from the back of a combat vehicle, have a maximum range of 52km. In contrast, following the February 24 invasion, Russia has launched numerous cruise missiles, like the 9M727 Cruise Missile into cities in Western Ukraine like Kiev and Vinnytsia (Ostrovsky and Troyanovsky 2022; Gauthier-Villars et al. 2022). These cruise missiles have guidance systems, are very accurate, and have a range of 400-500km. In the first few months, Russia is estimated to have fired more than 3,650 missiles and guided rockets (Gauthier-Villars et al. 2022). We have been unable to find evidence of Russian cruise missile strikes in Ukraine when Russia was engaged in gray zone conflict. Additionally, for example, Ukraine now has networked anti-aircraft missile systems and long-range (like the NASAMS batteries) and high-precision rocket artillery systems (like the HIMARS), which it did not possess during the period of gray zone conflict (Ostrovsky and Troyanovsky 2022). Overall, we observe a change from lower-technology systems used during gray zone conflict to long-range systems that are deployed well beyond battle-lines during the war. All this is to suggest that Russia (and Ukraine) are deploying different capabilities during gray zone conflict and war, suggesting these different forms of conflict generate different costs.

### 2.1.3 Model Setup, Equilibrium, C’s Payoffs

The model setup is largely the same, except now capabilities are correlated. We modify D’s costs to low-level responses to now be  $\beta_D = \beta + \zeta$ , with  $\beta \geq 0$  and  $\zeta \geq 0$ . Within the analysis below, we will vary  $\zeta$ . We also modify D’s costs of war, and set  $\underline{\kappa}_D = \kappa + \alpha\zeta$  and  $\bar{\kappa}_D = \kappa + \alpha\zeta + \psi$ , with  $\kappa > 0$ ,  $\psi > 0$  and  $\alpha \geq 0$ . This means that as D becomes more effective at a limited response ( $\zeta$  decreases), D’s costs of war ( $\underline{\kappa}_D$  and  $\bar{\kappa}_D$ ) also decreases, but this change to D’s costs of war is moderated by  $\alpha$ . Additionally, we might also expect some connection between D’s limited response capabilities and C’s costs to go to war. We also set  $\kappa_C = \tau - \nu\zeta$ , where  $\tau > 0$ , and  $\nu \geq 0$ . The term  $\kappa_C$  (C’s costs from going to war) is decreasing in  $\zeta$  because we might anticipate that as D becomes better at conflict (lower  $\zeta$ ), C’s costs from conflict would increase (hence  $\kappa_C$  decreasing in  $\zeta$ ). These linkages across costs of war are a little atypical—in the analysis below, we show that we can still get the results of interest (decreases in  $\zeta$  can incentivize C to start risking war) when  $\nu = 0$ . We assume that the values are such that  $\beta_D \geq 0$ ,  $\underline{\kappa}_D > 0$ , and  $\kappa_C > 0$ .

Our new setup allows for the costs of war and the costs of gray zone conflict to be related. Minus a few notation changes, this setup matches the model in the main paper whenever  $\zeta = 0$  or  $\alpha = \tau = 0$ . We believe the version explored here better represents settings where gray zone conflict and war have some overlap in techniques and capabilities.

With this new structure, all assumptions and the equilibrium as stated in the paper remain the same. Put another way, when, in equilibrium, C never risks war and C’s external deterrent threat binds, C still sets  $g_C^* = g_C^{GZ,EDT} = \frac{1}{4\beta_D} + \underline{\kappa}_D + \frac{\kappa_C}{\theta}$ ; if we were being more complete, we can expand the analysis out and say that  $g_C^{GZ,EDT} = \frac{1}{4(\beta+\zeta)} + \kappa + \alpha\zeta + \frac{\tau-\nu\zeta}{\theta}$ . We do not list the full equilibrium here because this is covered in the main paper. To summarize C’s actions and utilities, we list below what actions C takes in Stage 2 and what C’s final payoffs are. C will select a challenge that results in the following:

*Sometimes War, Deterrent Threat Binds:* C sets  $g_C = g_C^{W,EDT} = \frac{1}{4(\beta+\zeta)} + (\kappa + \alpha\zeta + \psi) + \frac{\tau-\nu\zeta}{\theta}$ . C attains utility  $U_C = \omega(\theta\rho - (\tau - \nu\zeta)) + (1-\omega)\theta\left(\rho - \frac{1}{4(\beta+\zeta)} + (\kappa + \alpha\zeta + \psi)\right) - \beta_C\left(\frac{1}{4(\beta+\zeta)} + (\kappa + \alpha\zeta + \psi) + \frac{\tau-\nu\zeta}{\theta}\right)^2$

*Sometimes War, Internal Efficiency Binds:* C sets  $g_C = g_C^{W,IE} = \frac{(1-\omega)\theta}{2\beta_C}$ . C attains utility  $U_C = U_C(g_C = g_C^{W,IE} | g_C^{W,DTB} \geq g_C^{W,IE}) = \omega(\theta\rho - (\tau - \nu\zeta)) + (1-\omega)\left(\theta\rho + \frac{(1-\omega)\theta^2}{2\beta_C} - \frac{\theta}{2(\beta+\zeta)} - (\tau - \nu\zeta)\right) - \beta_C\left(\frac{(1-\omega)\theta}{2\beta_C}\right)^2$



*Always Gray Zone, Deterrent Threat Binds:* C sets  $g_C = g_C^{GZ,EDT} = \frac{1}{4(\beta+\zeta)} + (\kappa + \alpha\zeta) + \frac{\tau-\nu\zeta}{\theta}$ . C attains utility  $U_C = \theta \left( \rho - \frac{1}{4(\beta+\zeta)} + (\kappa + \alpha\zeta) \right) - \beta_C \left( \frac{1}{4(\beta+\zeta)} + (\kappa + \alpha\zeta) + \frac{\tau-\nu\zeta}{\theta} \right)^2$

*Always Gray Zone, Internal Efficiency Binds:* C sets  $g_C = g_C^{GZ,IE} = \frac{\theta}{2\beta_C}$ . C attains utility  $U_C = \theta\rho + \frac{\theta^2}{4\beta_C} - \frac{\theta}{2(\beta+\zeta)} - (\tau - \nu\zeta)$

*Status Quo:* C sets  $g_C = 0$ . C attains utility  $U_C = \theta\rho - (\tau - \nu\zeta)$ .

## 2.2 Model Extension 1: When Improvements in Gray Zone Capabilities Lead to More War

For this new model, we create a new Observation:

**Observation 8:** *Decreases in  $\zeta$  can lead the challenger to switch from always conducting gray zone conflict to sometimes risking war.*

Whenever  $\zeta$  decreases,  $\beta_D$  also decreases. This is consistent with D becoming better at gray zone conflict, as D has lower costs to conducting it. And, whenever  $\zeta$  decreases, both  $\underline{\kappa}_D$  and  $\bar{\kappa}_D$  also decreases. This is consistent with D becoming better at war because now D has lower costs to conducting war. And finally, whenever  $\zeta$  decreases,  $\kappa_C$  increases. This is consistent with D becoming worse at war.

To prove Observation 8, it suffices to show a set of parameters where this holds. We offer several parameters. First, the effects are the cleanest to see when  $\beta_C = 0$ . We then offer a parameter set where  $\beta_C > 0$ . Finally, we include one set of parameters where D's costs of gray zone conflict only influences D's costs of war ( $\nu = 0$ ).

### 2.2.1 Demonstrating Observation 8 when $\beta_C = 0$

We let  $\kappa = 0$ ,  $\alpha = 0.5$ ,  $\beta = 2$ ,  $\psi = 0.05$ ,  $\rho = 0.1$ ,  $\tau = 0.3$ ,  $\nu = 0.6$ ,  $\rho = 0.1$ ,  $\beta_C = 0$ ,  $\theta = 1.6$ , and  $\omega = 0.3$ . We first let  $\zeta = 0.4$ . This implies that  $\kappa_C = 0.04$ ,  $\underline{\kappa}_D = 0.2$ ,  $\bar{\kappa}_D = 0.25$ , and  $\beta_D = 2.4$ . C's utility outcomes for selecting between the status quo, always selecting into gray zone conflict, and sometimes risking war are listed below. Recall that when  $\beta_C = 0$ , C does best either setting  $g_C = 0$ ,  $g_C = g_C^{GZ,EDT}$ , or  $g_C = g_C^{W,EDT}$ . In other words, whenever  $\beta_C = 0$ , C's internal efficiency never binds.

$\zeta = 0.4$ : C's Utility Outcomes (Parameter Set 1)					
Outcome	Status Quo C Risks War,	Always GZC	C Risks War	C Risks War, C Gets War	C Risks War, C Gets GZC
C sets $g_C =$	$g_C = 0$	$g_C^{GZ,EDT} = 0.3417$	$g_C^{W,EDT} = 0.3917$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.2083	-	-	0.2083
$g_C - g_D^*$	0	0.1333	-	-	0.1833
C's Payoffs	0.1	0.3133	0.3053	0.1	0.3933

For this set of parameters, C does best setting  $g_C^* = g_C^{GZ,EDT}$  and always selecting into gray zone conflict. In other words, here C will never risk going to war.

We can now keep all other parameters constant, but lower  $\zeta$ . Now, consider how the game plays out when  $\zeta = 0.1$ . This now means that  $\kappa_C = 0.24$ ,  $\underline{\kappa}_D = 0.05$ ,  $\bar{\kappa}_D = 0.1$ , and  $\beta_D = 2.1$ . So, under this new  $\zeta$ , C's costs from war have increased and D's costs from war and low level conflict have decreased. The game plays out in the following way.

$\zeta = 0.1$ : C's Utility Outcomes (Parameter Set 1)					
Outcome	Status Quo C Risks War,	Always GZC	C Risks War	C Risks War, C Gets War	C Risks War, C Gets GZC
C sets $g_C =$	$g_C = 0$	$g_C^{GZ,EDT} = 0.319$	$g_C^{W,EDT} = 0.369$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.2381	-	-	0.2381
$g_C - g_D^*$	0	0.081	-	-	0.131
C's Payoffs	0	0.0495	0.0667	-0.08	0.1295

Now C does best setting  $g_C^* = g_C^{W,EDT}$ . Why? The logic is essentially the same as it was earlier, with a few new caveats. In short, as D becomes better at conducting limited responses (i.e. as D has a lower  $\zeta$ ), C does worse when the game ends in gray zone conflict. Because as  $\zeta$  decreases D is willing to take on more costs to conducting gray zone conflict, this hinders C's ability to use limited challenges to gain an advantage. Whereas under  $\zeta = 0.4$  C did well from the "Always GZC" challenge, C does much worse under "Always GZC" when  $\zeta = 0.1$ . And, while C's wartime payoff is getting worse (moving from  $U_C = 0.1$  to  $U_C = -0.08$ ), it's not getting worse fast enough to disincentivise C from not being willing to risk war under  $\zeta = 0.1$ . Essentially, moving from  $\zeta = 0.4$  to  $\zeta = 0.1$  means that all of C's conflict payoffs get worse, but under  $\zeta = 0.1$  war is not bad enough to disincentivise C from risking the greater challenge and getting war sometimes.

When  $\beta_C = 0$ , we can run some comparative statics on when C does better risking war. We define function  $M$  as the difference between C's payoff from risking war and C's payoff from not risking war.

$$M = \omega(\theta\rho - (\tau - \nu\zeta)) + (1 - \omega)\theta\left(\rho - \frac{1}{4(\beta + \zeta)} + (\kappa + \alpha\zeta + \psi)\right) - \theta\left(\rho - \frac{1}{4(\beta + \zeta)} + (\kappa + \alpha\zeta)\right)$$

Or more simply

$$M = \omega(\theta\rho - (\tau - \nu\zeta)) - \omega\theta\left(\rho - \frac{1}{4(\beta + \zeta)} + (\kappa + \alpha\zeta + \psi)\right) + \theta\psi$$

We can take first order conditions with respect to  $\zeta$

$$\frac{d}{d\zeta}M = \omega\nu\zeta - \theta\omega\left(\alpha + \frac{1}{4\beta\zeta^2}\right)$$

This expression could be positive or negative. The effect of interest is when, as D becomes better at low level conflict (decreasing  $\zeta$ , which also alters  $\kappa_C$ ,  $\underline{\kappa}_D$  and  $\bar{\kappa}_D$ ), C becomes more willing to risk war—in other words, when the derivative is negative. When  $\theta$  is high,  $\alpha$  is high, and  $\nu$  is low, the derivative is smaller or negative. This means that when  $\beta_C = 0$ , the stronger the link between D's low-level conflict costs and D's costs of war—which is  $\alpha$ , because  $\underline{\kappa}_D = \kappa + \alpha\zeta$  and  $\bar{\kappa}_D = \kappa + \alpha\zeta + \psi$ —under more conditions will we observe the relationship of note (decreasing  $\zeta$  leads C to switch to risking war from never risking war). And, when  $\beta_C = 0$ , the weaker the link between D's low-level conflict costs and C's costs of war—which is  $\nu$  because  $\kappa_C = \tau - \nu\zeta$ —under more conditions will we observe the relationship of note (decreasing  $\zeta$  leads C to switch to risking war from never risking war).

Of course, this analysis only holds when  $\beta_C = 0$  or, under the right conditions, when  $\beta_C$  is small enough. We do not include an analysis when  $\beta_C \neq 0$  because the derivative expressions become very convoluted.

### 2.2.2 Demonstration Observation 8

We naturally do not want to give the impression the result only holds when  $\beta_C = 0$ . It does not—the analysis in the previous subsection is there so that the effects can be most clearly observed. Now we will present a new set of parameters which illustrates, again, that as  $\zeta$  decreases—which leads to the aggregate effects of D becoming better at war and gray zone conflict and C becoming worse at war—we can still get the same effects.

We let  $\kappa = 0$ ,  $\alpha = 0.5$ ,  $\beta = 2$ ,  $\psi = 0.05$ ,  $\rho = 0.1$ ,  $\tau = 0.3$ ,  $\nu = 0.6$ ,  $\rho = 0.1$ ,  $\beta_C = 0.2$ ,  $\theta = 1.6$ , and  $\omega = 0.3$ . We first let  $\zeta = 0.45$ . This implies that  $\kappa_C = 0.03$ ,  $\underline{\kappa}_D = 0.225$ ,  $\bar{\kappa}_D = 0.275$ , and  $\beta_D = 2.45$ . C's utility outcomes for selecting between the status quo, always selecting into gray zone conflict, and sometimes risking war are listed below.

$\zeta = 0.45$ : C's Utility Outcomes (Parameter Set 2)					
Outcome	Status Quo C Risks War,	Always GZC	C Risks War	C Risks War, C Gets War	C Risks War, C Gets GZC
<b>C sets <math>g_C =</math></b>	$g_C = 0$	$g_C^{GZ,EDT} = 0.3458$	$g_C^{W,EDT} = 0.3917$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.2041	-	-	0.2041
$g_C - g_D^*$	0	0.1417	-	-	0.1917
<b>C's Payoffs</b>	0.13	0.3328	0.3134	0.0987	0.4054

For this set of parameters, C does best setting  $g_C^* = g_C^{GZ,EDT}$  and always selecting into gray zone conflict. In other words, here C will never risk going to war.

We can now keep all other parameters constant, but lower  $\zeta$ . Now, consider how the game plays out when  $\zeta = 0.1$ . This now means that  $\kappa_C = 0.24$ ,  $\underline{\kappa}_D = 0.05$ ,  $\bar{\kappa}_D = 0.1$ , and  $\beta_D = 2.1$ . So, under this new  $\zeta$ , C's costs from war have increased and D's costs from war and low level conflict have decreased. The game plays out in the following way.

$\zeta = 0.1$ : C's Utility Outcomes (Parameter Set 2)					
Outcome	Status Quo C Risks War,	Always GZC	C Risks War	C Risks War, C Gets War	C Risks War, C Gets GZC
C sets $g_C =$	$g_C = 0$	$g_C^{GZ,EDT} = 0.319$	$g_C^{W,EDT} = 0.369$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.2381	-	-	0.2381
$g_C - g_D^*$	0	0.081	-	-	0.131
C's Payoffs	0	0.0292	0.0394	-0.107	0.1023

Now C does best setting  $g_C^* = g_C^{W,EDT}$ . Once more, as  $\zeta$  decreases, C switches to sometimes risking war.

### 2.2.3 Demonstrating Observation 8 when $\nu = 0$

We let  $\kappa = 0$ ,  $\alpha = 0.5$ ,  $\beta = 2$ ,  $\psi = 0.05$ ,  $\rho = 0.1$ ,  $\tau = 0.3$ ,  $\nu = 0.6$ ,  $\rho = 0.1$ ,  $\beta_C = 0.2$ ,  $\theta = 1.6$ , and  $\omega = 0.3$ . We first let  $\zeta = 1$ . This implies that  $\kappa_C = 0.04$ ,  $\underline{\kappa}_D = 0.5$ ,  $\bar{\kappa}_D = 0.55$ , and  $\beta_D = 3$ . C's utility outcomes for selecting between the status quo, always selecting into gray zone conflict, and sometimes risking war are listed below.

$\zeta = 1$ : C's Utility Outcomes (Parameter Set 3)					
Outcome	Status Quo C Risks War,	Always GZC	C Risks War	C Risks War, C Gets War	C Risks War, C Gets GZC
C sets $g_C =$	$g_C = 0$	$g_C^{GZ,EDT} = 0.6083$	$g_C^{W,EDT} = 0.6583$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.1667	-	-	0.1667
$g_C - g_D^*$	0	0.4417	-	-	0.4917
C's Payoffs	0.12	0.7527	0.584	0.0333	0.82

For this set of parameters, C does best setting  $g_C^* = g_C^{GZ,EDT}$  and always selecting into gray zone conflict. In other words, here C will never risk going to war.

We can now lower  $\zeta$  to equal  $\zeta = 0.2$ . This in turn results in  $\kappa_C = 0.04$ ,  $\underline{\kappa}_D = 0.1$ ,  $\bar{\kappa}_D = 0.15$ , and  $\beta_D = 2.2$ . C's utility outcomes for selecting between the status quo, always selecting into gray zone conflict, and sometimes risking war are listed below.

$\zeta = 0.2$ : C's Utility Outcomes (Parameter Set 3)					
Outcome	Status Quo	Always GZC	C Risks War	C Risks War, C Gets War	C Risks War, C Gets GZC
C sets $g_C =$	$g_C = 0$	$g_C^{GZ,EDT} = 0.2386$	$g_C^{W,EDT} = 0.2886$	$g_C^{W,EDT}   \kappa_D = \underline{\kappa}_D$	$g_C^{W,EDT}   \kappa_D = \bar{\kappa}_D$
$g_D^* =$	0	0.2273	-	-	0.2273
$g_C - g_D^*$	0	0.0114	-	-	0.0614
C's Payoffs	0.12	0.1268	0.1721	0.1033	0.2015

As can be seen, C now does better risking war.

## 3 Extension 2: Commitment Problems

In the main model in the text, we considered a deterrence setup with incomplete information. Through that model, we found that as the defender becomes better at conducting gray zone conflict, this can lead to a challenger switching from an equilibrium where the challenger never risks war to an equilibrium where the challenger will, with positive probability, go to war.

This finding—that the defender becoming better gray zone conflict can lead to war—can also exist in models without incomplete information. Consider a scenario where the defender, D, is a rising power. Through the

standard logic of commitment problems, this scenario can rule out peaceful equilibria, and lead a challenger, C, to declare war on the defender. And, through the logic discussed in Schram (2021), an alternative equilibrium can also exist where C degrades D’s rise through low-level conflict, thereby ruling out the need to declare war. This can serve as an explanation for why low-level conflict exists in the international system despite it being inefficient: when peaceful equilibria do not exist due to a rising power, low-level conflict can exist as a “solution” to commitment problems that is cheaper than war. *Importantly, this gray-zone-conflict-equilibrium will only be supported so long that D is not too effective at gray zone conflict. If D gets too effective at gray zone conflict, then this could render C ineffective at slowing D’s rise through low-level conflict; this, in turn, could result in C switching over to fight a war. In other words, in this setup with commitment problems (where peaceful equilibria fail to exist), D becoming better at low-level conflict can eliminate equilibria where gray zone conflict occurs and result in the game ending in war.*

We offer one version of how this works. A more extensive model with additional modeling technology (for example, that also considering C’s resolve) is possible, but we choose to keep things more streamlined here to highlight the key idea: in an environment with commitment problems (and no incomplete information), D becoming better at gray zone conflict can motivate C to declare war. The logic is similar to that in Observation 3 in Schram (2021); if C becomes less effective at slowing D’s rise, then C may revert to going to war.

To offer some discussion of where things are going, we will analyze a model and establish conditions describing a scenario where a peace equilibrium does not exist and an equilibrium with gray zone conflict does exist. We consider these conditions as being those that most faithfully represent equilibria where states cannot simply bargain their way out of gray zone conflict. Then, we make D more effective at gray zone conflict. We can then show, sometimes that this results in the gray zone conflict equilibria failing to exist; in this case, to handle the rising power, C may prefer going to war. This result offers one possible explanation for Ukraine that is consistent with Schram (2021). After Euromaidan, Ukraine’s potential westward shift represented a power shift (in the eyes of Russia). Russia responded by conducting gray zone conflict. Then, as Ukraine got better at gray zone conflict, this change may have incentivised Russia to switch over to war. The logic for this switch remains the same as it was above; as C gets worse at gray zone conflict, C will consider alternate options, with one of those options being war.

### 3.1 Extension 2 Model: Game form and utilities

Two states, C and D, are in a full-information, infinite-horizon crisis bargaining game where D is a rising power. State D begins the game in control of a fully divisible asset (or has proposal power for a policy) with per-period normalized value of 1. So long that D has not lost in a war, D decides the share of the asset that C receives. This game structure resembles the infinite horizon games explored in Fearon (1995) and Powell (1999), with the crucial modification that both states can conduct low-level conflict against the other. In this regard, the model here is also similar to Schram (2021), only here we will consider a shock to D’s ability to conduct gray zone conflict.

We assign  $t \in \{1, 2, 3, \dots\}$  to denote periods. If period  $t - 1$  did not experience war, then period  $t$  begins with C selecting a limited action (or not) that will degrade D’s rise. This is denoted  $g_{C,t} \in \{0, 1\}$ . D can then respond with their own limited action (or not) that will counter at least some of C’s limited action, denoted  $g_{D,t} \in \{0, 1\}$ . D also then makes C an offer from a continuum,  $x_t \in [0, 1]$ . Finally in the period, both actors simultaneously choose to “accept” the offer and low-level conflict or to declare “war.” If both states accept in period  $t$ , the game moves forward to period  $t + 1$ . If either state declares war, then the states fight and the outcome of war is determined probabilistically. After war, states continue receiving their war payoffs but they no longer make offers or engage in low-level conflict. The salient feature of war as modeled here is that states will fight and the political issue will be resolved. Under this game form, low-level conflict is not a crisis-ending move, but declaring war is. To summarize, the game form is as follows.

1. Period  $t = 1$  begins.
2. C selects  $g_{C,t} \in \{0, 1\}$ .
3. D selects  $g_{D,t} \in \{0, 1\}$  and  $x_t \in [0, 1]$ .

4. C and D simultaneously select whether to “accept” or declare “war.” If either or both actors declare war, the game terminates and actors receive their war payoffs. If both actors accept, utilities for the period are realized, the period is updated to  $t = t + 1$ , and the game returns to the second step (where C selects  $g_{C,t}$ ).

The function  $P(t, g_{C,t-1}, g_{D,t-1}, \alpha_C, \alpha_D)$  defines the probability that C wins in war in period  $t$ . We assume functional form  $P(t, g_{C,t-1}, g_{D,t-1}, \alpha_C, \alpha_D) = \rho - 1_{t=2}\mu - 1_{t \geq 3}\bar{\mu} + \max\{\alpha_C g_{C,t-1} - \alpha_D g_{D,t-1}, 0\}$ . Here D is a rising power that rises across two periods, with the  $P$  function denoting C’s likelihood of winning should a war occur. The first indicator function,  $1_{t=2}$  implies that, in period 2, C’s likelihood of winning in war decreases by  $\mu$ . The second indicator function,  $1_{t \geq 3}$  implies that in period 3 (and all periods forward) C’s likelihood of winning in war decreases by  $\bar{\mu}$ , where  $\bar{\mu} > \mu$ . For technical convenience, we go slightly beyond this and assume  $\delta\bar{\mu} > \mu$ .<sup>23</sup> We assume that  $\rho > \bar{\mu}$ , which keeps final probability values greater than zero. We also treat C and D undertaking low-level conflict as affecting future payoffs. C conducting gray zone conflict in the previous period ( $t - 1$ ) increases C’s likelihood of winning in war by  $\alpha_C > 0$ , and D countering this in the previous period then decreases C’s likelihood of winning in war by  $\alpha_D > 0$ . We assume that D cannot ever make C more of a declining state through gray zone conflict (the “max” function implies that these actions bottom-out at 0). We also assume that  $\mu \geq \alpha_C > \alpha_D$ , which means that C is unable to completely stop D’s rise over time, and that D is only partially effective at countering C’s efforts to degrade the rise. Lastly, we assume that no low-level conflict occurs before the crisis begins, or that  $g_{C,0} = g_{D,0} = 0$ .

Both hassling and war create costs. War generates a fixed cost for both states, with  $\kappa_C > 0$  and  $\kappa_D > 0$  denoting C’s and D’s costs of war. Following Fearon (1995), we model the costs of war as a one-time cost.<sup>24</sup> The cost of low-level conflict are represented by  $\beta_C > 0$  and  $\beta_D > 0$  and are modeled as per-period costs and appear in actor’s utility functions in period  $t$  as  $\beta_C g_{C,t}$  and  $\beta_D g_{D,t}$ . We also assume the costs of low-level conflict cannot exceed the costs of war. This is consistent with treating low-level conflict as being limited relative to war. These assumptions are  $\frac{\beta_C}{1-\delta} \leq \kappa_C$  and  $\frac{\beta_D}{1-\delta} \leq \kappa_D$ .

Finally, we assume C conducting limited conflict against D where D does not respond as generating additional costs for D. D faces additional per-period costs  $\gamma_D G(g_{C,t}, g_{D,t})$ , where  $\gamma_D > 0$ ,  $G(0, 0) = 0$ ,  $G(1, 1) = 0$  and  $G(1, 0) = 1$ . Below we will assume that D not countering gray zone conflict (i.e. setting  $g_{D,t} = 0$  when  $g_{C,t} = 1$ ) is quite bad for D. Specifically, we will assume that D incurs more costs by just letting gray zone conflict happen to them relative to going to war. Specifically, this means that  $\frac{\gamma_D}{1-\delta} > \kappa_D$ ; essentially, D would rather go to war than let C conduct limited challenges against them without any response.

If either state declares war, then states go to war. State C’s expected present value of going to war in period  $t$  is  $-\kappa_C + P(t, g_{C,t-1}, g_{D,t-1}, \alpha_C, \alpha_D)/(1 - \delta)$  and D’s expected present value is  $-\kappa_D + (1 - P(t, g_{C,t-1}, g_{D,t-1}, \alpha_C, \alpha_D))/(1 - \delta)$ . Following “accept-accept,” in period  $t$ , C receives payoff  $x_t - \beta_C g_{C,t}$  and D receives payoffs  $1 - x_t - \beta_D g_{D,t} - G(g_{C,t}, g_{D,t})$ .

### 3.2 Extension 2 Model: Equilibria Types and Existence

Below we reference several types of subgame perfect equilibria. In the first type of equilibria, war occurs.

**Definition:** In a **War Equilibrium**, in some period  $t \in \{1, 2, 3, \dots\}$ , C, D, or both declare war.

A War Equilibrium always exists. When both C and D declare war, there are no unilateral deviations that generate better outcomes because both states would still be at war.

**Proposition 1:** A War Equilibrium, with war occurring in some period  $t \in \{1, 2, 3, \dots\}$ , always exists.

Proposition 1 facilitates the existence of Non-War Equilibria. The Non-War Equilibria below rely on a strategy where deviations from an action schedule (in other words, a pair low-level conflict actions and offers across every period) results in the “punishment” of both parties committing to war. Because both states

<sup>23</sup>This assumption simplifies the dynamic analysis and where it is important will be noted below.

<sup>24</sup>An alternate framework used in Powell (1999) better captures the feature that war destroys some portion of the asset and that destruction costs persist over time, while the Fearon (1995) framework (used here) better represents war generating large up-front costs. If instead we were to adopt the Powell (1999), 86-90 approach to modeling the costs of war, the Propositions below would be slightly different but the results would not substantively change.

declaring war in any period is a subgame perfect equilibrium, both states declaring war after a deviation from on-path play is a credible off-path punishment.

In the second type of equilibria, neither hassling nor war occurs.

**Definition:** In a **Peace Equilibrium**, for each  $t \in \{1, 2, 3, \dots\}$ ,  $C$  selects  $g_{C,t} = 0$ ,  $D$  selects  $g_{D,t} = 0$ ,  $D$  offers  $x_t \in [0, 1]$ , and both players accept for all  $t$ .

Under select conditions, Peace Equilibria can be sustained through off-path punishments. Proposition 2 replicates the well-acknowledged finding that the threat of a rising power can lead to a preventive war.

**Proposition 2:** *If and only if*

$$1 \geq \rho + \frac{\delta\bar{\mu} - \mu}{1 - \delta} - \delta\kappa_D - \kappa_C \quad (6)$$

and

$$1 \geq \rho + \frac{\mu\delta}{1 - \delta} - \delta\kappa_D - \kappa_C, \quad (7)$$

then a Peaceful Equilibrium exists.

*Proof:* See below.

In the third type of equilibria, gray zone conflict occurs. We assume this takes on a specific form, similar to the “stable hassling equilibrium” discussed in Schram (2021).

**Definition:** In a **Stable Gray Zone Equilibrium**,  $D$  offers  $x'_1 \in [0, 1]$  in period  $t = 1$  and  $x'_2 \in [0, 1]$  for period  $t = 2$ , and  $x'_3 \in [0, 1]$  for all  $t \in \{3, 4, 5, \dots\}$ .  $C$  sets  $g_{C,t} = 1$  and  $D$  sets  $g_{D,t} = 1$  for all  $t$ . So long that both players remain on the defined offer-low-level-conflict schedule, both players accept; otherwise, if there is a deviation in period  $t$ , then both states declare war in that period.

And we can define the existence conditions.

**Proposition 3:** *If and only if*

$$1 \geq \rho + \frac{\delta\bar{\mu} - \underline{\mu} + \delta\beta_D + \beta_C}{1 - \delta} + \alpha_C - \alpha_D - \delta\kappa_D - \kappa_C \quad (8)$$

and

$$1 \geq \rho - \kappa_C - \delta\kappa_D + \frac{\mu\delta + \beta_C + \delta(\alpha_D - \alpha_C + \beta_D)}{1 - \delta} \quad (9)$$

then a Stable Gray Zone Equilibrium exists. *Proof:* See below.

It is worthwhile noting that many of these equilibria can exist simultaneously. However, we can see how the conditions for a Stable Gray Zone Equilibrium and a Peace Equilibrium are different. We would expect to find gray zone conflict for parameters where a Peace Equilibrium fails to exist but a Stable Gray Zone Equilibrium does exist. Essentially here, states cannot bargain their way to an equilibrium without conflict, but can find a more efficient outcome than resorting to a war (i.e. engaging in low-level conflict).

Note that from these conditions alone, we have the result that as  $D$  gets better at gray zone conflict ( $\alpha_D$  increases), the Stable Gray Zone Equilibrium may fail to exist (because the right-hand-side of 9 is increasing in  $\alpha_D$ ). Thus, in this model, we have established that  $D$  becoming better at gray zone conflict can eliminate the existence of equilibria with gray zone conflict; when a peaceful equilibrium also doesn't exist, this can result in war. The result here is similar to Observation 3 in Schram (2021); as  $C$  becomes less effective at gray

zone conflict (here due to D becoming better at gray zone conflict), an equilibrium with gray zone conflict will not longer exist, and D may go to war.

To offer one example, under parameters  $\rho = 0.9$ ,  $\delta = 0.63$ ,  $\kappa_C = 0.2$ ,  $\kappa_D = 0.3$ ,  $\alpha_C = 0.25$ ,  $\alpha_D = 0.2$ ,  $\beta_C = 0.001$ ,  $\beta_D = 0.001$ ,  $\gamma = 0.7$ ,  $\mu = 0.3$ ,  $\bar{\mu} = 0.4$ . Under these parameters, no peaceful equilibrium exists (Condition 7 breaks) but a Stable Gray Zone Conflict Equilibrium exists. But, if  $\alpha_D$  is raised to  $\alpha_D = 0.24$ , then the Stable Gray Zone Conflict Equilibrium breaks down; however, a war equilibrium still exists, meaning the change might result in a switch from an equilibrium where actors always conduct gray zone conflict to an equilibrium where actors go to war.

### 3.3 Proving Proposition 3: Stable Gray Zone Equilibrium

We begin by stating a useful lemma that will be used throughout the existence proofs.

Lemma 1: *Let action triplets  $(x_1, g_{C,1}, g_{D,1})$ ,  $(x_2, g_{C,2}, g_{D,2})$ ,  $(x_3, g_{C,3}, g_{D,3})$ , ... denote on-path equilibrium play. These actions can constitute a pure strategy Non-War Equilibrium if the following constraints hold:*

$$\sum_{i=1}^{\infty} \delta^{i-1} (x_i - \beta_C g_{C,i}) \geq \frac{P(1, 0, 0, \alpha_C, \alpha_D)}{1 - \delta} - \kappa_C, \quad (10)$$

$$\sum_{i=j}^{\infty} \delta^{i-2} (x_i - \beta_C g_{C,t}) \geq -\delta^{j-2} \kappa_C + \sum_{i=j}^{\infty} \delta^{i-2} (P(i, g_{C,i-1}, g_{D,i-1}, \alpha_C, \alpha_D)) \quad \forall j \in \{2, 3, 4, \dots\}, \quad (11)$$

$$\sum_{i=1}^{\infty} \delta^{i-1} (1 - x_i - \beta_D g_{D,i} - \gamma_D G(g_{C,t}, g_{D,t})) \geq -\kappa_D + \frac{1 - P(1, 0, 0, \alpha_C, \alpha_D)}{1 - \delta}, \quad (12)$$

$$\sum_{i=j}^{\infty} \delta^{i-2} (1 - x_i - \beta_D g_{D,i} - \gamma_D G(g_{C,t}, g_{D,t})) \geq -\delta^{j-2} \kappa_D + \sum_{i=j}^{\infty} \delta^{i-2} (1 - P(i, g_{C,i-1}, g_{D,i-1}, \alpha_C, \alpha_D)) \quad \forall j \in \{2, 3, 4, \dots\}. \quad (13)$$

Note: Lemma 1 states the most general conditions for Non-War Equilibrium existence. Because the War-War equilibrium is the worst possible feasible subgame perfect ‘punishment’ equilibrium, it is used to support the largest possible set of equilibria.

#### 3.3.1 Deriving the Conditions

In this subsection, we offer some intuition behind how the conditions are derived.

Based on Lemma 1,<sup>25</sup> the following must hold for C (starting in period 3 and working backwards):

$$\begin{aligned} \frac{x_3 - \beta_C}{1 - \delta} &\geq \frac{\rho - \bar{\mu} + \alpha_C - \alpha_D}{1 - \delta} - \kappa_C, \\ x_2 - \beta_C + \frac{\delta(x_3 - \beta_C)}{1 - \delta} &\geq \frac{\rho - \mu + \alpha_C - \alpha_D}{1 - \delta} - \kappa_C, \\ x_1 - \beta_C + \delta(x_2 - \beta_C) + \frac{\delta^2(x_3 - \beta_C)}{1 - \delta} &\geq \frac{\rho}{1 - \delta} - \kappa_C. \end{aligned}$$

<sup>25</sup>We do not consider  $x_3 \neq x_4 \neq x_5 \neq \dots$  because existence relies on the most generous possible offers to C, which are based on D's wartime payoffs. Because D's wartime payoffs in periods 3 and onward are the same in the Peace Equilibrium, we do not need to consider different offers made in periods  $t = 3$  onward.

And the following must hold for D (starting in period 3 and working backwards):

$$\begin{aligned}\frac{1 - x_3 - \beta_D}{1 - \delta} &\geq \frac{1 - \rho + \bar{\mu} + \alpha_D - \alpha_C}{1 - \delta} - \kappa_D, \\ 1 - x_2 - \beta_D + \frac{\delta(1 - x_3 - \beta_D)}{1 - \delta} &\geq \frac{1 - \rho + \underline{\mu} + \alpha_D - \alpha_C}{1 - \delta} - \kappa_D, \\ 1 - x_1 - \beta_D + \delta(1 - x_2 - \beta_D) + \frac{\delta^2(1 - x_3 - \beta_D)}{1 - \delta} &\geq \frac{1 - \rho}{1 - \delta} - \kappa_D,\end{aligned}$$

Because C is the declining power here, we consider the smallest offers that can be made to D in the later periods (where C has declined)—this way C is getting the largest possible payoffs in the future, and has the greatest incentives to stick to peace. We calculate this for period 3 onward, using D's constraints from above:

$$\frac{1 - x_3 - \beta_D}{1 - \delta} \geq \frac{1 - \rho + \bar{\mu} + \alpha_D - \alpha_C}{1 - \delta} - \kappa_D$$

or

$$\rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D \geq x_3$$

This statement implies that  $x_3 = \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D$  is the largest  $x_3$  value that will keep D from declaring war (i.e. the most generous offer to C that will keep D from declaring war). Note that because  $\rho > \bar{\mu}$ ,  $\alpha_C > \alpha_D$ , and  $(1 - \delta)\kappa_D > \beta_D$ , the left-hand side is never negative.

Next, we can move to solving for the smallest possible  $x_2$  by substituting in the smallest  $x_3$  that was solved for immediately above:

$$1 - x_2 - \beta_D + \frac{\delta(1 - \rho + \bar{\mu} + \alpha_D - \alpha_C)}{1 - \delta} - \delta\kappa_D \geq \frac{1 - \rho + \underline{\mu} + \alpha_D - \alpha_C}{1 - \delta} - \kappa_D$$

or

$$\rho - \alpha_D + \alpha_C - \beta_D + \kappa_D(1 - \delta) + \frac{\delta\bar{\mu} - \underline{\mu}}{1 - \delta} \geq x_2.$$

Note that because  $\delta\bar{\mu} > \underline{\mu}$ , this expression is always non-negative.<sup>26</sup>

These inequalities represent upper bounds on what can be given to C while still minimally satisfying D. We can then turn to C's second and first period payoffs. We only consider these periods because there C is stronger and possesses the incentive to go to war. This is in contrast to period 3, when the shift in power (out of C's favor) has already occurred). For C to not go to war in the second period, the following must hold:

$$x_2 - \beta_C + \frac{\delta(x_3 - \beta_C)}{1 - \delta} \geq \frac{\rho - \underline{\mu} + \alpha_C - \alpha_D}{1 - \delta} - \kappa_C$$

or, with plugging in the most generous  $x_3$  offer to C ( $x_3 = \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D \geq x_3$ ):

$$x_2 - \beta_C + \frac{\delta(\rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D - \beta_C)}{1 - \delta} \geq \frac{\rho - \underline{\mu} + \alpha_C - \alpha_D}{1 - \delta} - \kappa_C,$$

---

<sup>26</sup>This is where the earlier assumption is relevant.



or

$$\rho + \frac{\delta\bar{\mu} - \underline{\mu} + \delta\beta_D + \beta_C}{1 - \delta} + \alpha_C - \alpha_D - \delta\kappa_D - \kappa_C \leq x_2.$$

And for C not to go to war in the first period, it must be that

$$x_1 - \beta_C + \delta(x_2 - \beta_C) + \frac{\delta^2(x_3 - \beta_C)}{1 - \delta} \geq \frac{\rho}{1 - \delta} - \kappa_C$$

\*\* or, with plugging in the most generous  $x_2$  and  $x_3$  offers to C and simplifying:

$$\rho - \kappa_C - \delta\kappa_D + \frac{\mu\delta + \beta_C + \delta(\alpha_D - \alpha_C + \beta_D)}{1 - \delta} \leq x_1.$$

So long that there exists an  $x_2$  and an  $x_1$  satisfying these inequalities, C can be disincentivised from going to war. This will be possible when  $x_1 \leq 1$  and  $x_2 \leq 1$ ; if either of these conditions fail, then a Stable Gray Zone Conflict equilibrium will not hold.

### 3.3.2 Proving Proposition 3 $\leftarrow$

If some stable gray zone equilibrium exists, then 8 and 9 hold.

We proceed with the proof as follows. We have the existence of some stable gray zone equilibrium characterized by  $(x'_1, x'_2, x'_3)$  and where  $g_{D,t} = g_{C,t} = 1$  for all  $t$ . We then want to show that if this equilibrium exists, an alternate equilibrium  $(x_1^*, x_2^*, x_3^*)$  also exists. We then use the structure of this alternate equilibrium and the bounds in Lemma 1 to derive the equations 8 and 9.

Let  $(x'_1, x'_2, x'_3)$  denote the offers in a stable gray zone equilibrium. Because this is a stable gray zone equilibrium, it satisfies Lemma 1. We re-write Lemma 1 in the terms of the equilibrium. For period 1, it must be that

$$x'_1 - \beta_C + \delta(x'_2 - \beta_C) + \frac{\delta^2(x'_3 - \beta_C)}{1 - \delta} \geq \frac{\rho}{1 - \delta} - \kappa_C \quad (14)$$

and

$$1 - x'_1 - \beta_D + \delta(1 - x'_2 - \beta_D) + \frac{\delta^2(1 - x'_3 - \beta_D)}{1 - \delta} \geq \frac{1 - \rho}{1 - \delta} - \kappa_D. \quad (15)$$

For period 2 it must be

$$x'_2 - \beta_C + \frac{\delta(x'_3 - \beta_C)}{1 - \delta} \geq \frac{\rho - \underline{\mu} + \alpha_C - \alpha_D}{1 - \delta} - \kappa_C \quad (16)$$

and

$$1 - x'_2 - \beta_D + \frac{\delta(1 - x'_3 - \beta_D)}{1 - \delta} \geq \frac{1 - \rho + \underline{\mu} + \alpha_D - \alpha_C}{1 - \delta} - \kappa_D. \quad (17)$$

And for periods 3, 4, 5, ... it must be that

$$\frac{x'_3 - \beta_C}{1 - \delta} \geq \frac{\rho - \bar{\mu} + \alpha_C - \alpha_D}{1 - \delta} - \kappa_C \quad (18)$$

and

$$\frac{1 - x'_3 - \beta_D}{1 - \delta} \geq \frac{1 - \rho + \bar{\mu} + \alpha_D - \alpha_C}{1 - \delta} - \kappa_D. \quad (19)$$

I re-write equations 18 and 19 to identify bounds on  $x'_3$ . These are

$$x'_3 \geq \rho - \bar{\mu} - \alpha_D + \alpha_C - (1 - \delta)\kappa_C + \beta_C \quad (20)$$

$$x'_3 \leq \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D. \quad (21)$$

It is also possible to re-write 16 and 17 to identify bounds on  $x'_2$ . These are

$$x'_2 \geq \frac{\rho - \underline{\mu} + \alpha_C - \alpha_D + \beta_C - \delta x'_3}{1 - \delta} - \kappa_C \quad (22)$$

$$x'_2 \leq \frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x'_3}{1 - \delta} + \kappa_D \quad (23)$$

It is also possible to re-write 14 and 15 to identify bounds on  $x'_1$ . These are

$$x'_1 \geq -\delta x'_2 + \frac{\rho + \beta_C - \delta^2 x'_3}{1 - \delta} - \kappa_C \quad (24)$$

$$x'_1 \leq -\delta x'_2 + \frac{\rho - \beta_D - \delta^2 x'_3}{1 - \delta} + \kappa_D \quad (25)$$

Using the six equations above, it is more clear to see what conditions are needed on  $(x'_1, x'_2, x'_3)$ . We now claim that if  $(x'_1, x'_2, x'_3)$  constitutes a stable hassling equilibrium, there also exists a (not necessarily unique) stable gray zone equilibrium  $(x_1^*, x_2^*, x_3^*)$  that takes the form  $x_3^* = \min\{\rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D, 1\}$ ,  $x_2^* = \min\left\{\frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D, 1\right\}$ ,  $x_1^* = \max\left\{-\delta x_2^* + \frac{\rho + \beta_C - \delta^2 x_3^*}{1 - \delta} - \kappa_C, 0\right\}$ . For  $(x_1^*, x_2^*, x_3^*)$  to constitute a stable gray zone equilibrium, we must show the  $(x_1^*, x_2^*, x_3^*)$  satisfies Lemma 1. We skip the step where we redefine each inequality in Lemma 1 in terms of the new triplet and instead directly express the reduced set of inequalities necessary for  $(x_1^*, x_2^*, x_3^*)$  to be a Stable Gray Zone Equilibrium. In addition to  $x_1^* \in [0, 1]$ ,  $x_2^* \in [0, 1]$ , and  $x_3^* \in [0, 1]$ ,

$$x_3^* \geq \rho - \bar{\mu} - \alpha_D + \alpha_C - (1 - \delta)\kappa_C + \beta_C, \quad (26)$$

$$x_3^* \leq \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D, \quad (27)$$

$$x_2^* \geq \frac{\rho - \underline{\mu} + \alpha_C - \alpha_D + \beta_C - \delta x_3^*}{1 - \delta} - \kappa_C, \quad (28)$$

$$x_2^* \leq \frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D, \quad (29)$$

$$x_1^* \geq -\delta x_2^* + \frac{\rho + \beta_C - \delta^2 x_3^*}{1 - \delta} - \kappa_C \quad (30)$$

$$x_1^* \leq -\delta x_2^* + \frac{\rho - \beta_D - \delta^2 x_3^*}{1 - \delta} + \kappa_D \quad (31)$$

We now show that this is a Stable Gray Zone Conflict Equilibrium. By their definitions,  $x_3^* \leq 1$ ,  $x_2^* \leq 1$ ,  $x_1^* \geq 0$ , and inequalities 27, 29 and 30 are satisfied. By the assumptions on cost terms and the extent of the shifts, inequalities  $x_2^* \geq 0$  [Let  $x_3^* = \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D$ . Then

$$x_2^* = \frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta(\rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D)}{1 - \delta} + \kappa_D$$

$$x_2^* = \frac{(1 - \delta)(\rho + \alpha_C - \alpha_D) - \underline{\mu} + \delta\bar{\mu}}{1 - \delta} - \beta_D + (1 - \delta)\kappa_D$$

which is always positive.] and  $x_3^* \geq 0$ ; similarly, inequalities 26 and 28 are satisfied. And, by how  $x_1^*$  is defined and by 24,  $x_1^* \leq x_1'$ ; because  $x_1'$  is (by assumption) part of a Stable Hassling Equilibrium,  $x_1' \leq 1$ , which makes  $x_1^* \leq 1$ .

Next, we need to show that inequality 31 holds. We do this by cases. In the first case, when  $x_1^* = -\delta x_2' + \frac{\rho + \beta_C - \delta^2 x_3'}{1 - \delta} - \kappa_C$ , inequality 31 holds because  $\kappa_C \geq \frac{\beta_C}{1 - \delta}$  and  $\kappa_D \geq \frac{\beta_D}{1 - \delta}$ . In the second case, when  $x_1^* = 0$ , we can substitute  $x_2^* = \frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D$  and  $x_3^* = \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D$ , which re-states inequality 31 as  $\rho - \beta_D + \frac{\delta(\underline{\mu} + \alpha_D - \alpha_C)}{1 - \delta} + (1 - \delta)\kappa_D \geq 0$ . Because  $\kappa_D \geq \frac{\beta_D}{1 - \delta}$ , and  $\underline{\mu} > \alpha_C$ , the inequality holds.

Because  $(x_1^*, x_2^*, x_3^*)$  is a Stable Gray Zone Equilibrium, we can use inequalities 30 and 28 to show inequalities 8 and 9 hold. First, whenever  $x_3^* = \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D$  and  $x_2^* = \frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D$ , it is straightforward to see 8 and 9 hold. When  $x_3^* = 1$ ,  $x_2^* = 1$ , or both, then we can use inequalities 30 and 28 to derive 8 and 9. For example, using inequality 28 when  $x_3^* = 1$  gives

$$x_2^* \geq \frac{\rho - \underline{\mu} + \alpha_C - \alpha_D + \beta_C - \delta}{1 - \delta} - \kappa_C,$$

where, because  $1 \leq \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D$ , it must also be that

$$x_2^* \geq \rho + \frac{\delta\bar{\mu} - \underline{\mu} + \delta\beta_D + \beta_C}{1 - \delta} + \alpha_C - \alpha_D - \delta\kappa_D - \kappa_C,$$

which, because  $x_2^* \leq 1$ , implies 9. A similar logic holds for 8 holding when  $x_3^* = 1$ ,  $x_2^* = 1$ , or both because  $x_3^*$  and  $x_2^*$  are subtracted from the right hand side of inequality 30.

### 3.3.3 Proving Proposition 3 $\rightarrow$

If  $1 \geq \rho + \frac{\delta\bar{\mu} - \underline{\mu} + \delta\beta_D + \beta_C}{1 - \delta} + \alpha_C - \alpha_D - \delta\kappa_D - \kappa_C$  and  $1 \geq \rho - \kappa_C - \delta\kappa_D + \frac{\underline{\mu}\delta + \beta_C + \delta(\alpha_D - \alpha_C + \beta_D)}{1 - \delta}$  hold, then a Stable Gray Zone Equilibrium can exist.

To prove this, we define triple  $(x_1^*, x_2^*, x_3^*)$  that takes the form  $x_3^* = \min\{\rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D, 1\}$ ,  $x_2^* = \min\left\{\frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D, 1\right\}$ ,  $x_1^* = \max\left\{-\delta x_2^* + \frac{\rho + \beta_C - \delta^2 x_3^*}{1 - \delta} - \kappa_C, 0\right\}$ .  $(x_1^*, x_2^*, x_3^*)$  is a Stable Gray Zone Equilibrium if  $x_1^* \in [0, 1]$ ,  $x_2^* \in [0, 1]$ ,  $x_3^* \in [0, 1]$ , and the following hold:

$$x_3^* \geq \rho - \bar{\mu} - \alpha_D + \alpha_C - (1 - \delta)\kappa_C + \beta_C, \quad (32)$$

$$x_3^* \leq \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D, \quad (33)$$

$$x_2^* \geq \frac{\rho - \underline{\mu} + \alpha_C - \alpha_D + \beta_C - \delta x_3^*}{1 - \delta} - \kappa_C, \quad (34)$$

$$x_2^* \leq \frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D, \quad (35)$$

$$x_1^* \geq -\delta x_2^* + \frac{\rho + \beta_C - \delta^2 x_3^*}{1 - \delta} - \kappa_C \quad (36)$$

$$x_1^* \leq -\delta x_2^* + \frac{\rho - \beta_D - \delta^2 x_3^*}{1 - \delta} + \kappa_D \quad (37)$$

By definition,  $x_1^* \geq 0$ ,  $x_2^* \leq 1$ ,  $x_3^* \leq 1$ , and inequalities 33, 29 and 36 hold. Additionally, by the assumptions on  $\kappa_C$ ,  $\beta_C$ ,  $\kappa_D$ , and  $\beta_D$ , and on the power shifts,  $x_3^* \geq 0$  and 32 holds.

To show 34 holds, we must consider cases. First, when  $x_2^* = \frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D$ , the condition holds based on the properties on  $\kappa_C$ ,  $\beta_C$ ,  $\kappa_D$ , and  $\beta_D$ . When  $x_2^* = 1$ , we must consider two sub-cases. First, when  $x_3^* = 1$ , we can say  $\frac{\rho - \underline{\mu} + \alpha_C - \alpha_D + \beta_C}{1 - \delta} - \kappa_C < 1$ . Second, when  $x_3^* = \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D$ , we can substitute this value in and the inequality holds by assumption (specifically  $1 \geq \rho + \frac{\bar{\mu} - \delta \underline{\mu} + \delta \beta_D + \beta_C}{1 - \delta} + \alpha_C - \alpha_D - \delta \kappa_D - \kappa_C$ ).

To show 37 holds, we must consider cases. First, let  $x_1^* = -\delta x_2^* + \frac{\rho + \beta_C - \delta^2 x_3^*}{1 - \delta} - \kappa_C$ . It is clear that inequality 37 holds for any  $x_2^*$  or  $x_3^*$ . Suppose  $x_1^* = 0$ . Let  $x_2^* = \frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D$  and  $x_3^* = \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D$ , we substitute these in to the RHS of the equation, giving

$$\begin{aligned} 0 &\leq -\delta \left( \frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D \right) + \frac{\rho - \beta_D - \delta^2 x_3^*}{1 - \delta} + \kappa_D \\ 0 &\leq \left( \frac{-\delta \rho + \delta \underline{\mu} + \delta \alpha_D - \delta \alpha_C + \delta \beta_D + \delta^2 x_3^* + \rho - \beta_D - \delta^2 x_3^*}{1 - \delta} \right) + \kappa_D - \delta \kappa_D \\ 0 &\leq \left( \frac{(1 - \delta)\rho + \delta \underline{\mu} + \delta \alpha_D - \delta \alpha_C - \beta_D(1 - \delta)}{1 - \delta} \right) + (1 - \delta)\kappa_D \\ 0 &\leq \rho + (1 - \delta)\kappa_D - \beta_D + \left( \frac{\delta \underline{\mu} + \delta \alpha_D - \delta \alpha_C}{1 - \delta} \right) \end{aligned}$$

Based on the conditions on costs and power shifts, the RHS is positive, meaning the inequality holds. And, we do not need to consider cases where  $x_2^* = 1$  or  $x_3^* = 1$ ; whenever (for example)  $x_2^* = 1$ , it implies that  $\frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D > 1$ , which implies that if the inequality holds for the larger value  $\frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D$  it will hold for the smaller value 1.

We must also show that  $x_1^* \leq 1$ .  $x_1^* \leq 1$  will hold whenever  $-\delta x_2^* + \frac{\rho + \beta_C - \delta^2 x_3^*}{1 - \delta} - \kappa_C \leq 1$ . Letting  $x_2^* = \frac{\rho - \underline{\mu} - \alpha_D + \alpha_C - \beta_D - \delta x_3^*}{1 - \delta} + \kappa_D$  and  $x_3^* = \rho - \bar{\mu} - \alpha_D + \alpha_C + (1 - \delta)\kappa_D - \beta_D$ , we substitute these values in to the original expression, we get  $1 \geq \rho - \kappa_C - \delta \kappa_D + \frac{\mu \delta + \beta_C + \delta(\alpha_D - \alpha_C + \beta_D)}{1 - \delta}$ , which holds by assumption. Whenever  $x_2^* = 1$  or  $x_3^* = 1$ , then  $x_1^*$  remains less than 1 (based on the logic of the previous paragraph).

### 3.3.4 Proving Proposition 2

This follows the proof of Proposition 3—if  $g_{C,t} = g_{D,t} = 0$  for all  $t$ , then this is equivalent to if  $\alpha_D = \alpha_C = \beta_C = \beta_D = 0$ . All calculations above would go through.

## 4 New data

The universe of cases was created by first identifying cases of Russian foreign interventions from 3 prior datasets; ICB (Brecher and Wilkenfeld 1997), DCID (Valeriano and Maness 2014), and REI (Casey and Way 2017). Code replicating those findings is provided in the appropriate RMarkdown files. These cases were then supplemented with additional cases of Russian interference the authors were able to identify.

### 4.1 Comparison of current datasets

A comparison of what cases were covered in each individual dataset is provided in Figure A4. Note that there are significant inconsistencies concerning the sample of post-1994 Russian interventions identified by the ICB, DCID, and REI datasets.

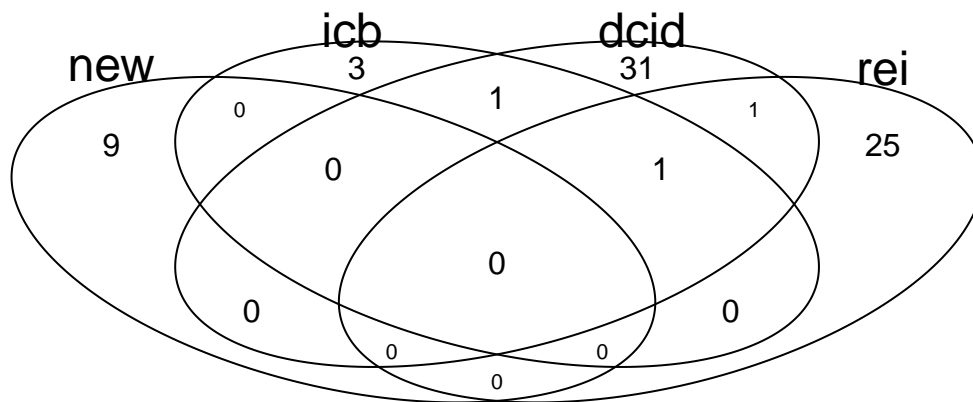


Figure A4: Venn diagram of case overlap among prior datasets

Aside from the cases covered, the intensity codings for current datasets are difficult to compare given their different scales. A more thorough analysis is provided in the appropriate R Markdown files, but a comparison of intensity codings in DCID (Valeriano and Maness 2014) and REI (Casey and Way 2017) is depicted in Figure A5. The DCID data identifies the United States, United Kingdom, Poland and Ukraine as targets of the most severe Russian cyber operations. In the cases documented by REI, the most severe Russian attacks occurred against France, Austria, and Ukraine. Part of this discrepancy is due to the respective foci of each dataset; DCID seeks out cases of cyber incidents and disputes while REI focuses on Russian electoral interference. While a majority of the REI cases include some form of Russian cyber activity, there are a few cases where only material support was provided (eg. Moldova 2014 and Belarus 1994).

This discrepancy exemplifies not only the challenges of relying on open source reporting for identifying cyber influence or disruption campaigns, but also differences in defining what counts as an attack. The only country-year that appears in both datasets is Ukraine 2014. We standardized codings across the two datasets using variable definitions from respective codebooks. A severity less than or equal to 2 in DCID's coding is synonymous in our recoding with REI's coding for disinformation, a severity between 3 and 7 equals REI's coding for cyberattack, and no cases in DCID have a severity greater than 7. We adopted Valeriano and Maness (2014)'s approach of sampling on intensity when there are multiple observations in a given time unit.

### 4.2 Variable codings

For each incident, we code whether Russia used conventional ground forces, conventional air or sea forces, paramilitary or covert forces, cyber disruption, and information operations. By distinguishing between these five types of aggression, we obtain a clearer picture of the intensity of each case of Russian intervention. The vast majority of cases include at least some type of cyber operations. In a few cases, data limitations preclude coding of non-kinetic activity by Russia or other actors. In Moldova 2005, for example, Russia provided material support for the Communist Party but there is no credible evidence of cyber activities.

The following binary coding criteria were used for each case:

**Intensity of Russian cyber attacks (2005-2017)  
Valeriano and Maness data**



**Intensity of Russian cyber attacks (1994-2017)  
Way and Casey data**



Figure A5: Comparison of coding for highest intensity Russian intervention in each target state

- **resp\_infoops** - Did Russia use information operations during this event? That includes propaganda, misinformation campaigns, theft of information, and other simple intrusions
- **resp\_cyberdisrup** - Did Russia use cyber attacks during this operation? That includes hacking, phishing, cyber espionage, DDoS attacks, etc. that constitute a system shut down rather than simple intrusions
- **resp\_paramil** - Did Russia use paramilitary troops during this event? Special forces, covert troops, speznatz, etc all count
- **resp\_convmil\_airsea** - Did Russia use conventional naval or air forces during this event?
- **resp\_convmil\_gro** - Did Russia use conventional ground troops like their army, artillery, tanks, etc during this event?

The complete dataset is provided in the appropriate .csv file. It includes sources used for the codings as well as justifications and explanations where needed.

### 4.3 Summary statistics

Although data was compiled on Russian intervention against all states from 1994-2018, the statistical analysis is limited to a sample from European states. In alignment with that, Table A7 present descriptive statistics of the sample used in the models provided in the main text.

Table A7: Covariate Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Intensity	1,000	0.1	0.4	0	5
NATO member	1,000	0.5	0.5	0	1
Dist. from Russia (minimum, log)	1,000	5.2	2.8	0.01	7.8
Democracy	926	0.9	0.3	0	1
Nuclear state	1,000	0.05	0.2	0	1
Population (log)	1,000	15.8	1.4	12.5	18.2
CINC ratio	754	0.1	0.1	0.0	0.4
GDP per capita	992	24.7	23.0	0.6	112.4
Military expenditure	962	7.3	13.3	0.0	59.8

Sample includes all European states (1994-2018). Binary variables converted to numeric.

The distribution of our dependent variable, intensity, is shown for the European sample in Figure A6. The figure only includes country-years with known attacks (omitting null cases) to allow an easier visual comparison of variation in attack intensity.

The bivariate correlations between the DV and the two EVs are shown in Figure A7. The intensity and NATO variables have been converted to numeric values to simplify visualizing the bivariate correlations.

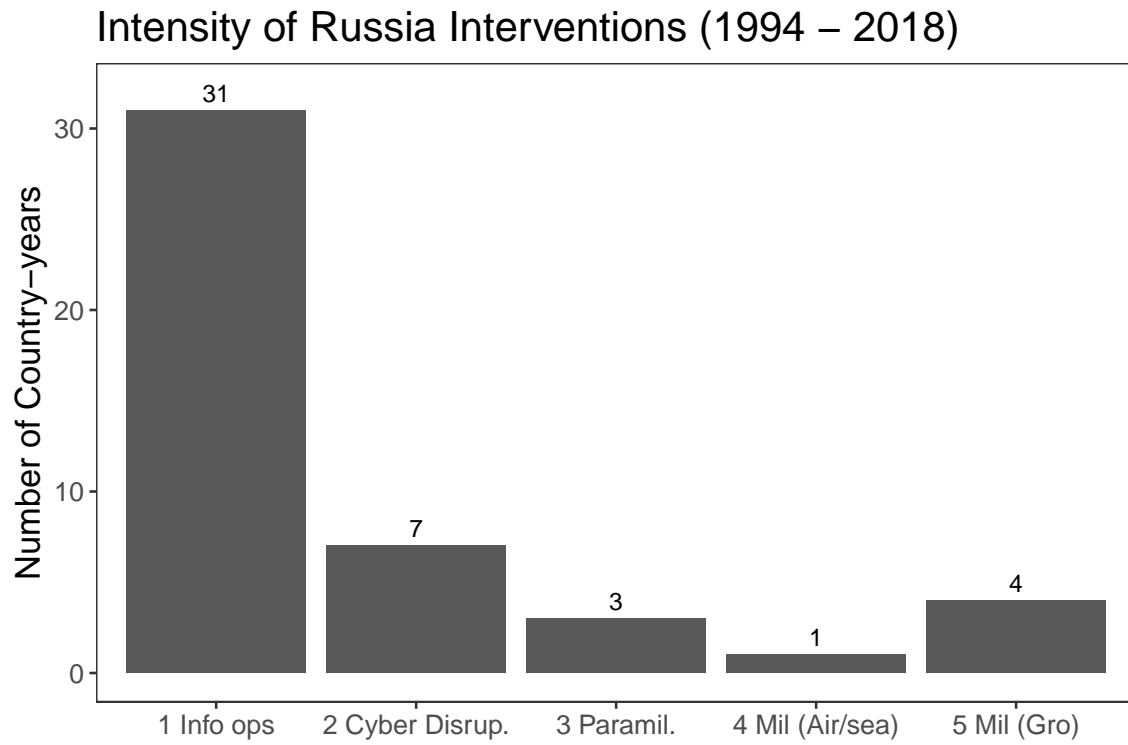


Figure A6: Intensity of Russian interventions

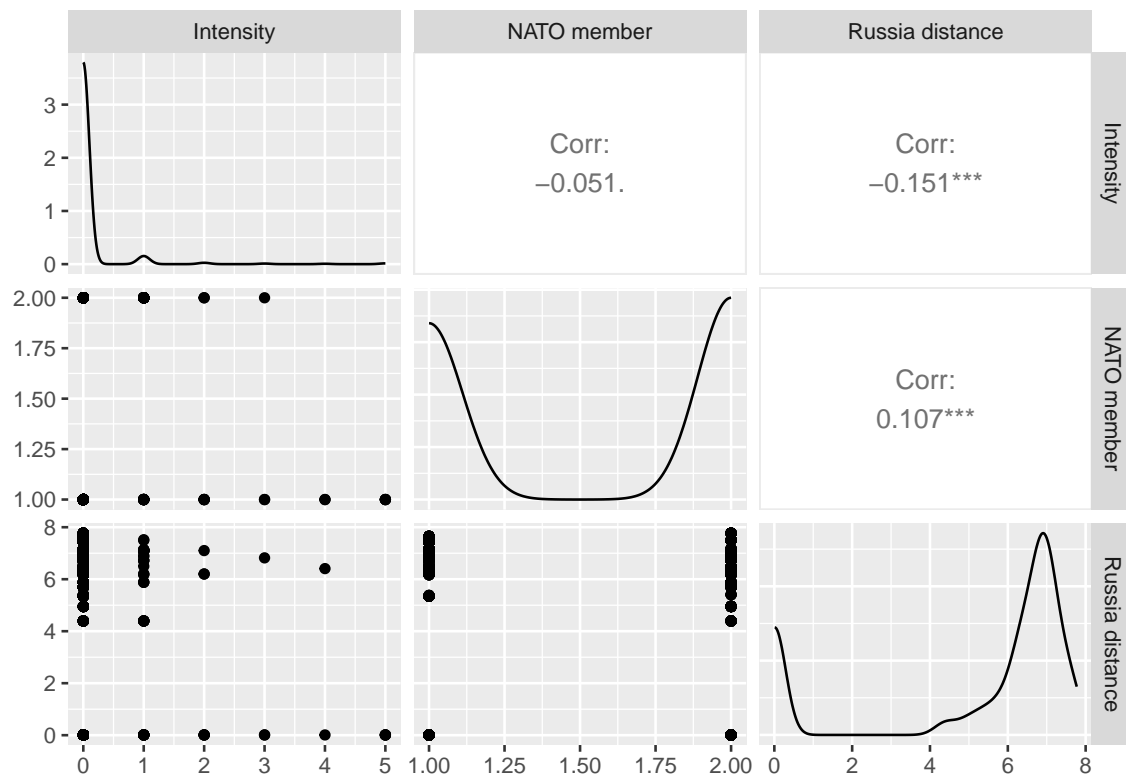


Figure A7: Bivariate correlation of dependent and independent variables



## 5 Alternate model specifications

We run a set of alternate model specifications as robustness checks. Our results are consistent across alternate modeling specifications including different regression models, control variables, and imputation strategies. We choose the ordered probit results as the main results given the appropriateness of that model specification and to ensure our primary results are not simply an artifact of our imputation strategy. Those results are shown below.

### 5.1 Alternate alliance measure

Because the dichotomous NATO/Non-NATO independent variable could be viewed as too coarse, we also re-run our six empirical models with an additional dummy variable for if a state was in a NATO Membership Action Plan (MAP), part of the NATO Partnership for Peace (PfP), or if the state was engaged in Intensified Dialogue (ID). Figure A8 illustrates how states' NATO status and pre-status change over time. In line with our hypothesis that NATO membership could serve as a deterrent to Russian aggression, we would expect that entering into a partnership with NATO could also deter Russia actions. In order to be invited into NATO, states must be in good standing with existing members, suggesting that some NATO states might come to the aid of a state with (for example) a MAP.

As can be seen in Table A8, introducing the pre-NATO dummy variable makes little substantive changes to our existing results. We find that NATO membership is still associated with a decrease in the intensity of Russian attacks. We also find, consistent with the discussion above, taking the steps to become a NATO member is also associated with a decrease in the intensity of Russian attacks. Once more, while we cannot claim our results are causal—just as NATO membership is not exogenous, NATO pre-membership is not exogenous—our results suggest that a state's warming relationship to NATO could decrease gray zone activity taken against it.

	Full sample			Relevant states sample		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Independent Variables						
NATO member	-0.37 (0.23)	-0.59*** (0.21)	-0.77*** (0.22)	-0.56** (0.26)	-0.67*** (0.26)	-0.81*** (0.25)
NATO pre-member	-0.86** (0.42)	-0.90** (0.38)	-1.04*** (0.39)	-4.78*** (0.00)	-4.43*** (0.00)	-4.61*** (0.00)
Russia distance	-0.10*** (0.04)	-0.09*** (0.03)	-0.11*** (0.04)	-0.04 (0.04)	-0.08** (0.03)	-0.08* (0.04)
Controls						
Democracy		0.24 (0.42)	0.53 (0.42)		0.21 (0.43)	0.50 (0.42)
Nuclear power		1.03** (0.46)	0.40 (0.48)		0.98* (0.52)	0.95 (0.76)
Population		0.15 (0.09)	0.08 (0.13)		0.13 (0.09)	0.15 (0.11)
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 A8: Pre-NATO Variable Robustness Check

It is also worthwhile mentioning: in some cases, the coefficient on NATO pre-membership has a point estimate

## Timeline for Stages of NATO Membership

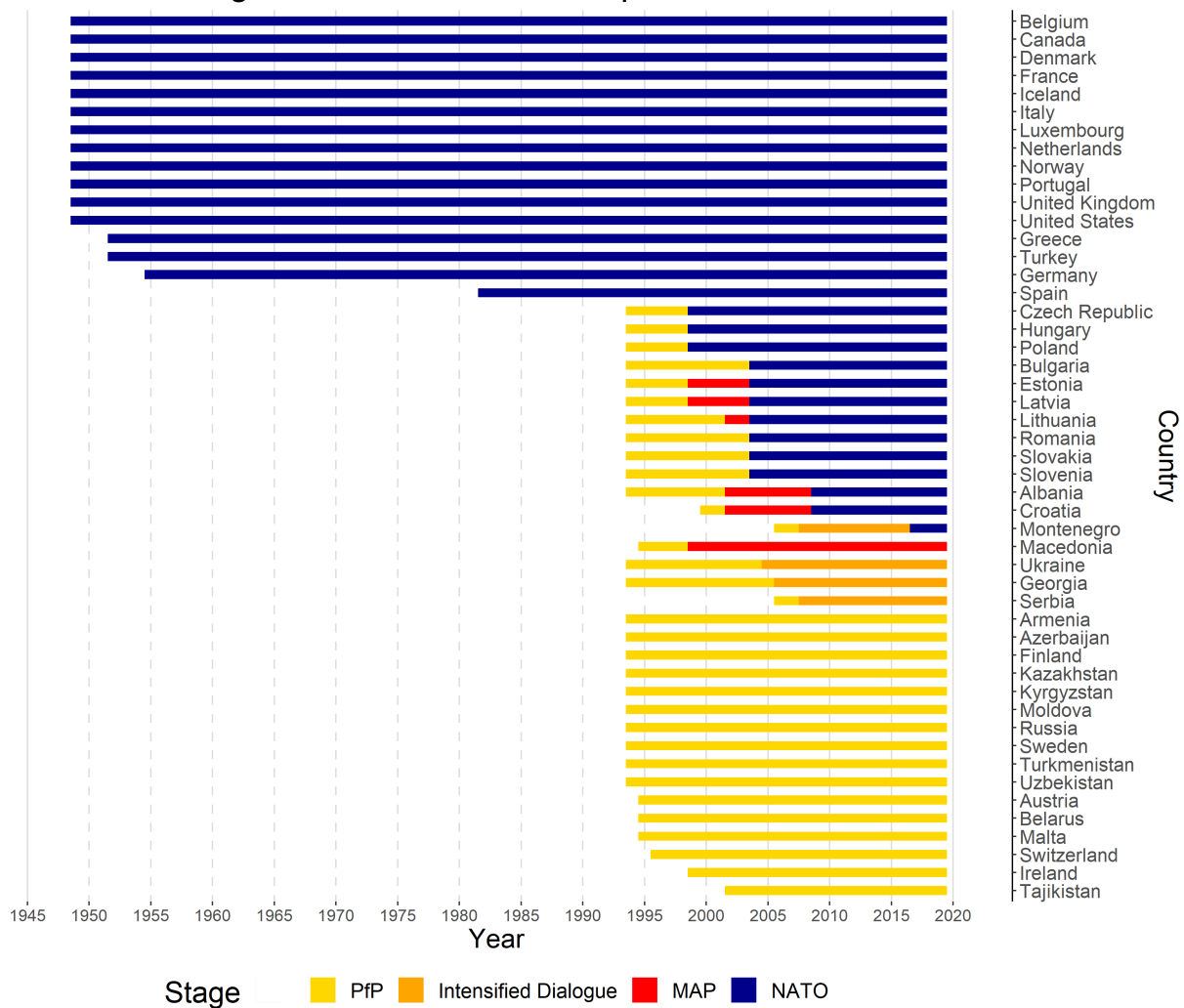


Figure A8: NATO membership timeline, including pre-NATO membership stages

that is larger in magnitude than the coefficient on NATO membership. Taken as true, this would suggest that NATO pre-membership could be a more effective deterrent than NATO membership. While for the first three models this difference is not statistically significant at the 10% level, in the last three models this difference is statistically significant at the 1% level. Ultimately, it is difficult to know how to interpret this finding. While it is plausible that this relationship is actually borne out empirically—that NATO leadership may take gray zone attacks against its pre-member states more seriously than gray zone attacks against some of its existing states—this result could also be driven by biases in the data. It could be that, in order for a state to be included in our “relevant” sample and part of Intensified Dialogue, that state has been screened by NATO to be particularly unlikely to be attacked by Russian gray zone efforts; this would bias our results, and could create spurious results. While we cannot fully ascertain what is driving this quirk in our results, this robustness check has ultimately presented further positive (albeit suggestive) evidence that is in-line with our hypothesis: gray zone behavior is being shaped by NATO’s deterrent threat. This does demonstrate that debates about both the causes and consequences of post-Cold War NATO expansion warrant important attention (Shiffrinson 2016; Lanoszka 2020).

## 5.2 Alternate distance measures

The primary models in the main text use NATO members as our observable indicator of Western deterrence and minimum distance to Russia as our observable indicator of Russian resolve and gray zone efficacy. An alternate measure would instead consider distance to NATO as a proxy for NATO’s external deterrence threat. We conduct additional models that use this distance variable instead. We take the natural log of each country-year’s capital to capital distance to the nearest NATO member. For NATO members, the value is set to 0.01 km to allow the natural log to calculate.

Table A9 shows the results of this alternate measure are consistent with the original model. In all model specifications with control variables, both distance measures are in the expected direction and statistically significant at at least the  $p < 0.05$  level.

	Full sample			Relevant states sample		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Independent Variables						
NATO distance	0.07* (0.04)	0.10*** (0.03)	0.13*** (0.03)	0.10** (0.04)	0.12*** (0.04)	0.14*** (0.04)
Russia distance	−0.10*** (0.03)	−0.10*** (0.03)	−0.12*** (0.03)	−0.04 (0.04)	−0.08** (0.04)	−0.09** (0.04)
Controls						
Democracy		0.15 (0.41)	0.43 (0.41)		0.09 (0.42)	0.40 (0.41)
Nuclear power		0.99** (0.44)	0.44 (0.47)		0.97* (0.50)	0.98 (0.80)
Population		0.18** (0.09)	0.12 (0.11)		0.16* (0.09)	0.17 (0.11)
GDP per cap		−0.01** (0.01)	−0.01** (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 A9: NATO Distance Measure Robustness Check

### 5.3 Odds ratios

Given the difficulty of interpreting ordered probit coefficients, Table A4 show the results as odds ratios with confidence intervals in parentheses when all other variables are held at their mean level. To use model 6 as an example for interpretation, for relevant NATO states, the odds of a non-cyber, non-information attack (categories 3, 4, or 5) versus a cyber attack, an information attack, or no attack is 49% lower.

Independent Variables	Full sample			Relevant states sample		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
NATO member	0.76 [0.39; 1.12]	0.62* [0.28; 0.95]	0.54* [0.18; 0.90]	0.62 [0.21; 1.04]	0.56* [0.13; 0.98]	0.50* [0.09; 0.91]
Russia distance	0.90* [0.84; 0.96]	0.90* [0.85; 0.95]	0.88* [0.83; 0.94]	0.95 [0.89; 1.02]	0.92* [0.86; 0.98]	0.91* [0.84; 0.98]
Controls						
Democracy		1.16 [0.46; 1.87]	1.56 [0.89; 2.24]		1.12 [0.39; 1.85]	1.53 [0.83; 2.23]
Nuclear power		2.67* [1.93; 3.41]	1.64 [0.86; 2.42]		2.63* [1.78; 3.48]	2.97* [1.59; 4.36]
Population		1.20* [1.05; 1.35]	1.14 [0.94; 1.34]		1.17 [1.00; 1.33]	1.19 [0.98; 1.40]
GDP per cap		0.99* [0.98; 1.00]	0.98* [0.97; 1.00]		0.99 [0.98; 1.00]	0.99 [0.97; 1.00]
Mil. spending			1.02 [1.00; 1.03]			1.00 [0.97; 1.03]
Observations	1,000	921	891	376	373	346

All models are ordered probits and include year-fixed effects with country-clustered standard errors in parentheses.

Table A10: Odds Ratios

### 5.4 OLS regression

Although an ordered probit model is most appropriate given the dependent variable (intensity) is ordinal, we ensure that the sign on our coefficients are consistent with an OLS model that treats intensity as a continuous variable. See Table A11.

### 5.5 Ordered logit

We also run all models as ordered logits instead of ordered probits. Both are generalized linear models appropriate for an ordinal dependent variable that differ only in whether they use a logit link function as opposed to inverse normal link function (Johnston, McDonald, and Quist 2020). The results of the ordered logit in Table A12 are almost identical to those of the ordered probit, as expected.

### 5.6 Multiple imputation

Models 2, 3, 5, and 6 lose some observations due to missing values for control variables; primarily those not available after 2012. Variables with missing data are shown in Figure A9, with all but CINC being used in the models in the main text. We do not use the CINC ratio variable because listwise deletion would lose 25% of our observations in a biased manner given the missingness is for all observations after 2012. Instead, the main model uses population and SIPRI military expenditure variables which adequately proxy for CINC given they are 2 of CINC's 6 components. When imputing missing values, we replace the population and military expenditure variables with CINC since it is more commonly used as an observable indicator for military power, the concept of interest.

	Full sample			Relevant states sample		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Independent Variables						
NATO member	-0.06*** (0.02)	-0.08*** (0.03)	-0.09*** (0.03)	-0.14*** (0.05)	-0.14*** (0.05)	-0.16*** (0.05)
Russia distance	-0.02*** (0.01)	-0.02*** (0.01)	-0.02*** (0.01)	-0.02** (0.01)	-0.02** (0.01)	-0.02*** (0.01)
Controls						
Democracy		-0.10 (0.12)	-0.06 (0.13)		-0.08 (0.13)	-0.02 (0.14)
Nuclear power		0.13*** (0.04)	0.10** (0.05)		0.15*** (0.06)	0.11 (0.10)
Population		-0.00** (0.00)	-0.00** (0.00)		-0.00** (0.00)	-0.00** (0.00)
GDP per cap		0.02 (0.01)	0.01 (0.02)		0.04 (0.03)	0.04 (0.04)
Mil. spending			0.00 (0.00)			0.00 (0.00)
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 A11: OLS Results

	Full sample			Relevant states sample		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Independent Variables						
NATO member	-0.43 (0.50)	-1.01** (0.46)	-1.26** (0.49)	-0.74 (0.54)	-1.14** (0.54)	-1.30** (0.54)
Russia distance	-0.18** (0.08)	-0.20*** (0.07)	-0.22*** (0.07)	-0.08 (0.09)	-0.17** (0.08)	-0.15 (0.10)
Controls						
Democracy		0.76 (0.81)	1.24* (0.75)		0.62 (0.81)	1.10 (0.72)
Nuclear power		1.78** (0.87)	1.07 (0.96)		1.73 (1.08)	2.51 (1.96)
Population		0.47** (0.23)	0.40 (0.29)		0.41* (0.24)	0.49 (0.31)
GDP per cap		-0.02* (0.01)	-0.03* (0.02)		-0.02 (0.02)	-0.02 (0.02)
Mil. spending			0.02 (0.02)			-0.02 (0.04)
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 A12: Ordered Logit Results

Missing values are calculated using bootstrap re-sampling across 10 different imputations using predictive mean matching (Buuren et al. 2006; White, Royston, and Wood 2011). The imputation predictions account for the temporal nature of the data. The same ordinal probit for each model is run across all 10 imputations

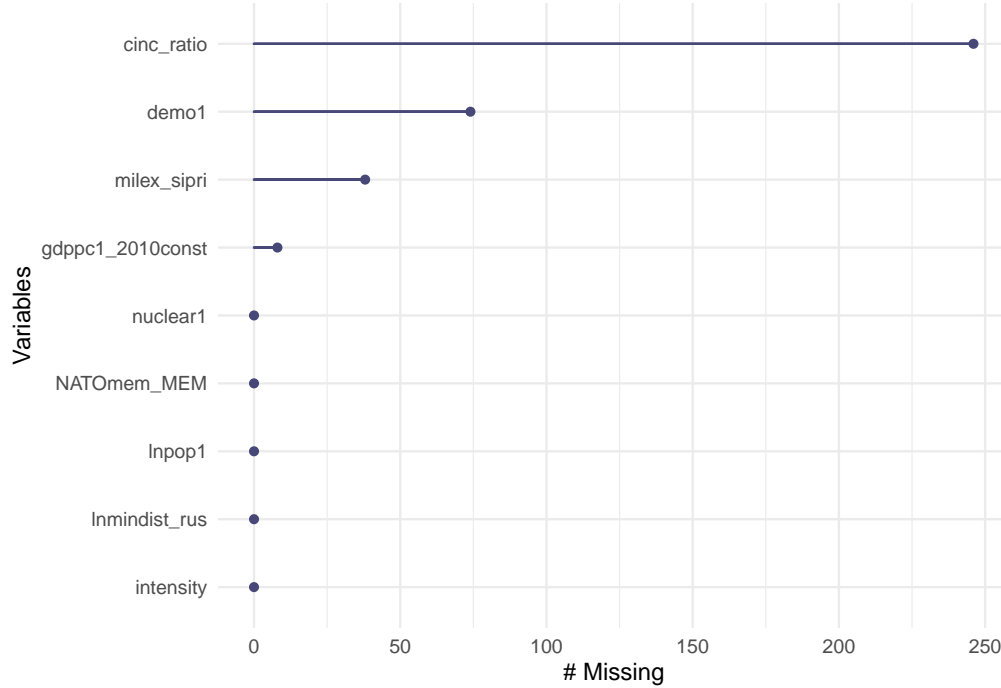


Figure A9: Number of missing observations for each variable in the dataset

and the coefficient estimates and standard errors are pooled across the 10 imputations to account for uncertainty produced by variation across the imputations. Variables not included in the main regression like ethno-linguistic fractionalization, and GDP per capita are included to increase the predictive performance of the imputation for variables that the literature suggests are correlated with the variables being imputed.

The results of the original models with imputed control variables are shown in Figure A10 and Table A13. We show the results for models 1 and 4 separately to make clear that these models have no imputed values. The coefficients in A13 are the same as those reported in models 1 and 4 in the main text and are only reported here to enable comparison with the imputed models in Figure A10.

	Full sample	Relevant states sample
	Model 1	Model 4
NATO member	0.76 [0.39; 1.12]	0.62 [0.21; 1.04]
Russia distance	0.90* [0.84; 0.96]	0.95 [0.89; 1.02]
Observations	1,000	376

All models are ordered probits and include year-fixed effects with country-clustered standard errors in parentheses.

Table A13: Odds Ratios (non-imputed models)

## 5.7 Targeted states sample

We run the same models on a third sample of just targeted states as empirical support for the endogenous bargaining and information asymmetry model extension offered in section 1.6 in the appendix. This includes only country-years that were targets of Russian aggression, meaning the intensity variable is greater than 0. Those results are shown in Table A14 and are consistent with the results produced by the other models.

Descriptive analysis of the same targeted states sample provides consistent results. It appears that higher intensity Russian interventions are associated with non-NATO targets, as shown in Figure A11.

<b>Models with Imputed Control Variables</b>				
	<b>Model 2</b>	<b>Model 3</b>	<b>Model 5</b>	<b>Model 6</b>
<i>Predictors</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>	<i>Odds Ratios</i>
NATO member	0.39 ** (0.16 – 0.94)	0.37 ** (0.16 – 0.87)	0.28 ** (0.09 – 0.85)	0.29 ** (0.11 – 0.79)
Russia distance	0.81 *** (0.72 – 0.92)	0.81 *** (0.71 – 0.91)	0.90 * (0.80 – 1.00)	0.90 * (0.81 – 1.00)
Democracy	2.12 (0.44 – 10.36)	1.93 (0.45 – 8.29)	1.98 (0.32 – 12.32)	1.66 (0.29 – 9.60)
Nuclear power	5.61 ** (1.15 – 27.34)	3.70 * (0.79 – 17.29)	1.95 (0.43 – 8.80)	1.45 (0.24 – 8.72)
GDP per cap	0.98 (0.96 – 1.00)	0.98 * (0.95 – 1.00)	1.00 (0.98 – 1.01)	1.00 (0.98 – 1.01)
Population	1.58 ** (1.04 – 2.42)		1.46 ** (1.02 – 2.10)	
CINC ratio		732.46 *** (12.60 – 42575.63)		120.07 * (0.79 – 18277.02)
Observations	1000	1000	376	376
R <sup>2</sup>	0.251	0.254	0.234	0.232

\*  $p < 0.1$  \*\*  $p < 0.05$  \*\*\*  $p < 0.01$

Figure A10: Imputed control variables results

## Intensity of Russian Interventions (1994–2018)

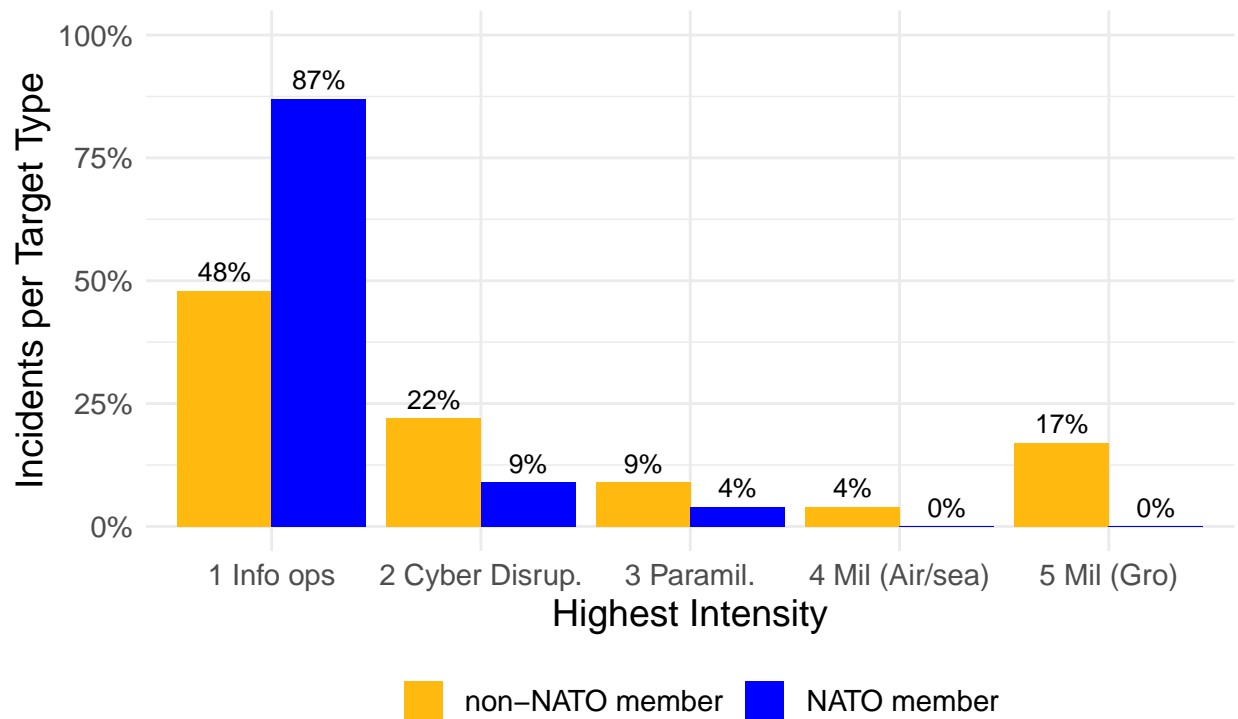


Figure A11: Intensity of Russian Interventions, by Target Type



	Model 1	Model 2	Model 3
Independent Variables			
NATO member	-2.15*** (0.64)	-39.01*** (1.84)	-48.05*** (1.82)
Russia distance	-0.19** (0.08)	-2.95*** (0.07)	-2.19*** (0.05)
Controls			
Democracy		-56.32*** (0.33)	-66.73*** (0.60)
Nuclear power		-2.90	51.08
Population		-13.98*** (0.04)	-12.54*** (0.07)
GDP per cap		0.78*** (0.12)	1.19*** (0.11)
Mil. spending			-1.57*** (0.47)

All models include year-fixed effects. \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$

Table A14: Targeted States Sample

## 6 Case Study: US 2016

The main text presents case studies of Russian interventions in Estonia, Ukraine, and Georgia, which are all contiguous to Russia and former Soviet republics, and thus more comparable. Yet we expect the logic of the argument to apply more generally. Thus Russian intervention should be even more restrained against targets that are further away and more capable. Intervention in the 2016 U.S. election is consistent with this expectation.

A U.S. intelligence assessment released soon after the 2016 election concluded with “high confidence” that “Russian President Vladimir Putin ordered an influence campaign in 2016 aimed at the US presidential election. Russia’s goals were to undermine public faith in the US democratic process, denigrate Secretary Clinton, and harm her electability and potential presidency. We further assess Putin and the Russian Government developed a clear preference for President-elect Trump” (Office of the Director of National Intelligence 2017). Moscow’s influence operations might thus be described as unrestrained, even brazen, and thus motivated entirely by efficiency calculations. Yet the choice to pursue this course of action in the first place was very much constrained by the implicit deterrence posture of the United States. Russia could safely assume that the most powerful military in the world would retaliate for armed attacks against U.S. vital interests. While the United States had not designated its electoral process as “critical infrastructure” to explicitly signal that cyber interference was proscribed, Russia still had to consider the potential for American retaliation. Russia thus sought opportunities to impose costs and seek benefits while minimizing the risk of escalation. It found them through covert manipulation of democratic discourse. Indeed, Russia’s electoral interference has gone essentially unpunished by the United States to date, aside from the expulsion of some Russian intelligence officers and the application of some additional sanctions to an already heavy regime put in place after Ukraine. Of course, if Trump’s victory in 2016 or any of his administration’s subsequent policies can ever be credited to active measures by the Russian Federation, even in part, it would amount to one of the most consequential intelligence coups in history. It is just as likely, however, that the Russian campaign simply added noise to one of the most chaotic campaigns in U.S. presidential history (Gelman and Azari 2017). Russian information operations appear to be a low-cost gamble to influence an over-determined outcome.

## References

- Atland, Kristian, and Ihor Kabanenko. 2020. "Russia and Its Western Neighbours: A Comparative Study of the Security Situation in the Black, Baltic and Barents Sea Regions." *Europe-Asia Studies* 72 (2): 286–313. <https://doi.org/10.1080/09668136.2019.1690634>.
- Beardsworth, James. 2022. "Russia Scrambles for Soldiers Amid Ukraine War Manpower Shortage." *The Moscow Times*. <https://www.themoscowtimes.com/2022/05/23/russia-scrambles-for-soldiers-amid-ukraine-war-manpower-shortage-a77750>.
- Brecher, Michael, and Jonathan Wilkenfeld. 1997. *A Study of Crisis*. University of Michigan Press.
- Buuren, S. Van, J. P. L. Brand, C. G. M. Groothuis-Oudshoorn, and D. B. Rubin. 2006. "Fully Conditional Specification in Multivariate Imputation." *Journal of Statistical Computation and Simulation* 76 (12): 1049–64. <https://doi.org/10.1080/10629360600810434>.
- Casey, Adam, and Lucan Ahmad Way. 2017. "Russian Electoral Interventions, 1991-2017." Scholars Portal Dataverse. <https://doi.org/10.5683/SP/BYRQQS>.
- Christie, Edward Hunter. 2015. "NATO Review - Sanctions After Crimea: Have They Worked?" *NATO Review*. <https://www.nato.int/docu/review/articles/2015/07/13/sanctions-after-crimea-have-they-worked/index.html>.
- "Conflict in Ukraine's Donbas: A Visual Explainer." 2021. <https://www.crisisgroup.org/content/conflict-ukraines-donbas-visual-explainer>.
- Dress, Brad. 2022. "Russia Using Separatist Forces in Eastern Ukraine, UK Says." Text. *The Hill*.
- Eichensehr, Kristen E. 2022. "United States and Allies Target Russia and Belarus with Sanctions and Other Economic Measures." *American Journal of International Law* 116 (3): 6140631.
- Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49 (3): 379–414. <https://doi.org/10.1017/S0020818300033324>.
- Foster, Hal, and Tatyana Goryachova. 2014. "Ukraine's Next Battle Is Donetsk, but No Bombs, Please." *USA TODAY*. <https://www.usatoday.com/story/news/world/2014/07/12/ukraine-donetsk-separatists/12506719/>.
- Garamone, Jim. 2022. "Russian Efforts to Raise Numbers of Troops 'Unlikely to Succeed,' U.S. Official Says." *U.S. Department of Defense*.
- Gauthier-Villars, David, Steve Stecklow, Maurice Tamman, and Andrew Macaskill. 2022. "As Russian Missiles Struck Ukraine, Western Tech Still Flowed." *Reuters*, August.
- Gelman, Andrew, and Julia Azari. 2017. "19 Things We Learned from the 2016 Election." *Statistics and Public Policy* 4 (1): 1–10. <https://doi.org/10.1080/2330443X.2017.1356775>.
- Holland, Steve. 2022. "Russia's Military Suffering Manpower Shortages -U.S. Intelligence." *Reuters*, August.
- Ilyushina, Mary, and Annabelle Timsit. 2022. "What Does Putin's Partial Military Mobilization Mean for Russia and Ukraine?" *Washington Post*, September.
- Johnston, Carla, James McDonald, and Kramer Quist. 2020. "A Generalized Ordered Probit Model." *Communications in Statistics - Theory and Methods* 49 (7): 1712–29. <https://doi.org/10.1080/03610926.2019.1565780>.
- Klinova, M. V., and E. A. Sidorova. 2019. "Economic Sanctions of the West Against Russia: Development of the Situation." *Studies on Russian Economic Development* 30 (3): 355–64.
- Lanoszka, Alexander. 2020. "Thank Goodness for NATO Enlargement." *International Politics* 57: 451–70. <https://doi.org/10.1057/s41311-020-00234-8>.
- Leicester, John, and Hanna Arhinova. 2022. "Russia Almost Fully in Control of Donbas Province, Now Holds 20 Percent of Ukraine." *PBS NewsHour*. <https://www.pbs.org/newshour/world/russia-almost-fully-in-control-of-donbas-province-now-holds-20-percent-of-ukraine>.
- Lokshina, Tanya. 2017. "Grad Rockets Return to Eastern Ukraine?" *Human Rights Watch*.
- Luzin, Pavel. 2022. "Russia's Military Manpower Crunch Will Worsen." *CEPA*. <https://cepa.org/article/russias-military-manpower-crunch-will-worsen/>.
- Macias, Amanda, and Gabriel Cortes. 2022. "A Look at the Black Sea Grain Initiative, the UN-backed Deal That Helped Reopen Ukraine's Ports to Feed Countries Around the World." *CNBC*. <https://www.cnn.com/2022/09/22/infographic-showing-black-sea-initiative-crops.html>.
- Miller, Christopher, and Paul McLeary. 2022. "Ukraine Has Hobbled Russia's Black Sea Fleet. Could It Turn the Tide of the War?" *POLITICO*. <https://www.politico.com/news/2022/08/29/russias-black-sea-fleet-stuck-struck-and-sinking-00054114>.

- Moravcsik, Andrew. 2019. "Power of Connection: Why the Russia–Europe Gas Trade Is Strangely Untouched by Politics." *Nature* 576 (7785): 30–31. <https://doi.org/10.1038/d41586-019-03694-y>.
- Nelson, Eshe. 2022. "Bucking the Global Trend, Russia Lowers Interest Rates Again." *The New York Times*, July.
- Neuman, Scott. 2020. "Nearly 6 Years After MH17 Was Shot Down, Dutch Prosecutors Say They Will Sue Russia." *NPR*, July.
- Office of the Director of National Intelligence. 2017. "Assessing Russian Activities and Intentions in Recent US Elections." Intelligence {{Community Assessment}} ICA 2017-01D. Washington, DC: National Intelligence Council.
- Ostroukh, Andrey, and Alexander Marrow. 2022. "Russian Central Bank Trims Key Rate to 7.5%, Says Easing Could Be over Soon." *Reuters*, September.
- Ostrovsky, Simon, and Yegor Troyanovsky. 2022. "Russia Ramps up Long-Range Strikes as Ukraine Gets Precision Weapons from the West." *PBS NewsHour*. <https://www.pbs.org/newshour/show/russia-ramps-up-long-range-strikes-as-ukraine-gets-precision-weapons-from-the-west>.
- Petrenko, Victoria. 2017. "Kyiv Says There Are about 6,000 Russian Soldiers, 40,000 Separatists in Donbas." *Kyiv Post*. <https://www.kyivpost.com/ukraine-politics/kyiv-says-6000-russian-soldiers-40000-separatists-donbas.html>.
- Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics*. Princeton, N.J: Princeton University Press.
- Reuters. 2022. "Russia Says over 200,000 Drafted into Army Since Putin's Decree." *Reuters*, October.
- "Russia and the Separatists in Eastern Ukraine." 2016. Briefing 79. Kyiv/Brussels: International Crisis Group.
- Schatz, Valentin J., and Dmytro Koval. 2019. "Russia's Annexation of Crimea and the Passage of Ships Through Kerch Strait: A Law of the Sea Perspective." *Ocean Development & International Law* 50 (2-3): 275–97. <https://doi.org/10.1080/00908320.2019.1605677>.
- Schram, Peter. 2021. "Hassling: How States Prevent a Preventive War." *American Journal of Political Science* 62 (2): 294–308. <https://doi.org/10.1111/ajps.12538>.
- Schuetze, Christopher F. 2022. "Europe Says It Can Weather the Winter as Russia Postpones Resuming Gas Flow." *The New York Times*, September.
- Shiffrinson, Joshua R. Itzkowitz. 2016. "Deal or No Deal? The End of the Cold War and the U.S. Offer to Limit NATO Expansion." *International Security* 40 (4): 7–44. [https://doi.org/10.1162/ISEC\\_a\\_00236](https://doi.org/10.1162/ISEC_a_00236).
- Stepanenko, Kateryna, Frederick W. Kagan, and Brian Babcock-Lumish. 2022. "Institute for the Study of War." Institute for the Study of War.
- Strzelecki, Jan, Tadeusz Iwanski, and Piotr Zochowski. 2018. "The Leader of the Donetsk Separatists Is Assassinated." *OSW Centre for Eastern Studies*. <https://www.osw.waw.pl/en/publikacje/analyses/2018-09-03/leader-donetsk-separatists-assassinated>.
- Troianovski, Anton, and Ivan Nechepurenko. 2022. "McDonald's Is Back, Moscow Style, as Russian Economy Stumbles On." *The New York Times*, June.
- "Ukraine Crisis: 'Thousands of Russians' Fighting in East." 2014. *BBC News*, August.
- Valeriano, Brandon, and Ryan C Maness. 2014. "The Dynamics of Cyber Conflict Between Rival Antagonists, 2001–11." *Journal of Peace Research* 51 (3): 347–60. <https://doi.org/10.1177/0022343313518940>.
- Vorobyov, Niko. 2022a. "Ukraine Crisis: Who Are the Russia-backed Separatists?" *Al Jazeera*. <https://www.aljazeera.com/news/2022/2/4/ukraine-crisis-who-are-the-russia-backed-separatists>.
- . 2022b. "Why Russia Recognised Separatist-Held Regions in East Ukraine." *Al Jazeera*, February.
- "Western Sanctions Will Eventually Impair Russia's Economy." 2022. *The Economist*, August.
- White, Ian R., Patrick Royston, and Angela M. Wood. 2011. "Multiple Imputation Using Chained Equations: Issues and Guidance for Practice." *Statistics in Medicine* 30 (4): 377–99. <https://doi.org/10.1002/sim.4067>.
- Wroughton, Lesley. 2014. "U.S. Says Russia Sends Tanks, Rocket Launchers into Ukraine." *Reuters*, June.