Appendix For "Conflicts that Leave Something to Chance"

January 15, 2025

Contents

Part I

Comments on Model Assumptions

1 Connections to Powell (2015)

This section expands on the "Comments on Model Assumptions" section in the main paper by elaborating on the similarities and differences between this paper and Powell.

In this paper, nuclear risk in a conventional conflict is determined indirectly through the defender's arming level. In contrast, in Powell, the defender is able to directly, publicly, and credibly manipulate the level of nuclear risk within a conventional war without altering its likelihood of winning in the conventional war.[1] A natural interpretation of the defender's choice in the Powell model would be manipulating Defense Readiness Condition (DEFCON) levels while in a conflict, In contrast, in the model presented here, the defender's arming level indirectly shapes the likelihood of a nuclear exchange by generating longer or shorter conflicts. While the model of Powell is groundbreaking, it includes several strong assumptions that may not apply to all settings. For example, Powell assumes it is possible to publicly and credibly manipulate the likelihood of a nuclear exchange within a crisis. Practically, doing so would be subjected to "cheap talk" concerns, as a defender may want to signal that they have implemented a high-risk system (that is, one with a high risk of nuclear exchange) to deter challengers when, in reality, they have not. Placing conventional forces, as the defender does in my model, is more visible and less subject to this kind of bluffing. As another example, Powell assumes it is possible to exclusively manipulate nuclear risk without altering the balance of conventional power. This puts the model in (Powell, 2015) outside of most standard discussions of tripwires (see Schelling (1966)). As Schelling describes, the positioning of conventional forces in Western Berlin is a tripwire in that they bid-up nuclear risk, but such force placement would mechanically alter the likelihood of conventional conflict success.

Second, these models also differ in how conventional forces generate nuclear risk. In Powell, whenever one actor adds additional conventional forces to a conflict, this *always* leads to a greater risk of escalation. Unlike the model presented here, Powell does not consider the possibility that undertaking a rapid and decisive deployment could potentially reduce the probability of a nuclear exchange by preventing a protracted affair.

The modeling choices in this paper were made in an attempt to best represent our scholarly understanding of conflicts over non-existential issues in the nuclear era. In adopting these

[1]Formally, after a challenger selects some level of arming, p, the defender selects function $r(p)$. p denotes the likelihood the C wins in a conventional conflict, and $r(p)$ denotes the likelihood the conventional conflict escalates to a nuclear exchange.

choices, I establish a series of results suggesting that actors may arm more (Remark 1), fight harder (Remark 4), and signal more effectively (Remark 7) than what is established in Powell (2015), and below I describe how these insights offer more nuance into substantive cases in the nuclear era. However, some readers may prefer some of the modeling choices made in Powell (2015), and I address these concerns while preserving the defender's choice in selecting a force posture in the following way. First, if readers prefer assuming that nuclear risk is monotonically increasing in the selected conventional force posture, my model can support this assumption;[2] all remarks below will still attain, excluding Remark 4, which I discuss more below. Second, if readers believe that the defender can also manipulate nuclear risk, I discuss in an extension where the defender manipulates both conventional arming and nuclear risk; again, similar results attain. Lastly, while the results here are robust to modifications to make the model more like that in Powell, these results are distinct from what is presented in Powell).

Part II

Complete Information Game

2 Complete Information Equilibrium

2.1 Deriving p^C and p^D

Outside of p^C and p^D, the equilibrium follows from construction. I first derive p^C, the force postures that would make C willing to challenge, conditional on D escalating in stage 4

$$0 \leq -\frac{n}{h}N_C + \frac{\alpha}{hp(1-p)}\left((1-p)v_C\right) - \frac{c_C}{h}$$

$$0 \leq -npN_C + \alpha v_C - c_C p$$

$$p \leq \frac{\alpha v_C}{c_C + nN_C}$$

This means that if D arms to level $p = p^C = \frac{\alpha v_C}{c_C + nN_C}$, C is indifferent between dropping out or not. In our equilibria, to prevent open set issues, whenever D arms to $p = p^C$, C will be deterred and will not challenge.

[2] This can be accommodated by shifting the p_0 and p_1 parameters to only consider regions where the relationship is monotonic or non-monotonic.

Next I derive p^D as the force posture that would make a type v_D D willing to escalate conditional on C challenging. The left-hand-side is D's payoffs (sans arming costs, which are sunk) if D does not escalate, and the right is D's payoff from fighting (sans arming costs).

$$0 \leq \frac{n}{h} * (-N_D) + \frac{\alpha}{hp(1-p)} (pv_D) - \frac{c_D}{h}$$

$$0 \leq -n(1-p)N_D + \alpha v_D - c_D(1-p)$$

$$p \geq 1 - \frac{\alpha v_D}{c_D + nN_D}.$$

This means that if D arms to level $p = p^D(v_D)$, D is indifferent between escalating or not.

3 Proving the Proposition and Remarks

3.1 Proposition 1, Remarks 1, 2, and 5 Proofs

Proposition 1 follows all player's best responses in a straitforward manner.

All remark follow from the equilibrium and the derivation of p^D and p^C.

3.2 Remark 3 Proof

3.2.1 Case 1: For n'', $p^C \leq p^D$

If for n'' $p^C \leq p^D$ holds, then under n'', war is not possible because there is no arming level where C would be willing to challenge and D would be willing to fight. For all parameters where $p^C(n'') \leq p^D(n'')$, the likelihood of war is weakly decreasing as n' shifts to n''.

3.2.2 Case 2: For n'', $p^C > p^D$

This proof is assisted by a helpful Lemma that applies to a subset of the parameter space within Case 2. When D is optimally choosing to fight, D selects some arming level p within the set S, where $S = \left[max\left\{p_0, p^D\right\}, min\left\{p^C, p_1\right\}\right]$. Intuitively, the set S defines feasible arming levels where D will fight if challenged, and C will not be deterred. Note that we will consider two levels of nuclear instability parameter n, which we denote n and n' (with $n < n'$). As defined, $S(n') \subset S(n)$.[3]

I introduce some new notation here. I let $\hat{U}(p,n) = -\frac{p(1-p)}{\alpha+np(1-p)}(nN_D + c_D) + \frac{\alpha}{\alpha+np(1-p)}(pv_D) -$

[3]Recall $p^C = \frac{\alpha v_C}{c_C+nN_C}$ and $p^D = 1 - \frac{\alpha v_D}{c_D+nN_D}$.

$K(p)$. I also define $p^*(a,b)$ as

$$p^*(a,b) \in argmax_{p\in S(a)} \hat{U}(p,b)$$

Note that whenever $a = b = n$, this is D optimizing an arming level at nuclear instability parameter n.[4]

Whenever D (optimally) selects a p and goes to war, I define D's value function as

$$\hat{V}_D(n) = max_{p\in S(n)} \hat{U}(p,n)$$

This allows us to set up a useful Lemma.

Nuclear Instability and War Lemma *$\hat{V}_D(n)$ is decreasing in n.*

Proof: With this structure in place, I can show that $\hat{V}_D(n') \leq \hat{V}_D(n)$. The proof proceeds as follows: $\hat{V}_D(n') = max_{p\in S(n')}\hat{U}(p,n') \leq max_{p\in S(n)}\hat{U}(p,n') \leq \hat{U}(p^*(n,n'),n) \leq max_{p\in S(n)}\hat{U}(p,n) = \hat{V}_D(n)$

The first inequality holds because $S(n') \subset S(n)$, meaning $\hat{U}$ is optimized over a smaller set under n'. The second inequality holds because $\hat{U}(p,n)$ is decreasing in n at a fixed arming level $p^*(n,n')$.[5] The third inequality holds because there D is selecting their optimal p. □

The Lemma above shows that as n increases, D recieves a lower utility from going to war. Note that if for n'' $p^C > p^D$, then it also must be that $p^C > p^D$ for n'. This means that for D to deter C through force posture, under both n' and n'', D will set p^C, which is decreasing in n. Together, this means that as n increases, D's utility from war (setting $\hat{p}$) is decreasing, D's utility from deterring (setting p^C) is increasing, and D's utility from arming then acquiescing (setting p_0) remains the same. This means that as n increases, D will arm with the intent of fighting weakly less. □

3.3 Remark 4 Proof

In the statement of Remark 4 in the text, I reference two conditions. First, the assumption that the solution set $\hat{p}$ is singleton matters because this means that I do not need to define every solution as satisfying a relevant condition. Second, the assumption that both actors place high value on the issue is in place to keep the bounds on p^D and p^C—which move inwards as n

[4]Note that we abuse notations and sometimes let this denote a set of arming levels; when this is the case, the proof functions for all individual elements of the set $p^*(a,b)$. Note, we need to keep this separate as part of the proof below.

[5]Taking first order conditions of $\hat{U}(p, v_D, n)$ with respect to n yields $\frac{(p-1)p(\alpha N_D + \alpha p \bar{v}_D - p(1-p)c_D)}{\left(-\alpha + n(p)^2 - sp\right)^2}$. Note that $p - 1 < 0$ and, because $p \geq p^D(v_D)$, we can say $0 \leq -n(1-p)N_D + \alpha\bar{v}_D - c(1-p)$.

and N_D increases—away from influencing the direct effect of these parameters on D's preferred arming choice.

The complete, more technical statement of the Remark is as follows.

Remark 4A. *Formally, consider nuclear cost parameters N'_D and N''_D, and nuclear instability parameters n' and n'', where $N'_D < N''_D$ and $n' < n''$.*

(a) If $p^*(N'_D) \leq \frac{1}{2}$, $p^*(N''_D) \leq \frac{1}{2}$ and $p^C > \frac{1}{2}$, then $p^*(N'_D) \geq p^*(N''_D)$.[6] *And if $p^*(N'_D) \geq \frac{1}{2}$, $p^*(N''_D) \geq \frac{1}{2}$, and $p^D(N''_D) < \frac{1}{2}$, then $p^*(N'_D) \leq p^*(N''_D)$.*

(b) If $p^*(n')$ and $p^*(n'')$ are small enough, and $p^C(n'') > \frac{1}{2}$,[7] *then $p^*(n') \geq p^*(n'')$. And if $p^*(n')$ and $p^*(n'')$ are large enough, $p^D(n'') < \frac{1}{2}$, and $p^C(n'') \geq p_1$, then $p^*(n') \geq p^*(n'')$.*

3.3.1 For N_D. Part A. Analyzing the Objective Function

Proving (a). Consider a the solution to D's optimization problem. This is

$$\hat{p} \in \underset{p \in [max\{p^D, p_0\}, min\{p^C, p_1\}]}{\arg\max} \left\{ -\frac{np(1-p)}{\alpha + np(1-p)} N_D + \frac{\alpha}{\alpha + np(1-p)} (pv_D) - \frac{c_D p(1-p)}{\alpha + np(1-p)} - K(p) \right\}.$$

First, note that the objective function exhibits decreasing differences in N_D and p when $p < 1/2$ and increasing differences in N_D and p when $p > 1/2$. Letting $N_D < N'_D$ and $p < p'$, this experiences increasing differences when

$$-\frac{np'(1-p')}{\alpha + np'(1-p')} N'_D - \left(-\frac{np(1-p)}{\alpha + np(1-p)} N'_D \right) > -\frac{np'(1-p')}{\alpha + np'(1-p')} N_D - \left(-\frac{np(1-p)}{\alpha + np(1-p)} N_D \right)$$

or

$$\frac{np'(1-p')}{\alpha + np'(1-p')} N_D - \frac{np(1-p)}{\alpha + np(1-p)} N_D > \frac{np'(1-p')}{\alpha + np'(1-p')} N'_D - \frac{np(1-p)}{\alpha + np(1-p)} N'_D$$

or more simply

$$(N_D - N'_D) \left(\frac{np'(1-p')}{\alpha + np'(1-p')} - \frac{np(1-p)}{\alpha + np(1-p)} \right) > 0.$$

The term $N_D - N'_D$ is negative. The expression $\frac{np'(1-p')}{\alpha+np'(1-p')} - \frac{np(1-p)}{\alpha+np(1-p)}$ is (weakly) negative so long that $p \geq 1/2$.[8] The expression is weakly positive so long that $p \leq 1/2$.

[6] As defined earlier, $\hat{p}$ may represent the minimum value of a set. Here, when I say $p^*(N'_D) \leq p^*(N''_D)$, I abuse notation and assume that every element of both $p^*(N'_D)$ and $p^*(N''_D)$ is less than or equal to $\frac{1}{2}$.

[7] We will clarify "small enough" and "large enough" in the appendix.

[8] Can be seen by taking the cross partial derivative, or $\frac{\partial^2}{\partial p \partial N_D} \frac{np(1-p)}{\alpha+np(1-p)} = \frac{\alpha n(2p-1)}{(n(p-1)p-\alpha)^2}$.

3.3.2 For N_D. Part B. Proving for $p^*(N_D) \leq \frac{1}{2}$ *and* $p^*(N_D') \leq \frac{1}{2}$

By Assumption, $p^*(N_D) \in [p_0, \frac{1}{2}]$ and $p^*(N_D') \in [p_0, \frac{1}{2}]$, with $p^C > \frac{1}{2}$. We demonstrate over the range

$$U_D(p; N_D) = \begin{cases} 0 - K(p) & if\, p < p^D \\ -\frac{np(1-p)}{\alpha+np(1-p)} N_D + \frac{\alpha}{\alpha+np(1-p)}(pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)} - K(p) & if\, p^D \leq p \leq \frac{1}{2} \end{cases}$$

the utility function has decreasing differences in p and N_D. In the Proof of Lemma 1 (see the proof of incomplete information), we demonstrated there are no open set issues to the optimization problem. Also, below we will reference "Regions." Region 1 is is any $p < p^D$, Region 2 is any $p^D \leq p \leq \frac{1}{2}$. Recall by assumption $p^C > \frac{1}{2}$.

I write out every case that I must consider, as characterized by what Region of the utility function that the considered p or p' and N_D or N_D' put the function into. Note that there is some structure to the cases that I consider; for example, if (p, N_D') puts the utility function into Region 1, then (p, N_D) must also fall within Region 1; similarly, if (p', N_D') puts the utility function into Region 1, then (p', N_D) must also fall within Region 1.

Cases	$U_D(p'; N_D')$	$U_D(p; N_D')$	$U_D(p'; N_D)$	$U_D(p; N_D)$
A	1	1	2	1
B	1	1	2	2
C	2	1	2	2
D	2	1	2	1
E	1	1	1	1
F	2	2	2	2

It is useful to describe several properties that will be used in the proofs below.

Property (a): If $p \geq p^D$, then $-\frac{np(1-p)}{\alpha+np(1-p)} N_D + \frac{\alpha}{\alpha+np(1-p)}(pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)} \geq 0$.[9]

Property (b): if $p \geq p^D$, then $-\frac{np(1-p)}{\alpha+np(1-p)} N_D + \frac{\alpha}{\alpha+np(1-p)}(pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)}$ is increasing in p.[10]

Property (c): I abuse notation and (sometimes below will) bring in the region numbers to the utility function, letting $U_D(p; 1) = -K(p)$ and $U_D(p; 2) = -\frac{p(1-p)}{\alpha+np(1-p)}(nN_D + c_D) + \frac{\alpha}{\alpha+np(1-p)}(pv_D) -$

[9]This holds based on how p^D is defined: when $p \geq p^D$, then D is willing to fight and attain utility $-\frac{np(1-p)}{\alpha+np(1-p)} N_D + \frac{\alpha}{\alpha+np(1-p)}(pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)}$ over acquiesce and attain utility 0.

[10]Taking first order conditions gives $\frac{d}{dp}\left(-\frac{p(1-p)}{\alpha+np(1-p)}(nN_D + c_D) + \frac{\alpha}{\alpha+np(1-p)}(pv_D)\right) = \frac{\alpha(2p-1)(c_D+nN_D)+\alpha v_D(\alpha+np^2)}{(\alpha-n(p-1)p)^2}$, or equal to $\frac{\alpha p(c_D+nN_D)+\alpha(1-p)(-c_D-nN_D)+\alpha v_D(\alpha+np^2)}{(\alpha-n(p-1)p)^2}$. The right-hand side will be positive whenever $-(1-p)(c_D + nN_D) + v_D(\alpha + np^2) \geq 0$, which will hold by Property (a).

$K(p)$ regardless of p's relationship to p^D or p^C; for example, I will let $U_D(p^C, 1) = -K(p^C)$. If $p < p^D(N_D)$, then $U_D(p; N_D, 2) < U_D(p; N_D, 1)$ (because p is fixed).

I now describe how decreasing differences $(U_D(p', N_D') - U_D(p, N_D') \leq U_D(p', N_D) - U_D(p, N_D))$ occurs across all cases listed above.

Case A. Property (a) implies $U_D(p'; N_D') \leq U_D(p'; N_D)$. Also, $U_D(p, N_D') = U_D(p, N_D)$. This case exhibits decreasing differences.

Case B. Note that $-(U_D(p'; N_D') - U_D(p; N_D')) - K(p') + K(p) = 0$. Thus, Property (b) implies $U_D(p'; N_D) - U_D(p; N_D) - (U_D(p'; N_D') - U_D(p; N_D')) \geq 0$. Re-arranging this term implies that the case exhibits decreasing differences.

Case C. Because the objective function exhibits decreasing differences, we have $U_D(p'; N_D') - U_D(p; N_D', 2) \leq U_D(p'; N_D) - U_D(p; N_D)$. Applying Property (c) implies that this case exhibits decreasing differences.

Case D. We have $U_D(p; N_D') = U_D(p; N_D)$. And, because $U_D(p; N_D)$ is decreasing in N_D, we have $U_D(p'; N_D') < U_D(p'; N_D)$. Thus, the case exhibits decreasing differences.

Case E. $U_D(p'; N_D') = U_D(p'; N_D)$ and $U_D(p; N_D') = U_D(p; N_D)$. This case holds trivially.

Case F. The objective function in this region exhibits decreasing differences.

We have now shown that the utility function in this region exhibits decreasing differences in p and N_D. Via Topkis Theorem, we can say the optimal choice correspondence is decreasing.

3.3.3 For N_D. Part C. Proving for $p^*(N_D) \geq \frac{1}{2}$ *and* $p^*(N_D') \geq \frac{1}{2}$

By Assumption, $p^*(N_D) \in [\frac{1}{2}, p_1]$ and $p^*(N_D') \in [\frac{1}{2}, p_1]$ with $p_D(v_D) < \frac{1}{2}$. We demonstrate over this range,

$$U_D(p; N_D) = \begin{cases} -\frac{np(1-p)}{\alpha+np(1-p)} N_D + \frac{\alpha}{\alpha+np(1-p)}(pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)} - K(p) & if\, \frac{1}{2} \leq p < p^C \\ v_D - K(p) & if\, p^C \leq p \end{cases}$$

the utility function has increasing differences in p and N_D. In the Proof of Lemma 1, we demonstrated there are no open set issues to the optimization problem. Also, below we will reference "Regions." Region 2 is any $p^D(v_D) \leq p < p^C$. Region 3 is any $p \geq p^C$.

Note that both regions of the utility function exhibit increasing differences when $U_D(p'; N_D')$, $U_D(p; N_D')$, $U_D(p'; N_D)$, and $U_D(p; N_D)$ are entirely in Region 2 or Region 3. And, because p^C is unchanging in N_D, we only need to consider the following case.

Cases	$U_D(p';N_D')$	$U_D(p;N_D')$	$U_D(p';N_D)$	$U_D(p;N_D)$
E	3	2	3	2

We can show that $U_D(p',N_D') - U_D(p,N_D') \geq U_D(p',N_D) - U_D(p,N_D)$. Note that $U_D(p';N_D') = U_D(p';N_D)$. Also note that in Region 2 $U_D(\cdot;N_D)$ is decreasing in N_D, meaning $U_D(p;N_D') \leq U_D(p;N_D)$. Thus, this region exhibits increasing differences.
We have now shown that the utility function in this region exhibits increasing differences in p and N_D. Via Topkis Theorem, we can say the optimal choice correspondence is increasing.

3.3.4 For n. Part A. Analyzing the Objective Function

For n, effects in Remark 4 are less precise (than it was for N_D) but still partially present. I can turn my attention to properties of the expression $\hat{G}_D = -\frac{np(1-p)}{\alpha+np(1-p)}N_D + \frac{\alpha}{\alpha+np(1-p)}(pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)}$, which, unlike K, is twice continuously differentiable. I take the cross partial of $\hat{G}_D$, giving

$$\frac{\partial^2}{\partial p \partial n}\hat{G}_D = \frac{\alpha(2p-1)(\alpha N_D - (1-p)p(2c_D + nN_D)) + \alpha p v_D\left(\alpha(3p-2) - n(1-p)p^2\right)}{\left(\alpha + n(1-p)p\right)^3}$$

This expression is ugly, but consider when $p_0 \approx 0$ and $p_1 \approx 1$. Taking the limits and eliminating terms that obviously go to zero yields:

$$lim_{p\to 0}\left[\frac{\partial^2}{\partial p \partial n}\hat{G}_D\right] = \frac{-(\alpha^2 N_D)}{\alpha^3}$$

$$lim_{p\to 1}\left[\frac{\partial^2}{\partial p \partial n}\hat{G}_D\right] = \frac{\alpha^2 N_D + \alpha^2 v_D}{\alpha^3}$$

While this is not nearly as clean as the N_D expression, but clearly here when p^* is close to zero for a fixed set of parameters, the $\hat{G}_D$ exhibits decreasing differences. And, when p^* is close to 1 for a fixed set of parameters, then $\hat{G}_D$ exhibits increasing differences.

I will proceed as follows. I'm going to refer to $\underline{p}$ as the upper-bound on the region where $\hat{G}_D$ only experiences decreasing differences. And, I will refer to $\bar{p}$ as the lower-bound on the region where $\hat{G}_D$ only experiences increasing differences.

Additionally, we want to show that $\hat{G}_D$ is decreasing in n. Taking only first order conditions yields

$$\frac{\partial}{\partial n}\hat{G}_D = \frac{(p-1)p\left(\alpha(N_D + pv_D) - (1-p)c_D p\right)}{\left(\alpha - n(p-1)p\right)^2}$$

Recall for any $p \geq p^D(v_D)$, it must be that $0 \leq -n(1-p^D)N_D + \alpha\bar{v}_D - c_D(1-p^D)$. This implies that the expression $(\alpha(N_D + pv_D) - (1-p)c_D p)$ is positive, meaning that $(p-1)$ times the expression is negative. Thus, for the range we are considering $\hat{G}_D$ is decreasing in n.

3.3.5 For n. Part B. Proving for $p^*(n) \leq \underline{\mathbf{p}}$ and $p^*(n') \leq \underline{\mathbf{p}}$

By Assumption, $p^*(n) \in [p_0, \underline{p}]$ and $p^*(n') \in [p_0, \underline{p}]$, with $p^C > \underline{p}$. We demonstrate over the range

$$U_D(p; N_D) = \begin{cases} 0 - K(p) & if\, p < p^D \\ -\frac{np(1-p)}{\alpha+np(1-p)}N_D + \frac{\alpha}{\alpha+np(1-p)}(pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)} - K(p) & if\, p^D \leq p \leq \underline{p} \end{cases}$$

the utility function has decreasing differences in p and n. In the Proof of Lemma 1, we demonstrated there are no open set issues to the optimization problem. Also, below we will reference "Regions." Region 1 is is any $p < p^D$, Region 2 is any $p^D \leq p \leq \underline{p}$.

I write out every case that I must consider, as characterized by what Region of the utility function that the considered p or p' and n or n' put the function into.

Cases	$U_D(p'; n')$	$U_D(p; n')$	$U_D(p'; n)$	$U_D(p; n)$
A	1	1	2	1
B	1	1	2	2
C	2	1	2	2
D	2	1	2	1

The proof for this part is nearly identical to the proof of showing the $p^*(N_D)$ was decreasing in the lower region. For that reason, I exclude this part of the proof.

Part C. For n. Part c. Proving for $p^*(n) \geq \bar{p}$ *and* $p^*(n') \geq \bar{p}$ In this case, $p^*(n)$ and $p^*(n')$ always fall within the range where $\bar{p} \leq p \leq p_1 < p^C$; therefore, because this region exhibits increasing differences, $p^*(n) \leq p^*(n')$.

3.4 Expanding Remarks 1 and 2

Until this point I have kept the discussion of Remarks 1 and 2 brief, highlighting that increasing n can have dual effects on arming and welfare. More precision is possible.

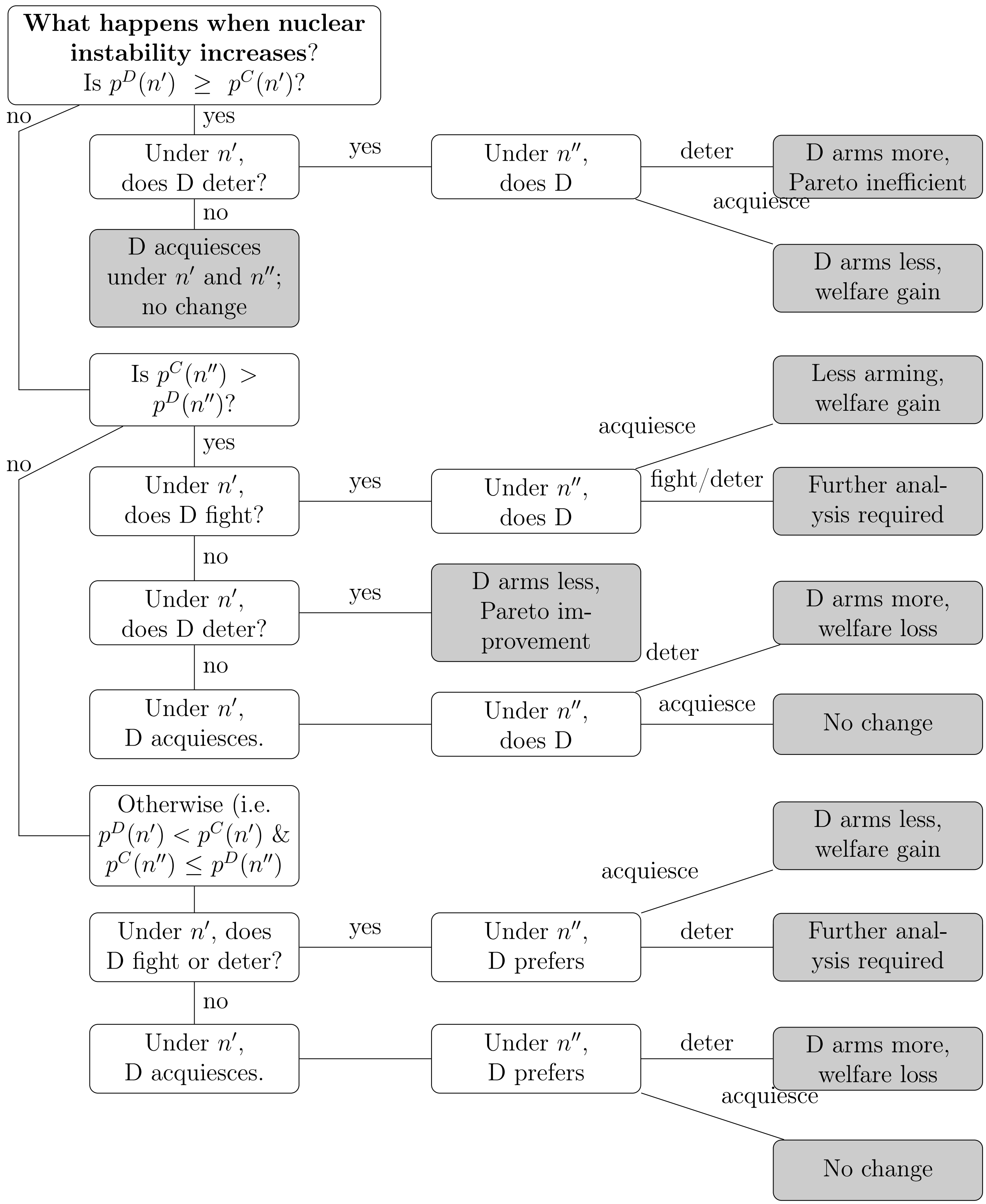

Figure 1: As n' shifts to n'' (where $n' < n''$), what happens to D's arming level and overall welfare? This analysis assumes that v_{DC} such that a transfer of the asset from one actor to the other can never be a welfare gain.

Figure 1 describes the effect of increasing nuclear instability, or formally, shifting $n = n'$ to $n = n''$, where $n' < n''$. This flowchart works by starting in the upper-left box and following the arrows based on what conditions hold. The terminal nodes (gray shading) describe the final effects of the change from n' to n'' when the conditions along the flow hold. Within this analysis, I assume that v_D and v_C are such that D's transferring the asset in dispute to C does not constitute a welfare gain or loss; any loss of welfare here comes through actors undertaking inefficient actions (arming or war) or doing worse within these inefficient actions (arming more, or doing worse in war).

To walk through how to use this flowchart, suppose under n' the condition $p^D \geq p^C$ holds. The next relevant question is how D behaves under n'. If D optimally deters under n' and also optimally deters under n'', then this is the scenario described in the discussion on Remark 1: as nuclear instability increases, D must arm more to achieve deterrence, which constitutes a welfare loss (or more specifically, the shift from n' to n'' is Pareto inefficient). In contrast, if $p^D \geq p^C$, D deters under n', and D acquiesces under n'', then the shift from n' to n'' is welfare-improving (because D is no longer arming), though D does worse (because D no longer gets the asset). Finally, if $p^D \geq p^C$ and D does not deter under n', then D's only other option is to acquiesce under n'. If D acquiesces under a lower level of nuclear instability (n'), D will also acquiesce under a higher level of nuclear instability (n''), meaning welfare and arming will not change.

Depending on the conditions and what D prefers under n' and n'', changing n can induce a wide range of changes in arming behaviors and welfare outcomes. While in some cases (labeled "Further analysis required") I cannot definitively say whether arming or welfare changes, the model makes specific predictions on arming and welfare for a wide range of parameters.

4 Figure Parameters

Figure 2: The cost function is $k * (p^* - p_0)/(p_1 - p^*)$ (but note that is not plotted). The non-illustrated parameter values are $c_D = 2$, $c_C = 1.5$, $N_C = 60$, $N_D = 40$, $\alpha = 0.1$, $n = 0.04$, $k = 8$, $\pi = 0.8$, $\overline{v}_D = 18$.

Figure 3: The cost function is $k*(p^*-p_0)^2$. The non-illustrated parameter values are $p_0 = 0.001$, $p_1 = 0.9$, $c_D = 6$, $c_C = 1.5$, $N_C = 30$, $N_D = 40$, $\alpha = 0.2$, $n = 0.03$, $k = 15$, $\pi = 0.8$, v_D range is 5 to 25, v_C range is 0.1 to 20.

Figure 4: The cost function is $k*(p^*-p_0)^2$. The non-illustrated parameter values are $p_0 = 0.001$, $p_1 = 0.9$, $c_D = 6$, $c_C = 1.5$, $N_C = 30$, $N_D = 40$, $\alpha = 0.2$, $k = 15$, $\pi = 0.8$, v_D range is 5 to 25, v_C range is 0.1 to 20.

Figure 5 (Remark 4): The cost function is $k * (p^* - p_0)/(p_1 - p_0)$. The non-illustrated parameter values are $c_D = 1.2$, $c_C = 1.5$, $N_C = 32$, $N_D = 10$, $\alpha = 0.12$, $k = 8$, $v_C = 40$, v_D range is 12 to 35.

Part III

Extension: Making n Endogenous

For this extension (and all others), I will add a simple assumption to rule out fairly uninteresting cases.

Extension Assumption *I will assume that* $p^D > p_0$.

It may be possible for D to manipulate both arming and the level of nuclear instability. This extension consider this possibility.

5 Game form

Two players, a challenger (C) and a defender (D), are in a deterrence game with complete information. The game order is as follows.

1. D selects a conventional force level that determines $p \in [p_0, p_1]$, which is D's likelihood of winning in a conventional conflict. I assume $0 < p_0 < p_1 < 1$. For ease, I will sometimes refer to this force level as D's "arming" level. D also selects $n_D \in \mathcal{N} \subset \mathbb{R}_+$, which denotes the nuclear instability parameter. $\mathcal{N}$ is assumed to be closed and compact.

2. C selects whether to challenge or not. If C does not challenge, the game ends with C receiving payoff 0 and D receiving payoff $v_D - K(p)$, where $K : \mathbb{R}_+ \to \mathbb{R}_+$ is D's costs from the conventional force level. I assume $K(p_0) = 0$, and K is continuous and increasing in p. If C does challenge, the game moves to the next stage.

3. D selects whether to acquiesce or escalate to conflict. If D acquiesces, C receives payoff v_C and D receives payoff $-K(p)$. If D escalates to conflict, then both states receive their conflict payoffs (described below).

The game here is nearly identical to the game form in the main text, only here D selects the nuclear instability parameter n_D, whereas previously this was given as fixed (as $n > 0$). For the sake of completeness, this means that $h(p) = n_D + \frac{\alpha}{p(1-p)}$, and C's expected utility from conflict is

$$\frac{n_D}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}((1-p)v_C) - \frac{c_C}{h(p)},$$

and D's expected utility from conflict is [11]

$$\frac{n_D}{h(p)} * (-N_D) + \frac{\alpha}{h(p)p(1-p)}(pv_D) - \frac{c_D}{h(p)} - K(p).$$

6 Equilibrium

Much of the intuition is the same as it was in the game in the main text. I highlight the differences here. When D does best deterring C, D can now (sometimes) use n_D to lower their arming costs. As it was earlier, to deter C, D must select an arming level where both D's war participation constraint and C's war cost constraint hold. However, now this is slightly different. I let $p^D(n_D)$ denote the following:

$$p^D(n_D) = 1 - \frac{\alpha v_D}{c_D + n_D N_D}$$

To make D most willing to go to war at the smallest possible arming level, D will select the lowest possible nuclear instability parameter. So long that the selected arming level is greater than or equal to this expression, D is willing to fight. For ease, I denote this arming level

$$p^D(\underline{n_D}) = 1 - \frac{\alpha v_D}{c_D + min\left\{\mathcal{N}\right\} N_D}$$

I also let $p^C(n_D)$ denote the following.

$$p^C(n_D) = \frac{\alpha v_C}{c_C + n_D N_C}.$$

To make C least willing to go to war at the smallest possible arming level, D will select the greatest possible nuclear instability parameter. For ease, I denote this arming level

[11]Or, for C and D (respectively), using $h(p) = \frac{\alpha + np(1-p)}{p(1-p)}$,

$$\frac{np(1-p)}{\alpha + np(1-p)} * (-N_C) + \frac{\alpha}{\alpha + np(1-p)}((1-p)v_C) - \frac{c_C p(1-p)}{\alpha + np(1-p)}$$

$$-\frac{np(1-p)}{\alpha + np(1-p)} N_D + \frac{\alpha}{\alpha + np(1-p)}(pv_D) - \frac{c_D p(1-p)}{\alpha + np(1-p)} - K(p)$$

$$p^C(\bar{n_D}) = \frac{\alpha v_C}{c_C + max\{\mathcal{N}\} N_C}.$$

Together, so long that D selects $p = max\left\{p^C(\underline{n_D}), p^D(\bar{n_D})\right\}$, deterrence will hold.

Sometimes D will prefer to fight a war. Now, when D does best going to war, D selects

$$\{\hat{p}, \hat{n_D}\} \in argmax_{p\in[max\{p^D(n_D),p_0\},min\{p^C(n_D),p_1\}]\times\hat{\mathcal{N}}} \left\{\frac{n_D}{h(p)} * (-N_D) + \frac{\alpha}{h(p)p(1-p)}(pv_D) - \frac{c}{h(p)} - K(p)\right\}.$$

For any given $n_D \in \mathcal{N}$ where $max\left\{p^D(n_D), p_0\right\} \leq min\left\{p^C(n_D), p_1\right\}$, a maximizer (or set of maximizers) exists. The function is continuous and optimized over a closed and compact set, meaning at least one maximizer exists. And, because $\mathcal{N}$ is closed and compact, at least one joint $\{\hat{p}, \hat{n_D}\}$ exists. Once again, if there are multiple possible maximizers, I assume D selects the smallest $\hat{p}$ from the set. I let $U_D(\hat{p}, \hat{n_D})$ denote D's utility from the above. Alternatively, it can also sometimes be the case that for some n_D $max\left\{p^D(n_D), p_0\right\} > min\left\{p^C(n_D), p_1\right\}$; when this is the case, the function is not defined. For that reason, I restrict the set of nuclear instability parameters $\mathcal{N}$ to $\hat{\mathcal{N}}$, which is the set of n_D values where $max\left\{p^D(n_D), p_0\right\} \leq min\left\{p^C(n_D), p_1\right\}$. If the set is empty, then D cannot ever select a $(\hat{p}, \hat{n_D})$ value where war occurs. Together, the equilibrium is as follows:

Proposition: *There exists an essentially unique*[12] *subgame perfect equilibrium taking the following form. Working backwards, if challenged, D will fight whenever $p \geq p^D(n_D)$ and will acquiesce otherwise. Before D fights or acquiesces, C will not challenge when $p \geq p^D(n_D)$ and $p \geq p^C(\bar{n_D})$ and will challenge otherwise. And, before C challenges or not, letting p^* denote equilibrium arming levels, D will select the following arming levels.*

- ***Case 1:*** *When $p^D(\bar{n_D}) < p^C(\underline{n_D}) \leq p_1$,*
  - *If $V_D - K(p^C(\underline{n_D})) \geq 0$ and $V_D - K(p^C(\underline{n_D})) \geq U_D(\hat{p}, \hat{n_D})$, then D selects $p^* = p^C(\bar{n_D})$ and C is deterred,*
  - *If $0 > V_D - K(p^C(\underline{n_D}))$ and $0 > U_D(\hat{p}, \hat{n_D})$, then D selects $p^* = p_0$, C challenges, and D acquiesces,*
  - *Otherwise, D selects $p^* = \hat{p}$ and $n_D^* = \hat{n_D}$, C challenges, and D fights.*
- ***Case 2 (Deterring C is Impossible):*** *When $p^D(\bar{n_D}) < p^C(\underline{n_D})$ and $p^C(\underline{n_D}) > p_1$,*

[12] *Two types of equilibria could also exist. First, mixed strategy equilibria can exist when players are indifferent over their actions. For example, if $U_D(\hat{p}) < 0 = V_D - k(p^C)$, D could mix over $p = p_0$ and $p = p^C$. Second, the optimal selected p conditional on D wanting to fight C could take on multiple values that D could mix over. Ultimately, we are assuming that both D and C are not mixing over actions that they are indifferent over when these cases arise.*

- *If* $U_D(\hat{p}, \hat{n_D}) \geq 0$, *then D selects* $p^* = \hat{p}$ *and* $n_D^* = \hat{n_D}$, *C challenges, and D fights,*
- *Otherwise, D selects* $p^* = p_0$, *C challenges, and D acquiesces.*

- ***Case 3:*** *When* $p^C(\underline{n_D}) \leq p^D(\bar{n_D})$ *and* $\hat{\mathcal{N}}$ *is non-empty*
  - *If* $V_D - K(p^D(\bar{n_D})) \geq 0$, *and* $V_D - K(p^D(\bar{n_D})) \geq U_D(\hat{p}, \hat{n_D})$, *then D selects* $p^* = p^D(\bar{n_D})$ *and C is deterred,*
  - *If* $V_D - K(p^D(\bar{n_D})) < 0$ *and* $0 \geq U_D(\hat{p}, \hat{n_D})$, *D selects* $p^* = p_0$, *C challenges, and D acquiesces.*
  - *Otherwise, D selects* $p^* = \hat{p}$ *and* $n_D^* = \hat{n_D}$, *C challenges, and D fights.*
- ***Case 4:*** *When* $p^C(\underline{n_D}) \leq p^D(\bar{n_D})$ *and* $\hat{\mathcal{N}}$ *is empty*
  - *If* $V_D - K(p^D(\bar{n_D})) \geq 0$, *then D selects* $p^* = p^D(\bar{n_D})$ *and C is deterred,*
  - *Otherwise, D selects* $p^* = p_0$, *C challenges, and D acquiesces.*

7 Cross-Equilibrium Analysis

When D can select the nuclear instability parameter, it can leads to greater or lower levels of nuclear instability and lower levels of arming. For example, consider $\mathcal{N} = \{\underline{n}, n, \bar{n}\}$, with $\underline{n} < n < \bar{n}$. Suppose it is optimal for D to deter C. In this new model (compared to the old model), here D either selects $p^* = p^D(\bar{n_D})$ or $p^* = p^C(\underline{n_D})$, where these values are lower than p^D and p^C (respectively).

Part IV

Extension: Endogenous Bargaining

The model in the main text was a deterrence model, much like Carter (2010), Baliga *et al.* (2020), and the paper this model is closest to, Powell (2015). However, some readers may have concerns about what the addition of bargaining would do to the paper's equilibrium actions and results. So long that the crisis bargaining setting has some kind of commitment problem, fighting is still possible and the crisis bargaining setting strongly resembles the deterrence setting. Here I modify the model to (a) allow for endogenous bargaining and (b) have a commitment problem stemming an exogenous power shift.

The model is as follows. Two players, a challenger (C) and a defender (D), are in a deterrence game with complete information over an infinite time horizon. The game order is as follows.

1. Period $t = 1$ begins.
2. D selects a conventional force level $p \in [p_0, p_1]$, which determines D's likelihood of winning in a conventional conflict in period $t = 1$. I assume $0 < p_0 < p_1 < 1$. Arming costs D one-time cost $K(p)$, with $K : \mathbb{R}_+ \to \mathbb{R}_+$ increasing and continuous, and $K(p_0) = 0$.
3. D selects $x_t \in [0, 1]$, which is some proposed split of the asset.
4. C selects whether to challenge or not. If C does not challenge, C receives share x_t of the asset, D receives share $1 - x_t$ of the asset, and the game moves to Step 6. If C challenges, the game moves to Step 5.
5. In response to C challenging, D selects whether to acquiesce or fight. If D acquiesces, the game ends and C receives the entirety of the good for all remaining periods. If D fights and escalates to conflict, then the game ends and both states receive their conflict payoffs (described below).
6. In response to C not challenging, period t ends, and actors receive a per-period payoff that is the split of the good. The period is updated to $t = t + 1$, and payoffs in period t are discounted by the common rate δ. The game re-starts at Step 3.

Suppose D makes some stream of offers x_t for all t and C does not challenge. Their payoffs are

$$U_C = \sum_{t=1}^{\infty} \delta^{t-1} x_t v_C$$

$$U_D = - K(p) + \sum_{t=1}^{\infty} \delta^{t-1} (1 - x_t) v_D$$

Next, suppose D makes some offer x_1, C challenges, and D acquiesces. The payoffs are

$$U_C = \frac{v_C}{1 - \delta}$$

$$U_D = - K(p)$$

Next, suppose D arms to level p, makes some offer x_1, C challenges (in the first round), and D fights. Letting $h(p) = n + \frac{\alpha}{p(1-p)}$, The payoffs here are as follows

$$U_C = \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}((1-p)v_C) - \frac{c_C}{h(p)}\right)$$

$$U_D = -K(p) + \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_D) + \frac{\alpha}{h(p)p(1-p)}(pv_D) - \frac{c_D}{h(p)}\right)$$

Finally, after the initial period, assume that D and C have engaged in a series of offers where C does not challenge. Then, in period $q > 1$, C challenges in response to x_q and D fights. I abuse notation and define the new distribution of power as exogenous parameter $\tilde{p} \in [p_0, 1)$.

$$U_C = \sum_{t=1}^{q} \delta^{t-1} x_t v_C + \frac{\delta^q}{1-\delta}\left(\frac{n}{h(\tilde{p})} * (-N_C) + \frac{\alpha}{h(\tilde{p})\tilde{p}(1-\tilde{p})}((1-\tilde{p})v_C) - \frac{c_C}{h(\tilde{p})}\right)$$

$$U_D = -K(\tilde{p}) + \sum_{t=1}^{q} \delta^{t-1}(1-x_t)v_D + \frac{1}{1-\delta}\left(\frac{n}{h(\tilde{p})} * (-N_D) + \frac{\alpha}{h(\tilde{p})\tilde{p}(1-\tilde{p})}(\tilde{p}v_D) - \frac{c_D}{h(\tilde{p})}\right)$$

Essentially, in this model, if D makes C an offer x_1 and C does not challenge, the game moves on with bargaining. A key feature here is that between periods 1 and 2, there is an exogenous power shift that might create a commitment problem. While bargaining could play out, C (or D) could decide that fighting in the first period would be optimal. Admittedly, this is a stark way of modeling a future shift in power: ultimately, this model is about showing how C and D being in a crisis in an environment with a power shift makes a crisis bargaining version of the model produce virtually identical results as the deterrence model (in the paper). While this could be done in many different, the modeling assumptions made here are designed to make this more complex model as tractable as possible and to offer the simplest intuition as to why the introduction of bargaining does not radically change the Remarks.

8 Equilibrium Assumptions & Equilibria

I only consider subgame perfect equilibria. I also make a series of assumptions about parameter features to simplify the analysis. I make two assumptions, one of which that is equivalent the assumptions in the main text. I assume that without any arming, D is unwilling to fight in the first period after C challenges D. Letting $p^D = 1 - \frac{\alpha v_D}{c_D + nN_D}$, this is $p^D > p_0$.[13] Additionally, I assume that in the first period, D will be willing to fight at some arming level, or that $p^D \leq p_1$.

The next assumption is new and relates to the power shift. I assume that if the power shift is allowed to happen, then this power shift favors D to the point where C is no longer willing to challenge in the future. This is

$$0 \geq \frac{1}{1-\delta}\left(\frac{n}{h(\tilde{p})} * (-N_C) + \frac{\alpha}{h(\tilde{p})\tilde{p}(1-\tilde{p})}\left((1-\tilde{p})v_C\right) - \frac{c_C}{h(\tilde{p})}\right).$$

Additionally, this power shift is such that D is always willing to fight when challenged, or

$$\tilde{p} \leq p^D$$

All these assumptions reduce the numbers of cases and make the ensuing analysis simpler.

The equilibria are as follows:

8.1 Periods $t \geq 2$

If the game enters into the second period, the game settles into a fixed equilibrium path. Here, D has experienced the power shift, meaning here D can extract the asset from C via bargaining (i.e. set $x_t = 0$ for all $t \geq 2$). When D makes this offer, C is at least indifferent between accepting and challenging, and will accept. From here, the game repeats such that D can keep offering C $x_t = 0$, and C will continue accepting.

8.2 Period $t = 1$

The bulk of strategic play happens in the first round. **Essentially, three things can happen. First, D could acquiesces**. This is D not arming ($p = p_0$) and setting any $x_1 \in [0, 1]$, then C challenging, and then D acquiescing. This gives C the asset in the first round and all future rounds. This will give D payoff

$$U_D(acquiesce) = 0.$$

[13]Note in this model the condition is $0 > \frac{1}{1-\delta}\left(\frac{n}{h(p_0)} * (-N_D) + \frac{\alpha}{h(p_0)p_0(1-p_0)}(p_0 v_D) - \frac{c_D}{h(p_0)}\right)$, which is simplified to the same p^D.

Second, D could deter C. If D is optimally deterring C, D will select a p and x_1 that optimizes the following (recall $h(p) = n + \frac{\alpha}{p(1-p)}$):

$$max_{p \in [p_0, p_1], x_1 \in [0,1]} \left\{ -K(p) + (1 - x_1)v_D + \frac{\delta}{1-\delta} v_D \right\}$$

conditional on the following holding

$$p \geq p^D$$

$$x_1 v_C \geq \frac{1}{1-\delta} \left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)} ((1-p)v_C) - \frac{c_C}{h(p)} \right)$$

I let $x_1 = \check{x}$ and $p = \check{p}$ denote the solution to the above optimization. There are several things to note here. Note that the first constraint (that $p \geq p^D$) is the same as it was in the main model. Essentially, when D is choosing between fighting and acquiescing, D is choosing between getting zero forever or getting D's wartime payoff forever, making the constraint equivalent to the one-period model.[14] Also note that the second constraint is different—because D can make some first-round offer to C to prevent war, C is deciding between fighting in the first round and accepting a first round offer (x_1) then getting nothing for all future periods given period 2 behavior. This dynamic is similar, but still somewhat different from the main model. Also note, as it was in the main model, there may not exist a feasible p and x_1 satisfying these these two inequalities. The second inequality may not be satisfied, which is akin to when $p^C > p_1$ in the main model. Lastly, note that this is still set up as two constraints on p, one where D must be willing to fight (the top constraint) and the second where C must do sufficiently bad from challenging conditional on D fighting.

Overall, when D deters C, I assume that D selects some $x_1 = \check{x}$ and $p = \check{p}$ that are the solution to the maximization problem above. This will give D utility

$$U_D(\check{x}, \check{p}) = -K(\check{p}) + (1 - \check{x})v_D + \frac{\delta}{1-\delta} v_D$$

Third, D could go to war. When this is the case, D will select a p and x_1 such that the following holds.

$$max_{p \in [p_0, p_1], x_1 \in [0,1]} \left\{ -K(p) + \frac{1}{1-\delta} \left(\frac{n}{h(p)} * (-N_D) + \frac{\alpha}{h(p)p(1-p)} (pv_D) - \frac{c_D}{h(p)} \right) \right\}$$

[14]D must select some p such that $0 * \frac{1}{1-\delta} \leq \frac{1}{1-\delta} \left(\frac{n}{h(p)} * (-N_D) + \frac{\alpha}{h(p)p(1-p)} (pv_D) - \frac{c_D}{h(p)} \right)$.

such that D would be willing to fight if challenged and C does not challenge, or (respectively)

$$p \geq 1 - \frac{\alpha v_D}{c_D + nN_D} = p^D$$

$$x_1 v_C < \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}((1-p)v_C) - \frac{c_C}{h(p)}\right)$$

I denote $x_1 = \tilde{x}$ and $p = \tilde{p}$ as the values that optimize the above. Note that I set this up as an optimization over all feasible p where the second constraint could be satisfied with equality. But, if D ever selects a x_1 and p that has the second constraint hold with equality, then this will deter C. Ultimately, this does not matter much; if D selects such a set of values, D would instead prefer deterring C from challenging, so this will never be optimal for D. For shorthand, I will denote D's utility from selecting this optimal $\tilde{x}$ and $\tilde{p}$ as $U_D(\tilde{x}, \tilde{p})$.

Together, I can describe D's equilibrium play.

Proposition 1B: *There exists an essentially unique* [15] *subgame perfect equilibrium taking the following form. Working backwards, in any round where $t \geq 2$, D will set $x_t = 0$ and C will accept. In the first round, if challenged, D will fight whenever $p \geq p^D$ and will acquiesce otherwise. Before D fights or acquiesces, if both $p \geq p^D$ and $x_1 v_C \geq \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}((1-p)v_C) - \frac{c_C}{h(p)}\right)$, then C will not challenge; otherwise C will challenge. And before C challenges or not, D will select the following arming levels (letting $x_1 = x^*$ and $p = p^*$ denote equilibrium arming levels):*

- ***Case 1:*** *When there exists some feasible x_1, p satisfying $x_1 v_C \geq \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}((1-p)v_C) - \frac{c_C}{h(p)}\right)$, and there exists some (different) feasible x_1, p satisfying both $x_1 v_C < \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}((1-p)v_C) - \frac{c_C}{h(p)}\right)$ and $p \geq p^D$*
  - *If $U_D(\check{x}, \check{p}) \geq U_D(\tilde{x}, \tilde{p})$ and $U_D(\check{x}, \check{p}) \geq 0$, then D selects $p^* = \check{p}$ and $x^* = \check{x}$ and C will not challenge (C is deterred).*
  - *If $U_D(\check{x}, \check{p}) < U_D(\tilde{x}, \tilde{p})$ and $U_D(\tilde{x}, \tilde{p}) \geq 0$, then D selects $p^* = \tilde{p}$ and $x^* = \tilde{x}$, C will challenge, and D will fight.*
  - *Otherwise, D will select $p^* = p_0$, C will challenge, and D will acquiesce.*

[15] *Two other types of equilibria could also exist. First, mixed-strategy equilibria can exist when players are indifferent over their actions. Second, the optimal selected p conditional on D wanting to fight C could take on multiple values that D could mix over. Ultimately, I assume assuming that neither D nor C is mixing over actions that they are indifferent over when these cases arise.*

- ***Case 2 (deterrence is impossible):*** *When there* <u>*does not*</u> *exist any feasible* x_1, p *satisfying* $x_1 v_C \geq \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}\left((1-p)v_C\right) - \frac{c_C}{h(p)}\right)$*, and there exists some feasible* x_1, p *satisfying both* $x_1 v_C < \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}\left((1-p)v_C\right) - \frac{c_C}{h(p)}\right)$ *and* $p \geq p^D$*.*
    - *If* $U_D(\tilde{x}, \tilde{p}) \geq 0$*, then D selects* $p^* = \tilde{p}$ *and* $x^* = \tilde{x}$*, C will challenge, and D will fight.*
    - *Otherwise, D selects* $p^* = p_0$*, C challenges, and D acquiesces.*
- ***Case 3 (fighting is impossible):*** *When there exists some feasible* x_1, p *satisfying* $x_1 v_C \geq \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}\left((1-p)v_C\right) - \frac{c_C}{h(p)}\right)$*, and there* <u>*does not*</u> *exist any feasible* x_1, p *satisfying both* $x_1 v_C < \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}\left((1-p)v_C\right) - \frac{c_C}{h(p)}\right)$ *and* $p \geq p^D$*.*
    - *If* $U_D(\check{x}, \check{p}) \geq 0$*, then D selects* $p^* = \check{p}$ *and* $x^* = \check{x}$ *and C will not challenge (C is deterred).*
    - *Otherwise, D selects* $p^* = p_0$*, C challenges, and D acquiesces.*

Proof: Follows by construction.

9 Results

9.1 Feasibility of Fighting/Game Outcomes

First, it is worthwhile pointing out that, in this model, fighting is entirely feasible. Consider the parameters $p_0 = 0.001$, $p_1 = 0.8$, $c_D = 10$, $c_C = 1$, $N_C = 10$, $N_D = 40$, $\alpha = 0.2$, $n = 0.02$, $K(p) = 30 * (p - p_0)^2$, $v_D = 20$, $v_C = 10$, and $\delta = 0.95$. Under these parameters, D cannot deter C: when D sets $p = 0.8$ and $x = 1$, C is still willing to fight; however, D does better here fighting rather than acquiescing, so in equilibrium, D will set $p = 0.8$, $x = 0$, C will challenge, and D will fight. Naturally, fighting is not the only possible outcome: if all parameters were the same by v_D was lowered to $v_D = 12$, then D would acquiesce. Below, I will discuss cases where D deters C.

9.2 Do the Remarks Still Hold?

Regarding Remarks 1, 2, and 5, the intuition described in the main text still holds. For deterrence, D must select a p such that (a) D is willing to fight, and (b) C must suffer enough from fighting. These two constraints are still influenced by n in opposite ways. For example, under the parameters $p_0 = 0.001$, $p_1 = 0.9$, $c_D = 8$, $c_C = 1.5$, $N_C = 50$, $N_D = 40$, $\alpha = 0.2$, $n = 0.02$, $K(p) = 15 * (p - p_0)^2$, $v_D = 15$, $v_C = 8$, and $\delta = 0.9$, D wants to deter C, and can do so by setting $p = 0.6591$ and $x_1 = 0$. However, if n increases to 0.035, then D switches to setting $p = 0.6809$ and $x_1 = 0$; in short, p increases, thus leading to greater arming costs. As a second example, under the parameters, $p_0 = 0.001$, $p_1 = 0.9$, $c_D = 10$, $c_C = 1.5$, $N_C = 30$, $N_D = 40$, $\alpha = 0.2$, $n = 0.02$, $K(p) = 35 * (p - p_0)^2$, $v_D = 15$, $v_C = 10$, and $\delta = 0.9$, D will deter C by setting $p = 0.7958$ and $x_1 = 0.3305$. However, if n increases to 0.035, then D switches to setting $p = 0.7502$ and $x = 0.1052$; in short, p decreases, thus leading to lower arming costs.

A partial form of Remark 3 can be replicated. The difficulty here lies in analyzing the constraint on C's arming level; whereas before the constraint could be written in terms of p ($p < p^C$ for C to be willing to fight), now this constraint is more complex.[16] There is not an easy resolution for this. Instead, I offer a different version of Remark 3 for this new model.

Remark 3B. *Consider nuclear instability parameters $n', n'' \in \mathbb{R}_+$ with $n' < n''$. If n' shifts to n'' and $x_1^*(n') = x_1^*(n'') = 0$, then the likelihood of war weakly decreases.*

Essentially, by assuming that the equilibrium the optimal offer is zero across both levels of nuclear instability, I am able to attain the nuclear peace result in this new model. Additionally, this result is easy to achieve in simulations; if I used the parameters from the "Feasibility of Fighting/Game Outcomes" subsection above but raised n to 0.2, then D would optimally deter C by setting $p = 0.7778$ and $x_1 = 0$, which would keep C from challenging (i.e. reduce the likelihood of war from certainty to zero).

Regarding Remark 4, the proof is identical to as it was in the main text (K and the $1/(1-\delta)$ terms drop out when signing the cross-partial derivative).

Part V

Extension: Incomplete Information Model and Discussion

10 Equilibrium Overview

The discussion of equilibrium behavior in the main paper was written in terms of strategic behavior. This was done in an attempt to make the presentation of strategic behavior as-clear-as-possible. Here, I discuss the equilibrium characterization in terms of specific arming levels, namely, considering every unique arming pair from both types of D. For the purpose of proving the various characterizations are equilibrium behavior within a set of parameters, this is better.

[16] This is now $x_1 v_C \leq \frac{1}{1-\delta}\left(\frac{n}{h(p)} * (-N_C) + \frac{\alpha}{h(p)p(1-p)}\left((1-p)v_C\right) - \frac{c_C}{h(p)}\right)$.

	Type $\underline{v}_D$ arming	Type $\bar{v}_D$ arming	**How is arming used? (Low-type first)**	**War with D? (Low-type first)**
Separating 1	p_0	$p^D(\bar{v}_D)$	Acquiesce, Deter	No, No
Separating 2	p_0	p^C	Acquiesce, Deter	No, No
Separating 3	p_0	$\hat{p}(\bar{v}_D)$	Acquiesce, Fight	No, **Yes**
Separating 4	p_0	$\bar{p}$	Acquiesce, Signal	No, No
Separating 5	$\hat{p}(\underline{v}_D)$	$\hat{p}(\bar{v}_D)$	Fight, Fight	**Yes, Yes**
Separating 6	$\hat{p}(\underline{v}_D)$	p^C	Fight, Deter	**Yes**, No
Pooling 1	p_0	p_0	Acquiesce, Acquiesce	No, No
Pooling 2	$p^D(\bar{v}_D)$	$p^D(\bar{v}_D)$	Bluff, Deter	No, No
Pooling 3	$\tilde{p}$	$\tilde{p}$	Bluff, Deter	No, No
Pooling 4	$p^D(\underline{v}_D)$	$p^D(\underline{v}_D)$	Deter, Deter	No, No
Pooling 5	p^C	p^C	Deter, Deter	No, No

Table 1: Equilibrium Summary. Note that implicit here is that $p_0 < p^D(\bar{v}_D)$.

What does equilibrium arming behavior look like? I summarize these various arming levels in the Table 1, which assumes $p_0 < p^D(\bar{v}_D)$.[17]

The way to read the table is as follows. The first column names the equilibrium, indicating whether it is pooling or separating. The second and third column specify the low-type's and high-type's arming levels. The forth column describes what the arming level accomplishes, using the terminology in the text. And the fifth column flags if war occurs or not.

To give a sense of what the game looks like, see Figure 2 displays the equilibrium for various parameters while allowing $\underline{v}_D$ and v_C to vary. $\underline{v}_D$ increases along the x-axis, and v_C increases along the y-axis. Note that these are the same parameters as Figure 1 in the main text, which allows for easy comparison between the labeling here and the labeling in the text. For example, the Separating 1 and Separating 2 equilibria form the "Deter-Acquiesce" equilibrium space, Pooling 2 and Pooling 3 form the "Deter-Bluff" equilibrium space, etc.

To give a sense of how the game plays out, recall that high-type D's are always willing to arm to level $p^D(\bar{v}_D)$ if this results in C not challenging by the parameter assumptions. In the bottom-left corner of Figure 1 (Separating 1), low-type D's care very little about the asset (low $\underline{v}_D$). In the Separating 1 parameter space, low-type D's are unwilling to arm to the level where they would imitate high-type D's, even if it led to them attaining the asset. In this parameter space, C will never challenge upon observing $p = p^D(\bar{v}_D)$ because C knows only high-types of D would be wiling to make this investment, and high-types would always fight after selecting

[17] I express the full equilibria without this assumption in the "Characterizing and Proving the Equilibria" section. To offer one example, the arming levels in Pooling 2 without this assumption would be $p = max\{p_0, p^D(\bar{v}_D)\}$.

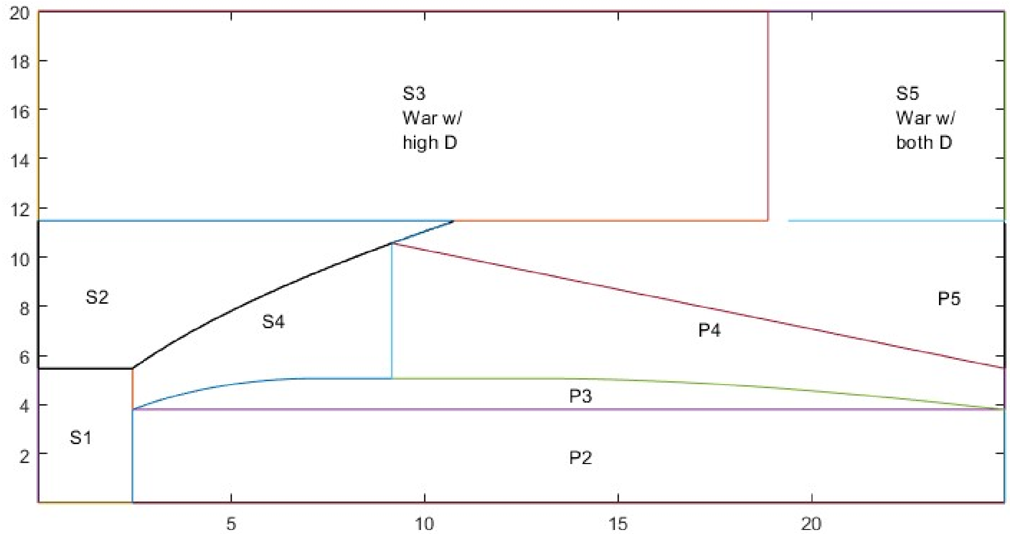

Figure 2: On the x-axis, I vary the values of $\underline{v}_D$, and on the y-axis I vary the values of v_C. "S1" is in reference to "Separating 1," and "P2" is in reference to "Pooling 2," etc. I add extra text to describe where war happens. The cost function is $k * (p^* - p_0)^2$. Note that Pooling 1 isn't visualized; because high-type D's going to war gives these D's a better payoff than selecting p_0 and acquiescing, this equilibrium space is ruled out (if war were more expensive, P1 would exist roughly where S3 and S5 is).

this investment level. Also here, C will always challenge upon observing $p = p_0$, because only low-types make the low investment and C knows that if they challenge, then they will attain the asset.

Moving to the right, low-type D's care more about the asset. Within Pooling 2, low-type D's are willing to select arming level $p = p^D(\bar{v}_D)$ if it results in them attaining the asset (i.e. C not challenging).[18] That being said, if low-type D's select arming level $p = p^D(\bar{v}_D)$ and were challenged, they would not fight because $p^D(\bar{v}_D) < p^D(\underline{v}_D)$; however, in this range, because C cares so little about the asset, C is unwilling to challenge at arming level $p^D(\bar{v}_D)$ even though C knows that by challenging all low-type D's would drop out. Moving up to Pooling 3, the logic is the same, only D's must pool on a slightly higher level of arming $p = \tilde{p}$ to deter C from challenging even though C knows low-type D's would drop out if challenged.

Moving up from Pooling 3 into the Separating 4 and Pooling 4 regions, C cares more about the asset, but is still unwilling to challenge if C knew that D would fight in response. Within this range of parameters, no arming level exists where (a) high-type D's would fight when challenged, (b) low-type D's would acquiesce when challenged, and (c) C would be deterred from challenging conditional on D's behavior as characterized in Pooling 2 and 3. When low-

[18] This holds because $\underline{v}_D - k(\bar{p}_D) \geq 0$ in this range.

type D's do not value the asset enough—as characterized by $v_D - k(p^D(\underline{v}_D)) < 0$)—a separating equilibrium (Separating 4) exists where high-type D's select an arming level that will insure low-type D's will not mimic them ($p = \bar{p}$), and low-type D's will select the lowest arming level $p = p_0$. In response, C would never challenge when observing $p = \bar{p}$, and would always challenge when observing $p = p_0$. When low-type D's value the asset more (Pooling 4)—as characterized by $v_D - k(p^D(\underline{v}_D)) \geq 0$—low-type D's attain a positive utility from arming to level $p = p^D(\underline{v}_D)$ and deterring C from challenging.

Moving up again, when $p^D(\bar{v}_D) < p^C$ (to Separating 2), then the level of arming that would convince a high-type D to fight after being challenged is less than the level of arming that would deter C from challenging conditional on D fighting with certainty. Thus, within this range, D must select a level of arming that exceeds $p^D(\bar{v}_D)$ to deter C. This level of arming will be p^C, the level that would make C not challenge (Separating 2). In the range of $\underline{v}_D$ values where $\bar{v}_D - k(p^D(\underline{v}_D)) \geq 0$, low-type D's begin caring enough and could select Pooling 5, where they choose p^C in order to deter C.

Finally, in the top region of the graph, high-type D's are no longer willing to arm to level p^C to deter C, or it becomes impossible with $p^C > p_1$. Under these parameters, then high-type D's will either select $p = p_0$ and acquiesce when challenged (Pooling 1), or will select $p = \hat{p}(\bar{v}_D)$ and fight when challenged (Separating 3), depending on which offers D a greater utility. When high-type D's optimally select $p = p_0$, low-type D's will always match high-type D's play and select $p = p_0$. When high-type D's optimally select $p = \hat{p}(\bar{v}_D)$, low-type D's will either select $p = p_0$ (when $\underline{v}_D$ is low, Separating 3), or will select $p = \hat{p}(\underline{v}_D)$ and will fight when challenged (Separating 5).[19]

11 Proving Lemma 1

Before I characterize and prove the equilibrium, I must prove Lemma 1. Lemma 1 establishes, for a set of parameters, that the selected level of arming p is increasing in D's type ($v_D \in \{\underline{v}_D, \bar{v}_D\}$). This Lemma is useful on two accounts. First, it is critical to Remark 1, which establishes the positive relationship between type and arming across the entire set of possible parameters. Second, the set of parameters included within Lemma 1 span several equilibria spaces. This allows me to refer to the monotonicity result within Lemma 1 at several points to make the equilibrium proofs more abbreviated.

Below I will refer to p^C and $p^D(v_D)$, where $v_D \in \{\underline{v}_D, \bar{v}_D\}$. I let $p = \frac{\alpha v_C}{c_C + nN_C}$ and $p^D(v_D) = 1 - \frac{\alpha v_D}{c_D + nN_D}$. These are the arming values where C would be deterred from challenging if D

[19]Note: Separating 5 may not actually be separating when low and high type D's optimally select the same arming level when fighting. For example, this can occur when $p^C > p_1$ and both low and high types arm to level $p = p_1$.

fought, and the arming value where type v_D D would be willing to fight.

***Lemma 1**: Suppose $\underline{v}_D - K(p^D(\underline{v}_D)) \geq 0$ and $p^D(\underline{v}_D) < p^C$. Given C's equilibrium behavior, within this region, high types select greater arming levels (i.e. $p^*(\bar{v}_D) \geq p^*(\underline{v}_D)$.*[20]

Proof: Consider what this range means for D. Due to $\underline{v}_D - K(p^D(\underline{v}_D)) \geq 0$, low types may become willing to deter or fight rather than just arm to level p_0 and let C take the asset.[21] And, due to $p^D(\underline{v}_D) < p^C$, each type D faces their own optimization problem with their arming decision that fully determines equilibrium play. To summarize equilibrium play in this region, if a type v_D D selects some $p \in [0, p^D(v_D))$, C will challenge and D will acquiesce. If D selects some $p \in [p^D(v_D), p^C)$, then C will challenge and D will fight. And, if D selects some $p \geq p^C$, then C will not challenge. Formally, under these parameters, both types of D face a non-continuous, non-concave optimization problem with respect to arming, where their utility function for all $p \in [p_0, p_1]$ is

$$U_D(p; v_D) = \begin{cases} 0 - K(p) & if\, p < p^D(v_D) \\ -\frac{np(1-p)}{\alpha+np(1-p)} N_D + \frac{\alpha}{\alpha+np(1-p)}(pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)} - K(p) & if\, p^D(v_D) \leq p < p^C \\ v_D - K(p) & if\, p^C \leq p \leq p_1 \end{cases}$$

Because I have kept things general and cannot identify an explicit solution, for Lemma 1 to hold, I must show that in this region enough structure exists where high type D's will always select weakly lower levels of arming than low-type D's. This part of the proof will utilize the Topkis Monotonicity Theorem (Topkis 1978; Milgrom and Shannon, Econometrica 1994). For ease, I define the relevant increasing differences condition:

*Definition: Function $U_D : [p_0, p_1] \times \{\underline{v}_D, \bar{v}_D\} \to \mathbb{R}$ has **increasing differences (ID)** in (p, v_D) if, for all $p' > p$ and $v'_D > v_D$, $U_D(p', v'_D) - U_D(p, v'_D) \geq U_D(p', v_D) - U_D(p, v_D)$.*

The Topkis Monotonicity Theorem can then clarify the relationship between the set of selected arming levels $p^*(v_D) = argmax_{p \in [p_0, p_1)} U_D(p; v_D)$ and D's private value v_D. This is defined as the following:

Topkis Monotonicity Theorem: If $U_D(p; v_D)$ has increasing differences (ID) in p and v_D, then $p^*(v_D)$ is non-decreasing.

To use the Topkis Theorem, I first show that an optimal p (or set of p's) exist by demon-

[20] Technically the set p^* is non-decreasing in private type.

[21] When $\underline{v}_D - K(p^D(\underline{v}_D)) < 0$, low types prefer setting $p = p_0$ and receiving payoff 0 to arming to level $\underline{v}_D$ and getting the good with certainty.

strating that there are no "open set" issues. In the first region of the utility function, or **Region 1** (the region where the selected $p < p^D(v_D)$), D's utility is strictly decreasing in p, meaning the optimal p for this region is p_0 (so long that $p_0 < p^D(v_D)$).[22] Next, **Region 3** (where $p^C \leq p$), D's utility is strictly decreasing, making p^C the optimal arming level. Finally, consider an analysis of a modified[23] **Region 2**. Consider the function $V(p) = -\frac{np(1-p)}{\alpha+np(1-p)}N_D + \frac{\alpha}{\alpha+np(1-p)}(pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)} - K(p)$ that is optimized over the closed set $p^D(v_D) \leq p \leq p^C$. So long that $V$is not maximized at p^C, then there is a clearly defined optimum to $U_D(p)$ over the span of Region 2 ($p^D(v_D) \leq p < p^C$). If V is maximized at p^C, then based on the parameter assumptions, D would do strictly better setting $p = p^C$ and attaining utility $v_D - K(p^C)$.[24] Together, this means that the discontinuities between Regions will never create open set issues, making this a well-define optimization problem with at least one solution.

Having established a non-empty set of optima exist for $U_D(p; v_D)$ as defined above, I must show that the above utility function exhibits increasing differences (ID) in v_D and p. It is straightforward to see that within Regions 1 and 3—in other words, for a p, p' pair such that both p fall within Region 1 (or Region 3)—(ID) holds with equality. Within Region 2, (ID) is equivalent to showing

$$\frac{\alpha}{\alpha + np'(1-p')}(p'v'_D) - \frac{\alpha}{\alpha + np(1-p)}(pv'_D) - \left(\frac{\alpha}{\alpha + np'(1-p')}(p'v_D) - \frac{\alpha}{\alpha + np(1-p)}(pv_D)\right) \geq 0$$

or

$$\left(\frac{\alpha p'}{\alpha + np'(1-p')} - \frac{\alpha p}{\alpha + np(1-p)}\right)(v'_D - v_D) \geq 0.$$

This will hold so long that

$$\frac{\alpha p'}{\alpha + np'(1-p')} - \frac{\alpha p}{\alpha + np(1-p)} \geq 0$$

or

$$\alpha^2 p' - \alpha^2 p + \alpha npp'(1-p) - \alpha npp'(1-p') \geq 0,$$

which will hold because $p' > p$.

Across regions is slightly more complicated. To show U_D has increasing differences, I write

[22] Consider the edge case where $p_0 = p^D(v_D)$. In this case, D's utility from selecting p_0 then acquiescing when challenged is the same as their utility from selecting p_0 then fighting when challenged. So, for the equilibrium that I consider, D will always fight.

[23] The modification here is that I optimize over the set $p^D(v_D) \leq p \leq p^C$ here rather than over the set $p^D(v_D) \leq p < p^C$ as it was above.

[24] This will be the case because $\alpha p/(\alpha + np(1-p))$ is always less than 1.

out every case I must consider, as characterized by what Region of the utility function that the considered p or p' and v_D or v'_D put the function into. Note that there is some structure to the cases that I consider; for example, if (p', v'_D) puts the utility function into Region 2, then (p', v_D) , (p, v'_D), and (p, v_D) must fall within Region 2 or Region 1, but not Region 3. As intuition, lowering p and v_D can never shift p^C downward or result in $p \geq p^C$ when I started with $p' < p^C$. And, if (p', v'_D) (or (p, v'_D)) puts the utility function into Region 3, then (p', v_D) (or (p, v_D)) must also fall within Region 3 because p^C is unchanging in v_D.[25] The set of cases that I must consider are shown in the Table below. To interpret what this means, Case A (below) implies that for a given p' and v'_D, the utility function is in the second region (where D is going to war); and, for (p, v'_D), (p', v_D), and (p, v_D), the utility function is in the second region.

Cases	$U_D(p'; v'_D)$	$U_D(p; v'_D)$	$U_D(p'; v_D)$	$U_D(p; v_D)$
A	2	1	1	1
B	2	2	1	1
C	2	1	2	1
D	2	2	2	1
E	3	2	3	2
F	3	2	3	1
G	3	1	3	1

Before I proceed showing U_D has increasing differences ($U_D(p', v'_D) - U_D(p, v'_D) \geq U_D(p', v_D) - U_D(p, v_D)$), note the following properties hold:

Property (a): if $p \geq p^D(v_D)$, then $-\frac{np(1-p)}{\alpha+np(1-p)} N_D + \frac{\alpha}{\alpha+np(1-p)} (pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)} \geq 0$.[26]

Property (b): $\frac{\alpha p}{\alpha+np(1-p)} < 1$ and $\frac{\alpha p'}{\alpha+np'(1-p')} < 1$.[27]

Property (c): if $p \geq p^D(v_D)$, then $-\frac{np(1-p)}{\alpha+np(1-p)} N_D + \frac{\alpha}{\alpha+np(1-p)} (pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)}$ is increasing in p.[28]

Property (d): I abuse notation and (sometimes below will) bring in the region numbers to the utility function, letting $U_D(p; v_D, 1) = -K(p)$, $U_D(p; v_D, 2) = -\frac{p(1-p)}{\alpha+np(1-p)} (nN_D + c_D) + \frac{\alpha}{\alpha+np(1-p)} (pv_D) - K(p)$, and $U_D(p; v_D, 3) = v_D - K(p)$, <u>regardless of p's relationship to $p^D(v_D)$</u>

[25] This latter point rules out, for example, a "Case H" where, in order, the regions are 3, 3, 3, and 1. This is ruled out because if $U_D(p; v'_D)$ falls in region 3, then it must also be the case that $U_D(p; v_D)$ falls in region 3.

[26] This holds based on how $p^D(v_D)$ is defined: when $p \geq p^D(v_D)$, then D is willing to fight and attain utility $-\frac{np(1-p)}{\alpha+np(1-p)} N_D + \frac{\alpha}{\alpha+np(1-p)} (pv_D) - \frac{c_D p(1-p)}{\alpha+np(1-p)}$ over acquiesce and attain utility 0.

[27] This holds by virtue of $p \in [0, 1]$.

[28] Taking first order conditions gives $\frac{d}{dp}\left(-\frac{p(1-p)}{\alpha+np(1-p)} (nN_D + c_D) + \frac{\alpha}{\alpha+np(1-p)} (pv_D)\right) = \frac{\alpha(2p-1)(c_D+nN_D)+\alpha v_D(\alpha+np^2)}{(\alpha-n(p-1)p)^2}$, or equal to $\frac{\alpha p(c_D+nN_D)+\alpha(1-p)(-c_D-nN_D)+\alpha v_D(\alpha+np^2)}{(\alpha-n(p-1)p)^2}$. The right-hand side will be positive whenever $-(1-p)(c_D + nN_D) + v_D(\alpha + np^2) \geq 0$, which will hold by Property (a).

or p^C; for example, I will let $U_D(p^C; v_D, 1) = -K(p^C)$. If $p < p^D(v_D)$, then $U_D(p; v_D, 2) < U_D(p; v_D, 1)$ (because p is fixed).

I now describe how U_D exhibits increasing differences across all cases listed above.

Case A: $U_D(p'; v'_D) > U_D(p'; v_D)$ by property (a), and $U_D(p; v'_D) = U_D(p; v_D)$ because they are in Region 1; therefore, (ID) holds.
Case B: by property (c) $U_D(p'; v'_D) - K(p') - (U_D(p; v'_D) - K(p)) > 0$; therefore (ID) holds.

Case C: $U_D(p'; v'_D) > U_D(p'; v_D)$ because, in Region 2, U_D is increasing in v_D. Also, $U_D(p; v'_D) = U_D(p; v_D)$;therefore, (ID) holds.

Case D: because Region 2 exhibits (ID), I can say $U_D(p'; v'_D) - U_D(p; v'_D) - (U_D(p'; v_D) - U_D(p; v_D, 2)) \geq 0$. By Property (d) $U_D(p; v_D, 2) < U_D(p; v_D, 1) = U_D(p; v_D)$; therefore (ID) holds.

Case E: ID in region 2 implies $U_D(p'; v'_D, 2) - U_D(p; v'_D) - (U_D(p'; v_D, 2) - U_D(p; v_D)) \geq 0$. By property (b) $(v'_D - v_D)\left(1 - \frac{\alpha p'}{\alpha + np'(1-p')}\right) > 0$; I can add this to the left hand side and (ID) will then hold.

Case F: (ID) in region 2 implies $U_D(p'; v'_D, 2) - U_D(p; v'_D) - (U_D(p'; v_D, 2) - U_D(p; v_D, 2)) \geq 0$. Because $(v'_D - v_D)\left(1 - \frac{\alpha p'}{\alpha + np'(1-p')}\right) > 0$, I can add this to the left hand side and get $U_D(p'; v'_D) - U_D(p; v'_D) - (U_D(p'; v_D) - U_D(p; v_D, 2)) \geq 0$. I use property (d) to say that $U_D(p; v_D, 2) < U_D(p; v_D, 1) = U_D(p; v_D)$, which will imply (ID) holds.

Case G: $U_D(p'; v'_D) > U_D(p'; v_D)$ trivially and $U_D(p; v'_D) = U_D(p; v_D)$, meaning (ID) holds.

Thus, increasing differences holds, and $p^*(v_D)$ is non-decreasing. Note that in this region it is not only that $p^*(\bar{v}_D) \geq p^*(\underline{v}_D)$, but also, for example, if $\bar{v}'_D > \bar{v}_D$, $p^*(\bar{v}'_D) \geq p^*(\bar{v}_D)$. As discussed in the text, this condition does not hold throughout.

12 Characterizing and Proving the Equilibrium

Here I fully characterize every equilibrium, and prove it's existence within the given parameter set.

12.1 Separating Equilibrium 1:

- This equilibrium occurs when $\underline{v}_D - K(p^D(\bar{v}_D)) \leq 0$, $p^D(\bar{v}_D) \geq p^C$,[29] $p^D(\bar{v}_D) > p_0$.
- Type $\bar{v}_D$ selects $p^* = p^D(\bar{v}_D)$, and type $\underline{v}_D$ selects $p^* = p_0$.
- C will challenge for all $p < p^D(\bar{v}_D)$ and will not challenge for all $p \geq p^D(\bar{v}_D)$.

[29]Note: recall that I am assuming (by the "Parameter Assumptions") $\bar{v}_D - K(p^D(\bar{v}_D)) > 0$, and $p^1 > p^D(\bar{v}_D)$.

- For this equilibrium and all other Equilibrium listed below (i.e. Separating 2, Separating 3, etc), each type of D would escalate if $p \geq p^D(v_D)$. Type $\underline{v}_D$ (who is challenged) will not escalate. Type $\bar{v}_D$ (who is not challenged) would escalate if challenged.
- C's Beliefs: If $p < p^D(\bar{v}_D)$, then C believes D is low-type with probability 1. If $p \geq p^D(\bar{v}_D)$, then C believes D is high-type with probability 1.
- Payoffs: Type $\bar{v}_D$ attains $\bar{v}_D - K(p^D(\bar{v}_D))$, type $\underline{v}_D$ attains 0.

Proof of Equilibrium:
For type $\underline{v}_D$: Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. For any $p \in [p^D(\bar{v}_D), p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = p^D(\bar{v}_D)$ dominates all other arming levels. Due to $\underline{v}_D - K(p^D(\bar{v}_D)) \leq 0$, low-type prefer selecting $p = p_0$ to $p = p^D(\bar{v}_D)$.

For type $\bar{v}_D$: Because after selecting $p \in [p_0, p^D(\bar{v}_D))$, it would be optimal for D to acquiesce rather than fight, [30] following the logic above, high types choose between p_0 and $p^D(\bar{v}_D)$. From the Parameter Assumptions $\bar{v}_D - K(p^D(\bar{v}_D)) > 0$, meaning high types prefer selecting $p = p^D(\bar{v}_D)$ to $p = p_0$.

For C: For $p \in [p_0, p^D(\bar{v}_D))$, C believes D is a low-type and would acquiesce if challenged, which gives C a strictly positive payoff. For $p \in [p^D(\bar{v}_D), p_1]$, the selected $p \geq p^C$, meaning C does weakly better not challenging.

12.2 Separating Equilibrium 2:

- This equilibrium occurs when
(a) $p^D(\underline{v}_D) < p^C$, $p^D(\bar{v}_D) < p^C$, $max\,\{0, U_D(\hat{p}(\bar{v}_D))\} \leq \bar{v}_D - K(p^C)$, $0 \geq \underline{v}_D - K(p^C)$, $0 > U_D(\hat{p}(\underline{v}_D))$, $p^C \leq p_1$, $p^C > p_0$, or
(b) $p^D(\underline{v}_D) \geq p^C$,$p^D(\bar{v}_D) < p^C$, $max\,\{0, U_D(\hat{p}(\bar{v}_D))\} \leq \bar{v}_D - K(p^C)$, $0 \geq \underline{v}_D - K(p^C)$,[31] $p^C \leq p_1$, $p^C > p_0$,
- Type $\bar{v}_D$ selects $p^* = p^C$, and type $\underline{v}_D$ selects $p^* = p_0$.
- C will challenge for all $p < p^C$ and will not challenge for all $p \geq p^C$.
- Type $\underline{v}_D$ (who is challenged) will not escalate. Type $\bar{v}_D$ (who is not challenged) would escalate if challenged.

[30] This follows from the definition of $p^D(\bar{v}_D)$.
[31] Note: this only goes up to p^C instead of $p^D(\underline{v}_D)$ because in equilibrium play it will be the case that playing p^C will keep C from challenging.

- C's Beliefs: If $p < p^C$, then C believes D is low-type with probability 1. If $p \geq p^C$, then C believes D is high-type with probability 1.

- Payoffs: Type $\bar{v}_D$ attains $\bar{v}_D - K(p^C)$, type $\underline{v}_D$ attains 0.

Proof of Equilibrium:
For type $\underline{v}_D$:
Case (a). Within the range $p \in [p_0, p^D(\underline{v}_D))$, C will challenge and D will acquiesce, making D's utility is strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\underline{v}_D), p^C)$, C will challenge and D will fight. Thus, in this range, $p = \hat{p}(\underline{v}_D)$ weakly dominates all other arming levels.[32] Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case, $0 \geq \underline{v}_D - K(p^C)$ and $0 > U_D(\hat{p}(\underline{v}_D))$, implying that D prefers p_0 to fighting or deterring.

Case (b). Within the range $p \in [p_0, p^C)$, C will challenge and D will acquiesce, making D's utility is strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p_C, p_1]$, C will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case $0 \geq \underline{v}_D - K(p^C)$, implying that D prefers setting $p = p_0$ to deterring.

For type $\bar{v}_D$: Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), p^C)$, C will challenge and D will fight; thus, in this range $p = \hat{p}(\bar{v}_D)$ weakly dominates all other arming levels.[33] Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case, $max\,\{0, U_D(\hat{p}(\bar{v}_D))\} \leq \bar{v}_D - K(p^C)$, implying that D prefers setting $p = p^C$ and deterring to fighting or acquiescing.

For C: For $p \in [p_0, p^C)$, C believes D is a low-type and would acquiesce (when $p < p^D(\underline{v}_D)$) or fight (when $p \geq p^D(\underline{v}_D)$) if challenged; regardless, C attains a strictly positive payoff for challenging (based on the p^C condition). For $p \in [p^C, p_1]$, C believes that D is a high type and would fight if challenged, which gives C a weakly negative payoff for challenging (based on the p^C condition).

[32] If $\hat{p}(\underline{v}_D) = p^C$, then D would optimally select p^C and deter. This will prevent any open set issues within this range.
[33] If $\hat{p}(\bar{v}_D) = p^C$, then D would optimally select p^C and deter. This prevents open set issues over this range.

12.3 Separating Equilibrium 3:

- This equilibrium occurs when
  (a) $p^C \leq p_1$, $p^D(\bar{v}_D) < p^C$, $U_D(\hat{p}(\bar{v}_D)) > \bar{v}_D - K(p^C)$, $U_D(\hat{p}(\bar{v}_D)) \geq 0$, and $0 > U_D(\hat{p}(\underline{v}_D))$,[34] or
  (b) $p^C > p_1$, $0 \leq U_D(\hat{p}(\bar{v}_D))$, and $0 > U_D(\hat{p}(\underline{v}_D))$[35]

- Type $\bar{v}_D$ selects $p = \hat{p}(\bar{v}_D)$, and type $\underline{v}_D$ selects $p = p_0$.

- C will challenge for all $p < p^C$, and will not challenge for all $p \geq p^C$.

- Type $\underline{v}_D$ (who is challenged) will not escalate. Type $\bar{v}_D$ (who is challenged) will escalate.

- C's Beliefs: If $p < \hat{p}(\bar{v}_D)$, then C believes D is low-type with probability 1. If $p \geq \hat{p}(\bar{v}_D)$, then C believes D is high-type with probability 1.

- Payoffs: Type $\bar{v}_D$ attains $U_D(\hat{p}(\bar{v}_D))$, type $\underline{v}_D$ attains 0.

Proof of Equilibrium:
For type $\underline{v}_D$:
Case (a),
Case (a.1). In addition to the conditions on case (a), assume $p^C > p^D(\underline{v}_D)$ and $\underline{v}_D - K(p^D(\underline{v}_D)) < 0$. Within the range $p \in [p_0, p^D(\underline{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\underline{v}_D), p^C)$, C will challenge and D will fight. Thus, in this range, $p = \hat{p}(\underline{v}_D)$ weakly dominates all other arming levels.[36] Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case, $0 > U_D(\hat{p}(\underline{v}_D))$, implying that D prefers p_0 to fighting. Additionally, I can use the conditions of this subcase $p^C > p^D(\underline{v}_D)$ and $\underline{v}_D - K(p^D(\underline{v}_D)) < 0$, which together imply $\underline{v}_D - K(p^C) < 0$, or that D prefers setting $p = p_0$ and acquiescing to setting $p = p^C$ and deterring.

Case (a.2). Assume $p^C > p^D(\underline{v}_D)$ and $\underline{v}_D - K(p^D(\underline{v}_D)) \geq 0$. This proof is identical up to the point before demonstrating that D prefers setting $p = p_0$ and acquiescing to setting $p = p^C$ and deterring. As I showed in Lemma 1, within the parameter set where $p^C > p^D(\underline{v}_D)$ and $\underline{v}_D - K(p^D(\underline{v}_D)) \geq 0$, D's selected p^* is non-decreasing in v_D. Thus, because high types optimally fight (as I discuss below), low-types would never prefer setting $p = p^C$ and deterring.

[34]This is assisted by Remark 1.
[35]Note: recall that I am assuming (by the "Parameter Assumptions") $\bar{v}_D - K(p^D(\bar{v}_D)) > 0$, and $p^1 > p^D(\bar{v}_D)$.
[36]If $\hat{p}(\underline{v}_D) = p^C$, then D would optimally select p^C and deter. This will prevent any open set issues within this range.

Case (a.3) Assume $p^D(\underline{v}_D) \geq p^C$. Within the range $p \in [p_0, p^C)$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. I can demonstrate that low-type D's always prefer setting $p = p_0$ to $p = p^C$. Starting with $U_D(\hat{p}(\bar{v}_D)) > \bar{v}_D - K(p^C)$, which is given by the conditions of the case, I use the definition of $U_D(\hat{p}(\bar{v}_D))$ to say

$$-\frac{\hat{p}(\bar{v}_D)(1-\hat{p}(\bar{v}_D))}{\alpha + n\hat{p}(\bar{v}_D)(1-\hat{p}(\bar{v}_D))}(nN_D + c_D) + \frac{\alpha\hat{p}(\bar{v}_D)}{\alpha + n\hat{p}(\bar{v}_D)(1-\hat{p}(\bar{v}_D))}\bar{v}_D - K(\hat{p}(\bar{v}_D)) > \bar{v}_D - K(p^C).$$

Because $\frac{\alpha\hat{p}(\bar{v}_D)}{\alpha+n\hat{p}(\bar{v}_D)(1-\hat{p}(\bar{v}_D))}(\bar{v}_D - \underline{v}_D) < \bar{v}_D - \underline{v}_D$, I can say

$$-\frac{\hat{p}(\bar{v}_D)(1-\hat{p}(\bar{v}_D))}{\alpha + n\hat{p}(\bar{v}_D)(1-\hat{p}(\bar{v}_D))}(nN_D + c_D) + \frac{\alpha\hat{p}(\bar{v}_D)}{\alpha + n\hat{p}(\bar{v}_D)(1-\hat{p}(\bar{v}_D))}\underline{v}_D - K(\hat{p}(\bar{v}_D)) > \underline{v}_D - K(p^C).$$

As how $\hat{p}(\bar{v}_D)$ is defined, it must be that $\hat{p}(\bar{v}_D) \leq p^C$, implying (by the conditions of case a.3) $\hat{p}(\bar{v}_D) \leq p^D(\underline{v}_D)$. Thus, from how $p^D(\underline{v}_D)$ is defined, $-\frac{\hat{p}(\bar{v}_D)(1-\hat{p}(\bar{v}_D))}{\alpha+n\hat{p}(\bar{v}_D)(1-\hat{p}(\bar{v}_D))}(nN_D + c_D) + \frac{\alpha\hat{p}(\bar{v}_D)}{\alpha+n\hat{p}(\bar{v}_D)(1-\hat{p}(\bar{v}_D))}\underline{v}_D \leq 0$. This in turn implies

$$0 - K(\hat{p}(\bar{v}_D)) > \underline{v}_D - K(p^C),$$

or $0 > \underline{v}_D - K(p^C)$. Thus, D prefers setting $p = p_0$ and acquiescing to $p = p^C$ and deterring.

Case b.
Case (b.1) In addition to the conditions on case (b), assume $p_1 \geq p^D(\underline{v}_D)$. Within the range $p \in [p_0, p^D(\underline{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\underline{v}_D), p_1]$ C will challenge and D will fight. Thus, in this range, $p = \hat{p}(\underline{v}_D)$ weakly dominates all other arming levels. By the conditions of the case, $0 > U_D(\hat{p}(\underline{v}_D))$, implying that D prefers setting $p = p_0$ to fighting.

Case (b.2) In addition to the conditions on case (b), assume $p_1 < p^D(\underline{v}_D)$. Within the range $p \in [p_0, p_1]$, C will challenge and D will acquiesce, making D's utility is strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels.

For type $\bar{v}_D$:
Case (a) Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), p^C)$, C will challenge and D will fight. Thus, in this

range, $p = \hat{p}(\bar{v}_D)$ weakly dominates all other arming levels.[37] Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case, $U_D(\hat{p}(\bar{v}_D)) > \bar{v}_D - K(p^C)$ and $U_D(\hat{p}(\bar{v}_D)) \geq 0$, implying that D prefers selecting $p = \hat{p}(\bar{v}_D)$ and fighting.

Case (b) Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), p_1]$, C will challenge and D will fight. Thus, in this range, $p = \hat{p}(\bar{v}_D)$ weakly dominates all other arming levels. By the conditions of the case $U_D(\hat{p}(\bar{v}_D)) \geq 0$, D prefers selecting $p = \hat{p}(\bar{v}_D)$ and fighting.

For C: For $p \in [p_0, \hat{p}(\bar{v}_D))$, C believes D is a low-type and would acquiesce if challenged, which gives C a strictly positive payoff for challenging (based on $\hat{p}(\bar{v}_D) < p^C$).

Case (a) For $p \in [\hat{p}(\bar{v}_D), p^C)$, C believes that D is a high type and would fight if challenged, which gives C a weakly positive payoff for challenging (based on the p^C condition). For $p \in [p^C, p_1]$, C believes that D is a high type and would fight if challenged, which gives C a weakly negative payoff for challenging (based on the p^C condition).

Case (b) For $p \in [\hat{p}(\bar{v}_D), p_1)$, C believes that D is a high type and would fight if challenged, which gives C a weakly positive payoff for challenging (based on the $p_1 < p^C$ condition).

12.4 Separating Equilibrium 4

- This equilibrium occurs when
  (a) $p^C \leq p^D(\bar{v}_D)$, $\underline{v}_D - K(p^D(\bar{v}_D)) > 0$, if $p^D(\underline{v}_D) \leq p_1$ then $\underline{v}_D - K(p^D(\underline{v}_D)) \leq 0$, $\bar{p} \leq p_1$, and, if $\tilde{P}$ is non-empty, $\underline{v}_D - K(\tilde{p}) \leq 0$
  (b) $p^C > p^D(\bar{v}_D)$, $\underline{v}_D - K(p^C) > 0$, if $p^D(\underline{v}_D) \leq p_1$ then $\underline{v}_D - K(p^D(\underline{v}_D)) \leq 0$, $\bar{p} \leq p_1$, and, if $\tilde{P}$ is non-empty, $\underline{v}_D - K(\tilde{p}) \leq 0$

- Type $\bar{v}_D$ selects $p = \bar{p}$, and type $\underline{v}_D$ selects $p = p_0$.

- C will challenge for all $p < \bar{p}$, and will not challenge for all $p \geq \bar{p}$.

- Type $\underline{v}_D$ (who is challenged) will not escalate. Type $\bar{v}_D$ is not challenged..

- C's Beliefs: If $p < \bar{p}$, then C believes D is low-type with probability 1. If $p \geq \bar{p}$, then C

[37] If $\hat{p}(\bar{v}_D) = p^C$, then D would optimally select p^C and deter. This will prevent any open set issues within this range.

believes D is high-type with probability 1.

- Payoffs: Type $\bar{v}_D$ attains $\bar{v}_D - K(\bar{p})$, type $\underline{v}_D$ attains 0.

Proof of Equilibrium:

For type $\underline{v}_D$: Within the range $p \in [p_0, \bar{p})$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. For any $p \in [\bar{p}, p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = \bar{p}$ dominates all other arming levels. By it's definition, $\underline{v}_D - K(\bar{p}) = 0$, meaning low-type D's weakly prefer selecting $p = p_0$ and acquiescing to $p = \bar{p}$ and bluffing.

For type $\bar{v}_D$: Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), \bar{p})$, C will challenge and D will fight. Thus, in this range, there exists some arming level or set of arming levels that dominates all others.[38] Within the range $p \in [\bar{p}, p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = \bar{p}$ dominates all other arming levels. Because $\underline{v}_D - K(\bar{p}) = 0$, it must be that $\bar{v}_D - K(\bar{p}) > 0$, meaning D prefers arming to $p = \bar{p}$ to setting $p = p_0$. To demonstrate that D prefers setting $p = \bar{p}$ to fighting, I must first define the value $\dot{p}$ as type $\bar{v}_D$'s optimal arming level conditional on the high type looking to fight, or

$$\dot{p} \in argmax_{p \in [p^D(\bar{v}_D), \bar{p}]} \left\{ -\frac{\dot{p}(1-\dot{p})}{\alpha + n\dot{p}(1-\dot{p})} (nN_D + c_D) + \frac{\alpha \dot{p}}{\alpha + n\dot{p}(1-\dot{p})} \bar{v}_D - K(\dot{p}) \right\}.$$

Note that because $\underline{v}_D - K(\bar{p}) = 0$ and $\underline{v}_D - K(p^D(\underline{v}_D)) \leq 0$, it must be that $\bar{p} \leq p^D(\underline{v}_D)$, meaning $\dot{p} \leq p^D(\underline{v}_D)$.

I start with a condition which follows from how $\bar{p}$ is defined:

$$\underline{v}_D - K(\bar{p}) = 0.$$

Using that $\dot{p} \leq p^D(\underline{v}_D)$, it must be that

$$\underline{v}_D - K(\bar{p}) \geq -\frac{\dot{p}(1-\dot{p})}{\alpha + n\dot{p}(1-\dot{p})} (nN_D + c_D) + \frac{\alpha \dot{p}}{\alpha + n\dot{p}(1-\dot{p})} \underline{v}_D - K(\dot{p}).$$

[38]Note this will not be $p = \hat{p}(\bar{v}_D)$ because the utility function is optimized over a different domain. Also, following the rationale discussed in prior footnotes, there will not be open set issues here.

Because $\bar{v}_D - \underline{v}_D > (\bar{v}_D - \underline{v}_D)\left(\frac{\alpha\dot{p}}{\alpha+n\dot{p}(1-\dot{p})}\right)$, I can say

$$\bar{v}_D - K(\bar{p}) \geq -\frac{\dot{p}(1-\dot{p})}{\alpha + n\dot{p}(1-\dot{p})}(nN_D + c_D) + \frac{\alpha\dot{p}}{\alpha + n\dot{p}(1-\dot{p})}\bar{v}_D - K(\dot{p}),$$

which implies that D always prefers setting $p = \bar{p}$ and deterring to selecting $p = \dot{p}$ and fighting.

For C: For $p \in [p_0, \bar{p})$, C believes D is a low-type and would acquiesce if challenged, which gives C a strictly positive payoff. For $p \in [\bar{p}, p_1]$, C believes D is a high-type and would fight if challenged. For both cases, $\bar{p} \geq p^C$,[39] meaning C would prefer to acquiesce rather than fight with arming level $\bar{p}$.

12.5 Separating 5 Equilibrium:

- This equilibrium occurs when
  (a) $p^C \leq p_1$, $p^D(\underline{v}_D) < p^C$, $U_D(\hat{p}(\bar{v}_D)) > \bar{v}_D - K(p^C)$, $U_D(\hat{p}(\bar{v}_D)) \geq 0$, and $U_D(\hat{p}(\underline{v}_D)) \geq 0$, or[40]
  (b) $p^C > p_1$, $p^D(\underline{v}_D) \leq p_1$, $0 \leq U_D(\hat{p}(\bar{v}_D))$, and $0 \leq U_D(\hat{p}(\underline{v}_D))$.
- Type $\bar{v}_D$ selects $p = \hat{p}(\bar{v}_D)$, and type $\underline{v}_D$ selects $p = \hat{p}(\underline{v}_D)$.
- C will challenge for all $p < p^C$, and will not challenge for all $p \geq p^C$.
- Both types will escalate when challenged.
- C's Beliefs: If $p < \hat{p}(\bar{v}_D)$, then C believes D is low-type with probability 1. If $p \geq \hat{p}(\bar{v}_D)$, then C believes D is high-type with probability 1.
- Payoffs: Type $\bar{v}_D$ attains $U_D(\hat{p}(\bar{v}_D))$, type $\underline{v}_D$ attains $U_D(\hat{p}(\underline{v}_D))$.

Proof of Equilibrium
Case (a).
For type $\underline{v}_D$: Within the range $p \in [p_0, p^D(\underline{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\underline{v}_D), p^C)$, C will challenge and D will fight. Thus, in this range, $p = \hat{p}(\underline{v}_D)$ weakly dominates all other arming levels. Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case, $U_D(\hat{p}(\underline{v}_D)) \geq 0$, implying that

[39] In case (a), this follows from how $\bar{p}$ is defined and $p^C \leq p^D(\bar{v}_D)$ and $\underline{v}_D - K(p^D(\bar{v}_D)) > 0$. In case (b), this follows from how $\bar{p}$ is defined and $\underline{v}_D - K(p^C) > 0$.
[40] The part on low-types only choosing between fight and acquiesce relies on Remark 1.

D prefers setting $p = \hat{p}(\underline{v}_D)$ and fighting to $p = p_0$ and acquiescing. The remainder relies on utilizing Lemma 1. I now show that the above conditions imply $\underline{v}_D - K(p^D(\underline{v}_D)) > 0$, which is needed for Lemma 1 to apply. Using $U_D(\hat{p}(\underline{v}_D)) \geq 0$, I can say

$$-\frac{\hat{p}(\underline{v}_D)(1-\hat{p}(\underline{v}_D))}{\alpha + n\hat{p}(\underline{v}_D)(1-\hat{p}(\underline{v}_D))}(nN_D + c_D) + \frac{\alpha\hat{p}(\underline{v}_D)}{\alpha + n\hat{p}(\underline{v}_D)(1-\hat{p}(\underline{v}_D))}\underline{v}_D - K(\hat{p}(\underline{v}_D)) \geq 0.$$

This implies that

$$\underline{v}_D - K(\hat{p}(\underline{v}_D)) > 0.$$

Because $\hat{p}(\underline{v}_D) \in [p^D(\underline{v}_D), p^C]$, I can say

$$\underline{v}_D - K(p^D(\underline{v}_D)) > 0.$$

Thus, Lemma 1 can apply here. Because high types prefer setting $p = \hat{p}(\bar{v}_D)$ and fighting to setting $p = p^C$ and deterring, low-types will never set $p = p^C$.

For type $\bar{v}_D$: Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), p^C)$, C will challenge and D will fight; thus, in this range $p = \hat{p}(\bar{v}_D)$ weakly dominates all other arming levels.[41] Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case, $U_D(\hat{p}(\bar{v}_D)) > \bar{v}_D - K(p^C)$ and $U_D(\hat{p}(\bar{v}_D)) \geq 0$ implying that D prefers selecting $\hat{p}(\bar{v}_D)$ and fighting to deterring or acquiescing.

For C: For $p \in [p_0, \hat{p}(\bar{v}_D))$, C believes D is a low-type and would acquiesce if challenged, which gives C a strictly positive payoff for challenging (based on $\hat{p}(\bar{v}_D) < p^C$). For $p \in [\hat{p}(\bar{v}_D), p^C)$, C believes that D is a high type and would fight if challenged, which gives C a weakly positive payoff for challenging (based on the p^C condition). For $p \in [p^C, p_1]$, C believes that D is a high type and would fight if challenged, which gives C a weakly negative payoff for challenging (based on the p^C condition).

Case (b). The proof is nearly identical, other than D can no longer select some $p \geq p^C$ and deter C.

[41] If $\hat{p}(\bar{v}_D) = p^C$, then D would optimally select p^C and deter. This will prevent any open set issues within this range.

12.6 Separating 6 Equilibrium:

- This equilibrium occurs when $p^D(\underline{v}_D) < p^C$, $p^C \leq p_1$, $\bar{v}_D - K(p^C) \geq U_D(\hat{p}(\bar{v}_D))$, $U_D(\hat{p}(\underline{v}_D)) \geq 0$ and $U_D(\hat{p}(\underline{v}_D)) > \underline{v}_D - K(p^C)$
- Type $\bar{v}_D$ selects $p = p^C$, and type $\underline{v}_D$ selects $p = \hat{p}(\underline{v}_D)$.
- C will challenge for all $p < p^C$, and will not challenge for all $p \geq p^C$.
- When challenged, low-types will escalate
- C's Beliefs: If $p < p^C$, then C believes D is low-type with probability 1. If $p \geq p^C$, then C believes D is high-type with probability 1.
- Payoffs: Type $\bar{v}_D$ attains $\bar{v}_D - K(p^C)$, type $\underline{v}_D$ attains $U_D(\hat{p}(\underline{v}_D))$.

Proof of equilibrium:
For type $\underline{v}_D$: Within the range $p \in [p_0, p^D(\underline{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\underline{v}_D), p^C)$, C will challenge and D will fight. Thus, in this range, $p = \hat{p}(\underline{v}_D)$ weakly dominates all other arming levels. Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case, $U_D(\hat{p}(\underline{v}_D)) \geq 0$ and $U_D(\hat{p}(\underline{v}_D)) > \underline{v}_D - K(p^C)$, implying that D prefers setting $p = \hat{p}(\underline{v}_D)$ and fighting to $p = p_0$ and acquiescing or p^C and deterring.

For type $\bar{v}_D$: Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), p^C)$, C will challenge and D will fight; thus, in this range $p = \hat{p}(\bar{v}_D)$ weakly dominates all other arming levels.[42] Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case, $\bar{v}_D - K(p^C) \geq U_D(\hat{p}(\bar{v}_D))$, implying that D prefers selecting $p = p^C$ and deterring to selecting $p = \hat{p}(\bar{v}_D)$ and fighting. And, as shown in the discussion of the Separating 5 equilibrium, $U_D(\hat{p}(\underline{v}_D)) \geq 0$ implies that $\underline{v}_D - K(p^D(\underline{v}_D)) > 0$. This means that Separating 6 falls within conditions for Lemma 1. Thus, D will never select $p = p_0$ and acquiesce because low types select $p = \hat{p}(\underline{v}_D)$.

For C: For $p \in [p_0, p^C)$, C believes D is a low-type and would either acquiesce if challenged or

[42] If $\hat{p}(\bar{v}_D) = p^C$, then D would optimally select p^C and deter. This will prevent any open set issues within this range.

fight when challenged: both give C a strictly positive payoff for challenging. For $p \in [p^C, p^1)$, C believes that D is a high type and would fight if challenged, which gives C a weakly negative payoff for challenging (based on the p^C condition).

12.7 Pooling Equilibrium 1:

- This equilibrium occurs when
(a) $p^C \leq p_1$, $p^D(\bar{v}_D) < p^C$, $p^C > p_0$, $p_0 < p^D(\bar{v}_D)$, $0 > \bar{U}_D(\hat{p}(\bar{v}_D))$, $0 > \bar{v}_D - K(p^C)$ or
(b) $p^C > p_1$, $p_0 < p^D(\bar{v}_D)$, $0 > \bar{U}_D(\hat{p}(\bar{v}_D))$
- Type $\bar{v}_D$ selects $p = p_0$, and type $\underline{v}_D$ selects $p = p_0$.
- C will challenge for all $p < p^C$, and will not challenge for all $p \geq p^C$.
- Neither type will escalate when challenged.
- C's Beliefs: If $p = p_0$, then C believes D is low-type with probability $1 - \pi$ and high-type with probability π. If $p \neq p_0$ and $p < p^C$, then C believes D is low-type with probability 1. If $p \geq p^C$, then C believes D is high-type with probability 1.
- Payoffs: Type $\bar{v}_D$ attains 0, type $\underline{v}_D$ attains 0.

Proof of Equilibrium
Case (a).
For type $\underline{v}_D$:
Case (a.1) In addition to the assumptions on case (a), also assume that $\hat{p}(\underline{v}_D) < p^C$. Within the range $p \in [p_0, p^D(\underline{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\underline{v}_D), p^C)$, C will challenge and D will fight. Thus, in this range, $p = \hat{p}(\underline{v}_D)$ weakly dominates all other arming levels. Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. Because high type D's prefer acquiescing to deterring ($0 > \bar{v}_D - K(p^C)$), it implies that low-types also prefer acquiescing to deterring. I can also that type $\underline{v}_D$ prefers setting $p = p_0$ and acquiescing to fighting. I start with $0 > \bar{U}_D(\hat{p}(\bar{v}_D))$, or

$$0 > -\frac{\hat{p}(\bar{v}_D)(1 - \hat{p}(\bar{v}_D))}{\alpha + n\hat{p}(\bar{v}_D)(1 - \hat{p}(\bar{v}_D))}(nN_D + c_D) + \frac{\alpha\hat{p}(\bar{v}_D)}{\alpha + n\hat{p}(\bar{v}_D)(1 - \hat{p}(\bar{v}_D))}\bar{v}_D - K(\hat{p}(\bar{v}_D)).$$

Using that $\hat{p}(\bar{v}_D)$ optimizes the expression on the right and using that $\hat{p}(\underline{v}_D) \in [p^D(\bar{v}_D), p^C]$,[43]

[43]Because $\hat{p}(\underline{v}_D) \in [p^D(\underline{v}_D), p^C]$, $p^D(\bar{v}_D) < p^D(\underline{v}_D)$, and $\hat{p}(\underline{v}_D) < p^C$.

I can say

$$0 > -\frac{\hat{p}(\underline{v}_D)(1-\hat{p}(\underline{v}_D))}{\alpha + n\hat{p}(\underline{v}_D)(1-\hat{p}(\underline{v}_D))}(nN_D + c_D) + \frac{\alpha \hat{p}(\underline{v}_D)}{\alpha + n\hat{p}(\underline{v}_D)(1-\hat{p}(\underline{v}_D))}\bar{v}_D - K(\hat{p}(\underline{v}_D)),$$

and also

$$0 > -\frac{\hat{p}(\underline{v}_D)(1-\hat{p}(\underline{v}_D))}{\alpha + n\hat{p}(\underline{v}_D)(1-\hat{p}(\underline{v}_D))}(nN_D + c_D) + \frac{\alpha \hat{p}(\underline{v}_D)}{\alpha + n\hat{p}(\underline{v}_D)(1-\hat{p}(\underline{v}_D))}\underline{v}_D - K(\hat{p}(\underline{v}_D)),$$

Thus, type $\underline{v}_D$ prefers setting $p = p_0$ and acquiescing to fighting.

Case (a.2) Assume that $\hat{p}(\underline{v}_D) \geq p^C$. Within the range $p \in [p_0, p^C)$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. Because $0 > \bar{v}_D - K(p^C)$, it implies that type $\underline{v}_D$ prefers setting $p = p_0$ and acquiescing to setting $p = p^C$ and deterring.

For type $\bar{v}_D$: Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), p^C)$, C will challenge and D will fight; thus, in this range $p = \hat{p}(\bar{v}_D)$ weakly dominates all other arming levels.[44] Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case, $0 > \bar{U}_D(\hat{p}(\bar{v}_D))$ and $0 > \bar{v}_D - K(p^C)$, implying that D prefers selecting p_0 and acquiescing to fighting or deterring .

For C: For $p = p_0$, C believes both types of D are selecting this arming level; thus, C's beliefs on type are the prior distribution of type. For $p \in (p_0, p^C)$, C believes D is a low-type and would acquiesce if challenged, which gives C a strictly positive payoff for challenging (based on $\hat{p}(\bar{v}_D) < p^C$). For $p \in [p^C, p_1]$, C believes that D is a high type and would fight if challenged, which gives C a weakly negative payoff for challenging (based on the p^C condition).

Case (b). The proof is nearly identical, other than D can no longer select some $p \geq p^C$ and deter C, and C believes that D is a low type for selecting any $p \in (p_0, p_1]$.

[44] If $\hat{p}(\bar{v}_D) = p^C$, then D would optimally select p^C and deter. This prevents open set issues over this range.

12.8 Pooling Equilibrium 2:

- This equilibrium occurs when $\underline{v}_D - K(max\left\{p^D(\bar{v}_D), p_0\right\}) > 0$, $p^C \leq max\left\{p^D(\bar{v}_D), p_0\right\}$,and the set $\tilde{P}$ is non-empty and $\tilde{p} \leq max\left\{p^D(\bar{v}_D), p_0\right\}$,[45]
- Type $\bar{v}_D$ selects $p = max\left\{p^D(\bar{v}_D),\ p_0\right\}$, and type $\underline{v}_D$ selects $p = max\left\{p^D(\bar{v}_D),\ p_0\right\}$.
- C will not challenge when observing $p \geq max\left\{p^D(\bar{v}_D),\ p_0\right\}$. C will challenge when observing $p < max\left\{p^D(\bar{v}_D),\ p_0\right\}$.
- Type $\underline{v}_D$ (who is not challenged) would not escalate if challenged. Type $\bar{v}_D$ (who is not challenged) would escalate if challenged.
- C's Beliefs: If $p = max\left\{p^D(\bar{v}_D),\ p_0\right\}$, then C believes D is low-type with probability $1 - \pi$ and high-type with probability π. If $p < max\left\{p^D(\bar{v}_D),\ p_0\right\}$,[46] then C believes D is low-type with probability 1. If $p > max\left\{p^D(\bar{v}_D),\ p_0\right\}$, then C believes D is high-type with probability 1.
- Payoffs: Type $\bar{v}_D$ attains $\bar{v}_D - K(max\left\{p^D(\bar{v}_D),\ p_0\right\})$, type $\underline{v}_D$ attains $\underline{v}_D - K(max\left\{p^D(\bar{v}_D),\ p_0\right\})$.

Proof of Equilibrium
For type $\underline{v}_D$:
Case 1. In addition to the assumptions, also assume that $p_0 < p^D(\bar{v}_D)$. Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), p_1]$, C will not challenge. Thus, in this range, $p = p^D(\bar{v}_D)$ dominates all other arming levels. By assumption $\underline{v}_D - K(p^D(\bar{v}_D)) > 0$, meaning D prefers setting $p = p^D(\bar{v}_D)$ and bluffing to setting $p = p_0$ and acquiescing.[47]

Case 2. Assume $p_0 \geq p^D(\bar{v}_D)$. For all $p \in [p_0, p_1]$, C will not challenge. Thus, in this range, $p = p_0$ dominates all other arming levels.

For type $\bar{v}_D$:
Case 1. In addition to the assumptions, also assume that $p_0 < p^D(\bar{v}_D)$. Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), p_1]$, C will not challenge. Thus, in this range, $p = p^D(\bar{v}_D)$ dominates all other

[45] Note: recall that I am assuming (by the "Parameter Assumptions") $\bar{v}_D - K(max\left\{p^D(\bar{v}_D), p_0\right\}) > 0$, and $p^1 > p^D(\bar{v}_D)$.

[46] Needless to say this belief structure could describe non-feasible actions (when $p_0 > p^D(\bar{v}_D)$). I keep this in place for simplicity.

[47] Because $p^D(\bar{v}_D) < p^D(\underline{v}_D)$, D prefers not fighting.

arming levels. By the Parameter Assumptions, $\bar{v}_D - K(p^D(\bar{v}_D)) > 0$, meaning D prefers setting $p = p^D(\bar{v}_D)$ and bluffing to setting $p = p_0$ and acquiescing.

Case 2. Assume $p_0 \geq p^D(\bar{v}_D)$. For all $p \in [p_0, p_1]$, C will not challenge. Thus, in this range, $p = p_0$ dominates all other arming levels.

For C: For $p = max\left\{p^D(\bar{v}_D),\, p_0\right\}$, C believes both types of D are selecting this arming level; thus, C's beliefs on type are the prior distribution of type. For $p \in [p_0, p^D(\bar{v}_D))$ (whenever $p_0 < p^D(\bar{v}_D)$), C believes D is a low-type and would acquiesce if challenged, which gives C a strictly positive payoff for challenging. For $p \in (p^D(\bar{v}_D), p_1]$, C believes that D is a high type and would fight if challenged, which gives C a weakly negative payoff for challenging (because $p^C < p^D(\bar{v}_D)$).

12.9 Pooling Equilibrium 3:

- This equilibrium occurs when the set $\tilde{P}$ is non-empty, $\tilde{p} > max\left\{p^D(\bar{v}_D),\, p_0\right\}$, $\underline{v}_D - K(\tilde{p}) > 0$, $\tilde{p} < p^D(\underline{v}_D)$, and $\tilde{p} \leq p_1$.
- Type $\bar{v}_D$ selects $p = \tilde{p}$, and type $\underline{v}_D$ selects $p = \tilde{p}$.
- C will not challenge when observing $p = \tilde{p}$, will challenge when observing $p < \tilde{p}$, and will not challenge when observing $p > \tilde{p}$.
- Type $\underline{v}_D$ (who is not challenged) would not escalate if challenged. Type $\bar{v}_D$ (who is not challenged) would escalate if challenged.
- C's Beliefs: If $p = \tilde{p}$, then C believes D is low-type with probability $1 - \pi$ and high-type with probability π. If $p < \tilde{p}$, then C believes D is low-type with probability 1. If $p > \tilde{p}$, then C believes D is high-type with probability 1.
- Payoffs: Type $\bar{v}_D$ attains $\bar{v}_D - K(\tilde{p})$, type $\underline{v}_D$ attains $\underline{v}_D - K(\tilde{p})$.

Proof of Equilibrium

For type $\underline{v}_D$: Within the range $p \in [p_0, \tilde{p})$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [\tilde{p}, p_1]$, C will not challenge. Thus, in this range, $p = \tilde{p}$ dominates all other arming levels. By assumption $\underline{v}_D - K(\tilde{p}) > 0$, meaning D prefers setting $p = \tilde{p}$ and bluffing to setting $p = p_0$ and acquiescing.[48]

For type $\bar{v}_D$: Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming

[48]Because $\tilde{p} < p^D(\underline{v}_D)$, D prefers not fighting.

levels. Within the range $p \in [p^D(\bar{v}_D), \tilde{p})$, C will challenge and D will fight, selecting some optimal arming level. In the range $p \in [\tilde{p}, p_1]$, C will not challenge. Thus, in this range, $p = \tilde{p}$ dominates all other arming levels. By assumption $\underline{v}_D - K(\tilde{p}) > 0$, implying that high types prefer setting $p = \tilde{p}$ and deterring to setting $p = p_0$ and acquiescing. To demonstrate that high types would never select $p \in [p^D(\bar{v}_D), \tilde{p})$ and fight, I define re-define $\dot{p}$ as

$$\dot{p} \in argmax_{p \in [p^D(\bar{v}_D), \tilde{p}]} \left\{ -\frac{\dot{p}(1-\dot{p})}{\alpha + n\dot{p}(1-\dot{p})} (nN_D + c_D) + \frac{\alpha \dot{p}}{\alpha + n\dot{p}(1-\dot{p})} \bar{v}_D - K(\dot{p}) \right\}.$$

I start using $\underline{v}_D - K(\tilde{p}) > 0$, which is given, and that $\tilde{p} < p^D(\underline{v}_D)$, which implies that low-types would do strictly worse selecting some $\dot{p}$ and fighting relative to setting $p = p_0$ and acquiescing. Using this observation and $\underline{v}_D - K(\tilde{p}) > 0$ gives

$$\underline{v}_D - K(\tilde{p}) > -\frac{\dot{p}(1-\dot{p})}{\alpha + n\dot{p}(1-\dot{p})} (nN_D + c_D) + \frac{\alpha \dot{p}}{\alpha + n\dot{p}(1-\dot{p})} \underline{v}_D - K(\dot{p})$$

I can then use that $\bar{v}_D - \underline{v}_D > (\bar{v}_D - \underline{v}_D) \left(\frac{\alpha \dot{p}}{\alpha + n\dot{p}(1-\dot{p})} \right)$, which gives

$$\bar{v}_D - K(\tilde{p}) > -\frac{\dot{p}(1-\dot{p})}{\alpha + n\dot{p}(1-\dot{p})} (nN_D + c_D) + \frac{\alpha \dot{p}}{\alpha + n\dot{p}(1-\dot{p})} \bar{v}_D - K(\dot{p}).$$

Thus, high type D's prefer arming to level $\tilde{p}$ and deterring to fighting.

For C: For $p = [p_0, \tilde{p})$, C believes D is a low-type and would acquiesce if challenged, which gives C a strictly positive payoff for challenging. For $p \in (\tilde{p}, p_1]$, C believes that D is a high type and would fight if challenged, which gives C a weakly negative payoff for challenging. For $p = \tilde{p}$, C's beliefs follow the priors, and C prefers not challenging based on how $\tilde{p}$ is defined.

12.10 Pooling Equilibrium 4:

- This equilibrium occurs when $max \left\{ p^D(\underline{v}_D), p_0 \right\} \geq p^C$, $\underline{v}_D - K(max \left\{ p^D(\underline{v}_D), p_0 \right\}) > 0$, $p^D(\underline{v}_D) \leq p_1$, and, when the set of $\tilde{P}$ is non-empty, $\tilde{p} \geq max \left\{ p^D(\underline{v}_D), p_0 \right\}$.
- Type $\bar{v}_D$ selects $p = max \left\{ p^D(\underline{v}_D), p_0 \right\}$, and type $\underline{v}_D$ selects $p = max \left\{ p^D(\underline{v}_D), p_0 \right\}$.
- C will not challenge when observing $p = max \left\{ p^D(\underline{v}_D), p_0 \right\}$, will challenge when observing $p < max \left\{ p^D(\underline{v}_D), p_0 \right\}$, and will not challenge when observing $p > max \left\{ p^D(\underline{v}_D), p_0 \right\}$.
- Both types would escalate if challenged.
- C's Beliefs: If $p = max \left\{ p^D(\underline{v}_D), p_0 \right\}$, then C believes D is low-type with probability $1 - \pi$ and high-type with probability π. If $p < max \left\{ p^D(\underline{v}_D), p_0 \right\}$ then C believes D is

low-type with probability 1. If $p > max\left\{p^D(\underline{v}_D), p_0\right\}$, then C believes D is a high-type with probability 1.

- Payoffs: Type $\bar{v}_D$ attains $\bar{v}_D - K(max\left\{p^D(\underline{v}_D), p_0\right\})$, type $\underline{v}_D$ attains $\underline{v}_D - K(max\left\{p^D(\underline{v}_D), p_0\right\})$.

Proof of Equilibrium

For type $\underline{v}_D$:
Case 1. In addition to the assumptions, also assume that $p_0 < p^D(\underline{v}_D)$. Within the range $p \in [p_0, p^D(\underline{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\underline{v}_D), p_1]$, C will not challenge. Thus, in this range, $p = p^D(\bar{v}_D)$ dominates all other arming levels. By assumption $\underline{v}_D - K(p^D(\underline{v}_D)) > 0$, meaning D prefers setting $p = p^D(\underline{v}_D)$ and deterring to setting $p = p_0$ and acquiescing.

Case 2. Assume $p_0 \geq p^D(\underline{v}_D)$. For all $p \in [p_0, p_1]$, C will not challenge. Thus, in this range, $p = p_0$ dominates all other arming levels.

For type $\bar{v}_D$:
Case 1. In addition to the assumptions, also assume that $p_0 < p^D(\underline{v}_D)$. I also assume $p^D(\bar{v}_D) > p_0$; relaxing this makes little difference to the proof, so I will not discuss this alternate case. Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), p^D(\underline{v}_D))$, C will challenge and D will fight, selecting some optimal arming level. In the range $p \in [p^D(\underline{v}_D), p_1]$, C will not challenge. Thus, in this range, $p = p^D(\underline{v}_D)$ dominates all other arming levels. By assumption $\underline{v}_D - K(p^D(\underline{v}_D)) > 0$, implying that high types also prefer setting $p = p^D(\underline{v}_D)$ and deterring to setting $p = p_0$ and acquiescing. To demonstrate that high types would never select $p \in [p^D(\bar{v}_D), p^D(\underline{v}_D))$ and fight, I define re-define $\dot{p}$ as

$$\dot{p} \in argmax_{p \in [p^D(\bar{v}_D), p^D(\underline{v}_D)]} \left\{ -\frac{\dot{p}(1-\dot{p})}{\alpha + n\dot{p}(1-\dot{p})} (nN_D + c_D) + \frac{\alpha \dot{p}}{\alpha + n\dot{p}(1-\dot{p})} \bar{v}_D - K(\dot{p}) \right\}.$$

I start using $\underline{v}_D - K(p^D(\underline{v}_D)) > 0$, which is given, and that $\dot{p} < p^D(\underline{v}_D)$, which implies that low-types would do strictly worse selecting some $\dot{p}$ and fighting relative to setting $p = p_0$ and acquiescing.

$$\underline{v}_D - K(p^D(\underline{v}_D)) > -\frac{\dot{p}(1-\dot{p})}{\alpha + n\dot{p}(1-\dot{p})} (nN_D + c_D) + \frac{\alpha \dot{p}}{\alpha + n\dot{p}(1-\dot{p})} \underline{v}_D - K(\dot{p})$$

I can then use that $\bar{v}_D - \underline{v}_D > (\bar{v}_D - \underline{v}_D)\left(\frac{\alpha\dot{p}}{\alpha+n\dot{p}(1-\dot{p})}\right)$, which gives

$$\bar{v}_D - K(p^D(\underline{v}_D)) > -\frac{\dot{p}(1-\dot{p})}{\alpha+n\dot{p}(1-\dot{p})}(nN_D + c_D) + \frac{\alpha\dot{p}}{\alpha+n\dot{p}(1-\dot{p})}\bar{v}_D - K(\dot{p}).$$

Thus, high types prefer arming to level $p^D(\underline{v}_D)$ than to fighting.

Case 2. Case 2. Assume $p_0 \geq p^D(\underline{v}_D)$. For all $p \in [p_0, p_1]$, C will not challenge. Thus, in this range, $p = p_0$ dominates all other arming levels.

For C: For $p = max\left\{p^D(\underline{v}_D),\, p_0\right\}$, C believes both types of D are selecting this arming level; thus, C's beliefs on type are the prior distribution of type. For $p \in [p_0, p^D(\underline{v}_D))$ (whenever $p_0 < p^D(\underline{v}_D)$), C believes D is a low-type and would acquiesce if challenged, which gives C a strictly positive payoff for challenging. For $p \in (p^D(\underline{v}_D), p_1]$, C believes that D is a high type and would fight if challenged, which gives C a weakly negative payoff for challenging (because $p^C < p^D(\underline{v}_D)$).

12.11 Pooling Equilibrium 5

This equilibrium occurs when $\underline{v}_D - K(p^C) > 0$ and $\underline{v}_D - K(p^C) \geq U_D(\hat{p}(\underline{v}_D))$, $max\left\{p^D(\underline{v}_D), p_0\right\} < p^C$, $p^C \leq p_1$

- Type $\bar{v}_D$ selects $p = p^C$, and type $\underline{v}_D$ selects $p = p^C$.
- C will not challenge when observing $p = p^C$, will challenge when observing $p < p^C$, and will not challenge when observing $p \geq p^C$.
- Both types would escalate if challenged.
- C's Beliefs: If $p = p^C$, then C believes D is low-type with probability $1 - \pi$ and high-type with probability π. If $p < p^C$, then C believes D is low-type with probability 1. If $p > p^C$, then C believes D is high-type with probability 1.
- Payoffs: Type $\bar{v}_D$ attains $\bar{v}_D - K(p^C)$, type $\underline{v}_D$ attains $\underline{v}_D - K(p^C)$.

Proof of Equilibrium

For type $\underline{v}_D$:

Within the range $p \in [p_0, p^D(\underline{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\underline{v}_D), p^C)$, C will challenge and D will fight. Thus, in this range, $p = \hat{p}(\underline{v}_D)$ weakly dominates all other arming levels. Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. By the conditions of the case, $\underline{v}_D - K(p^C) > 0$ and $\underline{v}_D - K(p^C) \geq U_D(\hat{p}(\underline{v}_D))$, implying that D prefers setting $p = p^C$ and deterring to $p = p_0$ and acquiescing or $\hat{p}(\underline{v}_D)$ and fighting.

For type $\bar{v}_D$:
Within the range $p \in [p_0, p^D(\bar{v}_D))$, C will challenge and D will acquiesce, making D's utility strictly decreasing in p. Thus, in this range, $p = p_0$ dominates all other arming levels. Within the range $p \in [p^D(\bar{v}_D), p^C)$, C will challenge and D will fight. Thus, in this range, $p = \hat{p}(\bar{v}_D)$ weakly dominates all other arming levels. Within the range $p \in [p^C, p_1]$, C will not challenge, making D's utility strictly decreasing in p. Thus, in this range, $p = p^C$ dominates all other arming levels. Because $p^D(\underline{v}_D) < p^C$ and $\underline{v}_D - K(p^C) > 0$, the conditions in Lemma 1 hold; thus, because low-types most prefer setting $p = p^C$ and deterring, high types also most prefer this.

For C: For $p = p^C$, C believes both types of D are selecting this arming level; thus, C's beliefs on type are the prior distribution of type. For $p \in [p_0, p^C)$ C believes D is a low-type and would acquiesce if challenged (when $p < p^D(\underline{v}_D)$) or would fight when challenged (when $p \geq p^D(\underline{v}_D)$); in either case, given $p < p^C$, these give C a strictly positive payoff for challenging. For $p \in (p^C, p_1]$, C believes that D is a high type and would fight if challenged, which gives C a weakly negative payoff for challenging (because $p > p^C$).

13 Demonstrating the Equilibrium Satisfies the Intuitive Criterion

For the Pooling 1, Separating 1, Separating 2, Separating 3, Separating 5, and Separating 6 equilibrium spaces, it is straightforward to see high types are doing as good as they can. For example, in Separating 1, high type D's must arm to level $p = max\,\{p_0, p(\bar{v}_D)\}$ to be willing to fight, and at this level C will not challenge and grant D the asset. If, for example, part of the Separating 1 spaces required D select some $p' > p$ for C to believe D is a high type, then this would not satisfy the intuitive criterion refinement; instead, for all these equilibria, high types D are doing as well as they can in the characterized separating equilibrium (or not separating, in the case of Pooling 1) from low-types.

Pooling 5 also has the feature where high types select the smallest possible value needed to deter C ($p = p^C$). Furthermore, as demonstrated in Lemma 1, high types will always select a weakly greater level of arming than low types; thus the $\underline{v}_D - K(p^C) > 0$ and $\underline{v}_D - K(p^C) \geq U_D(\hat{p}(\underline{v}_D))$ conditions imply that high types will do best selecting p^C over some $p = \hat{p}(\bar{v}_D)$ or $p = p_0$.

It is possible to demonstrate that Pooling 2, Pooling 3, Pooling 4, and Separating 4 all satisfy the intuitive criterion simultaneously. I do this in Lemma 3. To give a sense of what Lemma 3 means, Lemma 3 implies that within Pooling Equilibrium 4, high-type D's will never have an incentive to switch to some p'' where $p^D(\bar{v}_D) \leq p'' < p^D(\underline{v}_D)$ and fight with positive probability relative to arming to $p^D(\underline{v}_D)$ and attaining the asset. Given that Pooling 2-4 and Separating 4 all have the condition where high-types prefer arming to some level $p = p'$ that keeps C from challenging to arming to $p = p_0$ and acquiescing, proving Lemma 3 will imply that the equilibrium above satisfies the intuitive criterion.

Lemma 3: *Suppose an equilibrium exists where C will not challenge upon observing p' where $p = p' \in (p_0, p_1]$, $\underline{v}_D - K(p') \geq 0$, and $p' \leq p^D(\underline{v}_D)$. If for all $p'' \in [p^D(\bar{v}_D), p')$ either (a) C challenges with certainty upon observing p'' or (b) C challenges with probability $1 - \zeta \in (0, 1]$ after observing p'', then high-type D's prefer arming to level p' rather than selecting p'' and fighting with (a) certainty or (b) probability $1 - \zeta$.*

Proof: Any semi-separating equilibrium will take the form of high types arming to level p'' and always fighting when challenged,[49] and low-types mixing between arming to level p_0 and always acquiescing when challenged (where challenging happens with certainty), and arming to level p'' and acquiescing when challenged (where challenging happens with probability $1 - \zeta$).[50] For low-type D's to be indifferent between arming to p_0 and always getting challenged, and arming to p'' and getting challenged with probability $1 - \zeta$, the following must hold (lest ζ does not support a semi-separating equilibrium):

$$0 = \zeta\left(\underline{v}_D\right) + (1-\zeta)\left(0\right) - K(p'').$$

Also note that because $p'' < p^D(\underline{v}_D)$, low-type D's prefer acquiescing to going to war, implying that

$$0 > \zeta\left(\underline{v}_D\right) + (1-\zeta)\left(-\frac{(nN_D + c_D)p''(1-p'')}{\alpha + np''(1-p'')} + \frac{\alpha}{\alpha + np''(1-p'')}\left(p''\underline{v}_D\right)\right) - K(p'').$$

Because $\underline{v}_D - K(p') \geq 0$, I can say

$$\underline{v}_D - K(p') > \zeta\left(\underline{v}_D\right) + (1-\zeta)\left(-\frac{(nN_D + c_D)p''(1-p'')}{\alpha + np''(1-p'')} + \frac{\alpha}{\alpha + np''(1-p'')}\left(p''\underline{v}_D\right)\right) - K(p'')$$

I add $\bar{v}_D - \underline{v}_D$ to the left-hand-side, and I add $\zeta\left(\bar{v}_D - \underline{v}_D\right) + (1-\zeta)\frac{\alpha p''}{\alpha + np''(1-p'')}\left(\bar{v}_D - \underline{v}_D\right)$ to the

[49] High types fight due to $p'' \geq p^D(\bar{v}_D)$

[50] Low types acquiesce because $p'' < p^D(\underline{v}_D)$.

right-hand-side. The inequality is preserved because $\frac{\alpha p''}{\alpha + np''(1-p'')} < 1$. This gives

$$\bar{v}_D - K(p') > \zeta\left(\bar{v}_D\right) + (1-\zeta)\left(-\frac{(nN_D + c_D)p''(1-p'')}{\alpha + np''(1-p'')} + \frac{\alpha}{\alpha + np''(1-p'')}\left(p''\bar{v}_D\right)\right) - K(p'').$$

which implies that high-type D's prefer arming to p' and attaining the asset relative to arming to level p'' and fighting over the asset with some probability (as part of the semi-separating equilibrium).

Note that the proof above also functions for the case when C challenges with certainty (set $\zeta = 0$).

14 Proof of Incomplete Information Remarks

Remarks 1, 2 and 5 still hold in the incomplete information game via construction of $p^D(v_D)$ and p^C. The proof of Remark 3 for the incomplete information game is more complex and is included below.. Remark 4 holds via the proof above. Remark 6 holds via the equilibrium construction. Finally, Remark 7 clearly holds given the signalling equilibrium (discussed above).

15 Proof of Remark 3 (Incomplete Information)

Remark 3 (Nuclear Peace). *Increasing nuclear instability results in fewer instances of war. Formally, we define nuclear instability parameters $n', n'' \in \mathbb{R}_+$ with $n' < n''$. If n' shifts to n'', then the likelihood of war weakly decreases.*

Proof. Because this proof is involved, it is worthwhile outlining how I proceed. I begin by discussing "Case 1," which considers conditions where where war never happens under n''. I then proceed to the more complex case, "Case 2." I first establish a useful lemma, which demonstrates that as n increases, D's utility from war is decreasing. I then establish another useful Lemma, which characterizes the full set of inequalities where low types go to war in equilibrium. For example, one of these inequalities is that low-type D's must do better going to war than acquiescing. I then show the inequalities needed to support the equilibria where low types go to war are strained or break as n increases. Referring back to the example, because low-type D's war utility is decreasing and their "acquiesce" utility is unchanging, the inequality where D prefers fighting to acquiescing is strained or can break. I then repeat the process for high types.

15.1 Case 1: For n'', $p^C \leq p^D(\bar{v}_D)$

If for n'' $p^C \leq p^D(\bar{v}_D)$ holds, then for n'' $p^C \leq p^D(\underline{v}_D)$ also holds.[51] This implies that under n'', war is never possible because there is no arming level where C would be willing to challenge and D would be willing to fight. Therefore, even if n' were such that $p^D(\bar{v}_D) < p^C$ or $p^D(\underline{v}_D) < p^C$ (i.e. war was possible under n'), the likelihood of war would be (weakly) decreasing.

15.2 Case 2: For n'', $p^C > p^D(\bar{v}_D)$

This proof is assisted by a helpful Lemma that applies to a subset of the parameter space within Case 2.

When D is optimally choosing to fight, D selects some arming level p within the set $S = \left[max\left\{p_0, p^D(v_D)\right\}, min\left\{p^C, p_1\right\}\right]$. Intuitively, the set S defines feasible arming levels where D will fight if challenged, and C will not be deterred. Note that we will consider two levels of nuclear instability parameter n, which we denote n and n' (with $n < n'$). As defined, $S(n') \subset S(n)$.[52]

I introduce some new notation here. I let $\hat{U}(p, v_D, n) = -\frac{p(1-p)}{\alpha+np(1-p)}(nN_D+c_D)+\frac{\alpha}{\alpha+np(1-p)}(pv_D)-K(p)$. I also define $p^*(a, b)$ as

$$p^*(a, b) \in argmax_{p\in S(a)}\hat{U}(p, v_D, b)$$

note that whenever $a = b$, this is D optimizing an arming level at nuclear instability parameter n.[53]

Whenever D (optimally) selects a p and goes to war, I define D's value function as

$$\hat{V}_D(n, v_D) = max_{p\in S(n)}\hat{U}(p, v_D, n)$$

This allows us to set up a useful Lemma.

Remark 3 Lemma A: *For a fixed $v_D \in \{\underline{v}_D, \bar{v}_D\}$, $\hat{V}_D(n, v_D)$ is decreasing in n.*

Proof:

[51] Recall that $p^D(\bar{v}_D) < p^D(\underline{v}_D)$.

[52] Recall $p^C = \frac{\alpha v_C}{c_C+nN_C}$ and $p^D(v_D) = 1 - \frac{\alpha v_D}{c_D+nN_D}$.

[53] Note that we abuse notations and sometimes let this denote a set of arming levels; when this is the case, the proof functions for all individual elements of the set $p^*(a, b)$.

With this structure in place, I can show that $\hat{V}(n', v_D) \leq \hat{V}(n, v_D)$. The proof proceeds as follows.

$$\begin{aligned}
\hat{V}(n', v_D) =& max_{p \in S(n')} \hat{U}(p, v_D, n') \\
\leq& max_{p \in S(n)} \hat{U}(p, v_D, n') \\
\leq& \hat{U}(p^*(n, n'), v_D, n) \\
\leq& max_{p \in S(n)} \hat{U}(p, v_D, n) \\
=& \hat{V}(n, v_D)
\end{aligned}$$

The first inequality holds because $S(n') \subset S(n)$; this means that $\hat{U}$ is optimized over a smaller set under n'. The second inequality holds because $\hat{U}(p, v_D, n)$ is decreasing in parameter n at a fixed arming level $p^*(n, n')$.[54] The third inequality holds because there D is selecting their optimal p.$\square$

This Lemma means D does worse from fighting as n increases. To show that this shrinks the parameter set where war occurs, I must analyze the equilibrium cases defined above (Separating 1, Separating 2, etc). I do this in parts, first focusing on showing the low-types will fight less as n increases.

15.2.1 The Parameter Set Where Low Types Fight is Shrinking

I first define the following Lemma:

Remark 3 Lemma B*: If and only if*
(a) when $p^C \leq p_1$*, the following conditions hold:* $max\left\{p^D(\underline{v}_D), p_0\right\} < p^C$, $\underline{v}_D - K(p^C) < U_D(\hat{p}(\underline{v}_D))$ *and* $U_D(\hat{p}(\underline{v}_D)) \geq 0$*, or*

(b) when $p^C > p_1$*, the following conditions hold:* $U_D(\hat{p}(\underline{v}_D)) \geq 0$*,*

then low-type D's go to war.

[54] Taking first order conditions of $\hat{U}(p, v_D, n)$ with respect to n yields $\frac{(p-1)p(\alpha N_D + \alpha p \bar{v}_D - p(1-p)c_D)}{\left(-\alpha + n(p)^2 - sp\right)^2}$. Note that $p - 1 < 0$ and, because $p \geq p^D(v_D)$, we can say $0 \leq -n(1-p)N_D + \alpha\bar{v}_D - c(1-p)$.

Proof: The "iff" relies on how the conditions above are equivalent to the conditions for Separating equilibria 5 and 6, which are the only equilibria spaces where low-types go to war. From earlier (see the proof of Separating 5) I can say that $U_D(\hat{p}(\underline{v}_D)) \geq 0$ implies $\underline{v}_D - K(p^D(\underline{v}_D)) \geq 0$. This means that low-types will fight, and that Lemma 1 can be applied here. Based on Lemma 1, type $\bar{v}_D$ will either select into fighting (setting $p = \hat{p}(\bar{v}_D)$) or deterring (setting $p = p^C$). When $p^C \leq p_1$, $U_D(\hat{p}(\bar{v}_D)) > \bar{v}_D - K(p^C)$, and the above conditions hold, the equilibrium is Separating 5 (a). When $p^C \leq p_1$, $U_D(\hat{p}(\bar{v}_D)) \leq \bar{v}_D - K(p^C)$, and the above conditions hold, the equilibrium is Separating 6. And, when $p^C > p_1$ and the above conditions hold, then this is Separating 5 (b).

□

Based on Remark 3 Lemma B, low-types will only go to war when those constraints hold. From here, I can rely on examining how moving from n' to n'' will alter the constraints. Suppose for n' $p^C \leq p_1$. As n' increases to n'', $p^D(\underline{v}_D)$ is weakly increasing, p_0 is unchanging, and p^C is decreasing, thus making the $max\left\{p^D(\underline{v}_D), p_0\right\} < p^C$ inequality strained (or potentially break). Also as n increases, $\underline{v}_D - K(p^C)$ is increasing , $U_D(\hat{p}(\underline{v}_D))$ is decreasing (as shown above in the Nuclear Instability and War Lemma), and 0 is unchanging, thus making the inequalities $\underline{v}_D - K(p^C) < U_D(\hat{p}(\underline{v}_D))$ and $U_D(\hat{p}(\underline{v}_D)) \geq 0$ strained (or potentially break). Now suppose for n' $p^C > p_1$ holds; through the logic discussed above, the inequalities in this case are strained or could break. Because p^C is decreasing in n, the shift from $n'$to n'' could result in a move from (abusing notation) $p^C(n') > p_1$ to $p^C(n'') \leq p_1$. When this shift occurs, for war to still occur, the additional constraint $\underline{v}_D - K(p^C) < U_D(\hat{p}(\underline{v}_D))$ must also hold; thus, in the shift from n' to n'', all existing constraints become more difficult to satisfy and new constraints must be met, collectively making low-type D's less willing to go to war.

15.2.2 The Parameter Set Where High Types Fight is Shrinking.

As it was for low types, I must identify the constraints that fully characterize all the parameter spaces where high-types go to war. I do this in the following Lemma:

Remark 3, Lemma C: If and only if

(a) When $p^C \leq p_1$, the following conditions hold: $max\left\{p^D(\bar{v}_D), p_0\right\} < p^C$, $\bar{v}_D - K(p^C) < U_D(\hat{p}(\bar{v}_D))$ and $U_D(\hat{p}(\bar{v}_D)) \geq 0$, or

(b) when $p^C > p_1$, the following conditions hold: $U_D(\hat{p}(\bar{v}_D)) \geq 0$,

then high-type D's go to war.

Proof:

First, suppose $p^C \leq p_1$. It could also be that
(0) The set $[max\left\{p^D(\underline{v}_D), p_0\right\}, p^C]$ is empty
(1) The set $[max\left\{p^D(\underline{v}_D), p_0\right\}, p^C]$ is non-empty and $U_D(\hat{p}(\underline{v}_D)) \geq 0$; or
(2) The set $[max\left\{p^D(\underline{v}_D), p_0\right\}, p^C]$ is non-empty and $U_D(\hat{p}(\underline{v}_D)) < 0$.

Writing the conditions in (a) with the conditions in (0) and (2) (in other words, fully writing out conditions (0) and (2)) gives:

(0). $p^C \leq p_1$, $max\left\{p^D(\bar{v}_D), p_0\right\} < p^C$, $\bar{v}_D - K(p^C) < U_D(\hat{p}(\bar{v}_D))$ and $U_D(\hat{p}(\bar{v}_D)) \geq 0$ and the set $[max\left\{p^D(\underline{v}_D), p_0\right\}, p^C]$ is empty.

(2). $p^C \leq p_1$, $max\left\{p^D(\bar{v}_D), p_0\right\} < p^C$, $\bar{v}_D - K(p^C) < U_D(\hat{p}(\bar{v}_D))$ and $U_D(\hat{p}(\bar{v}_D)) \geq 0$ and the set $[max\left\{p^D(\underline{v}_D), p_0\right\}, p^C]$ is non-empty and $U_D(\hat{p}(\underline{v}_D)) < 0$.

Together, these are equivalent to Separating 3 (a).

The full set of conditions in (1) are the following: $p^C \leq p_1$, $max\left\{p^D(\bar{v}_D), p_0\right\} < p^C$, $U_D(\hat{p}(\bar{v}_D)) > \bar{v}_D - K(p^C)$, $U_D(\hat{p}(\bar{v}_D)) \geq 0$, the set $[max\left\{p^D(\underline{v}_D), p_0\right\}, p^C]$ is non-empty and $U_D(\hat{p}(\underline{v}_D)) \geq 0$.

These conditions are nearly equivalent to what is stated in Separating 5 (a). At first pass there appears to be two differences, but, as I show below, these difference are effectively ruled out.

First, the conditions for Separating 5 (a) states $max\left\{p^D(\underline{v}_D), p_0\right\} < p^C$, while the conditions on the set in (1) being non-empty imply $max\left\{p^D(\underline{v}_D), p_0\right\} \leq p^C$. In other words, (1) above states it is possible for $max\left\{p^D(\underline{v}_D), p_0\right\} = p^C$, the while Separating 5 (a) conditions do not state this is possible. However, note that the other conditions in (1) imply that this equality can never actually hold. If for high types $max\left\{p^D(\bar{v}_D), p_0\right\} < p^C$, it must be that $p^C > p_0$. Due to this, the remaining distinction between (1) and the conditions in Separating 5 (a) is that (1) also allows for $p^D(\underline{v}_D) = p^C$. However, it cannot ever be the case that $p^D(\underline{v}_D) = p^C$ and $U_D(\hat{p}(\underline{v}_D)) \geq 0$ simultaneously hold when $p_0 < p^C$. Based on how $p^D(\underline{v}_D)$ is defined, the following holds:

$$-\frac{p^D(\underline{v}_D)(1-p^D(\underline{v}_D))}{\alpha + np^D(\underline{v}_D)(1-p^D(\underline{v}_D))}(nN_D + c_D) + \frac{\alpha p^D(\underline{v}_D)}{\alpha + np^D(\underline{v}_D)(1-p^D(\underline{v}_D))}\underline{v}_D = 0.$$

Additionally, because $\hat{p}(\underline{v}_D)$ must fall between $p^D(\underline{v}_D)$ and p^C, when $p^D(\underline{v}_D) = p^C$, it must also be that $\hat{p}(\underline{v}_D) = p^D(\underline{v}_D) = p^C$. Expanding out the expression $U_D(\hat{p}(\underline{v}_D)) \geq 0$ and comparing it to the expression above (note that $U_D(\hat{p}(\underline{v}_D))$ has an additional cost term) gives

$$-\frac{p^D(\underline{v}_D)(1-p^D(\underline{v}_D))}{\alpha + np^D(\underline{v}_D)(1-p^D(\underline{v}_D))}(nN_D + c_D) + \frac{\alpha p^D(\underline{v}_D)}{\alpha + np^D(\underline{v}_D)(1-p^D(\underline{v}_D))}\underline{v}_D - K(p^C) \geq 0.$$

This cannot ever hold: if the top expression equals zero and the bottom expression has a new subtracted cost, then it cannot simultaneously be the case the $\hat{p}(\underline{v}_D) = p^D(\underline{v}_D) = p^C$ and $p_0 < p^C$.

Second, the conditions in Separating 5 (a) does not state that $U_D(\hat{p}(\bar{v}_D)) \geq 0$. However, because $U_D(\hat{p}(\underline{v}_D)) \geq 0$ (which is given in (1)), based on the proof of Separating 5, it implies that $\underline{v}_D - K(p^D(\underline{v}_D)) \geq 0$, meaning Lemma 1 can apply and I know that high-type D's will select a greater arming level. Additionally, $U_D(\hat{p}(\underline{v}_D)) \geq 0$ implies that low-type D's will either fight (set $p = \hat{p}(\underline{v}_D)$) or deter C (set $p = p^C$). Additionally, I know that high-type D's will not set $p = p^C$ due to $U_D(\hat{p}(\bar{v}_D)) > \bar{v}_D - K(p^C)$. Together, this implies that both types of D will fight, meaning $U_D(\hat{p}(\underline{v}_D)) \geq 0$. Thus, the conditions set out in (1) are equivalent to the conditions in Separating 5(a).

Now suppose $p^C > p_1$. It could also be that
(1) $U_D(\hat{p}(\underline{v}_D)) \geq 0$; or
(2) $U_D(\hat{p}(\underline{v}_D)) < 0$.

Writing out conditions (0) and (2) in full gives

(2) $p^C > p_1$, $U_D(\hat{p}(\bar{v}_D)) \geq 0$, and $U_D(\hat{p}(\underline{v}_D)) < 0$.

together, these are the conditions for Separating 3 (b).

Writing out conditions (1) in full gives

(1) $p^C > p_1$, $U_D(\hat{p}(\bar{v}_D)) \geq 0$, and $U_D(\hat{p}(\underline{v}_D)) \geq 0$.

These are the conditions for Separating 5 (b).

I have demonstrated that the conditions in the above Lemma are equivalent to the conditions for Separating 3, Separating 5 (a), and Separating 5 (b), the three settings where high type D's fight.

□

From here, I can examine how moving from n' to n'' will alter the constraints. Suppose for both n' and n'' $p^C \leq p_1$. As n increases, $p^D(\bar{v}_D)$ is weakly increasing, p_0 is unchanging, and p^C is decreasing, thus making the $max\left\{p^D(\bar{v}_D), p_0\right\} < p^C$ inequality strained (or potentially break). Also as n increases, $\bar{v}_D - K(p^C)$ is increasing , $U_D(\hat{p}(\bar{v}_D))$ is decreasing (as shown above), and 0 is unchanging, thus making the inequalities $\bar{v}_D - K(p^C) < U_D(\hat{p}(\bar{v}_D))$ and $U_D(\hat{p}(\bar{v}_D)) \geq 0$ strained (or potentially break). Now suppose for both $n'$and n'' $p^C > p_1$ holds; through the logic discussed above, $U_D(\hat{p}(\bar{v}_D)) \geq 0$ is strained or could break. Because p^C is decreasing in n, the shift from $n'$to n'' could result in a move from $p^C(n') > p_1$ to $p^C(n'') \leq p_1$. When this shift occurs, it imposes additional constraints $max\left\{p^D(\bar{v}_D), p_0\right\} < p^C$, $\bar{v}_D - K(p^C) < U_D(\hat{p}(\bar{v}_D))$ for fighting to still occur; thus, all existing constraints become more difficult to satisfy and new constraints must be met, collectively shrinking the set over which high-type D's go to war.

I have now demonstrated that as n increases, the constraints that result in selection into Separating 3 and Separating 5 all become more difficult to satisfy. Thus, in the shift from n' to n'', the war outcome occurs over a smaller set, or disappears altogether.

Part VI

Discussion: What if C Could Also Arm?

It would be possible to modify the game in the text to include C undertaking some (potentially costly) arming. Suppose, like Powell (2015), I assumed that C arms at the beginning of the game. For ease, suppose C chooses one of two arming levels that I denote $a_C \in \{\underline{a}_C, \bar{a_C}\}$, with $\underline{a}_C < \bar{a}_C$. Note that C's arming decision would need to factor into D's utility function; this could be accomplished by manipulating D's arming cost function, making it $K(p; a_C)$, where, for all p, $K(p, \underline{a}_C) < K(p, \bar{a}_C)$. Also note that I do not need to do anything to interact C's and D's arming level in the generation of nuclear risk; ultimately, D's selected p will still determine risk levels with more force parity leading to longer conflicts. However, also note that by modeling it in this way, p no longer cleanly represents the absolute arming level or force posture put forward by D; this will be important to the discussion below.[55]

[55]There is some subtlety here. Suppose under both $a_C = \underline{a}_C$ and $a_C = \bar{a_C}$, D optimally arms to $p^* = \frac{1}{2}$, meaning, in response to either of C's arming levels, D has armed to the level where D has a 50-50 chance of winning in a conventional war. However, in absolute terms, D would need to select a higher level of arming

Behavior in the game after C's arming decision would be identical to what is written in Proposition 1, only with $K(p; a_C)$ instead of $K(p)$. Regarding C's arming decision, some subtlety can arise. If, in the game, D would be willing to deter C even if C selected $\bar{a}_C$, then C has no reason to ever select the higher (more costly) arming level, and C will always select $\underline{a}_C$. Otherwise, depending on how costly arming is for C, sometimes C will select the low or high arming level (based on C's utilities), and actions will play out as it was in Proposition 1.

However, this is not to say that comparative statics (or the Remarks) would be the same.

Consider Remark 1, Suppose there was a setting where $p^C < p^D$ and C selected $a_C = \bar{a}_C$ in equilibrium, which resulted in D acquiescing. Now suppose n increases. If this is the case, then p^D would rise, which could potentially result in C initially setting $a_C = \underline{a}_C$ with D still acquiescing. While Remark 1 in the text claims that, when $p^D > p^C$, increasing n requires D to put forward a more expansive force posture for deterrence, in the modified model with this described equilibrium behavior, there are offsetting effects: p^D is increasing in n, but a_C is decreasing. If the latter effect is larger, then the absolute force posture that delivers a p^D probability of winning a war could be less in absolute terms under the new n. These competing effects cannot easily be disentangled, but if changes in n result in dramatic changes in C's arming levels, this is where the Remarks are most likely to being contradicted.

Essentially, the model with C arming could produce a range of new comparative statics; however, it is not as if these new comparative statics would eliminate the possibility of what is presented in the text.

References

Baliga, S., Bueno de Mesquita, E. and Wolitzky, A. (2020) Deterrence with imperfect attribution, *American Political Science Review*, **114**, 1155–1178. IV

Carter, D. B. (2010) The strategy of territorial conflict, *American Journal of Political Science*, **54**, 969–987. IV

Powell, R. (2015) Nuclear brinkmanship, limited war, and military power, *International Organization*, **69**, 589–626. (document), 1, IV

Schelling, T. C. (1966) *Arms and influence*, Yale University Press. 1

to achieve $p^* = \frac{1}{2}$ when C set $a_C = \bar{a_C}$ relative to when $a_C = \underline{a}_C$. This higher arming level would come with a higher cost, which is being captured in K. However, because the model does not assume that the level of nuclear risk in these two settings are different—because force parity or force imbalance drives nuclear risk—I do not need to modify nuclear instability or hazard rates.