When Can States Signal with Sunk Costs?

Noam Reich[*†]

January 13th, 2026

Abstract

I argue that states' ability to signal with sunk costs depends on the type of private information they have. The problem is that weak and unresolved states are the most incentivized to obtain a peaceful concession that avoids fighting and will mimic any signal that a strong or resolved state would be willing to send. Using a formal model, I show that communication breaks down unless strong or resolved states enjoy a signaling advantage that reduces their signaling costs relative to lower-quality types. I identify two common forms of signaling advantage – differential costs and index signals – that allow for communication of strength or resolve with sunk costs. However, these additional requirements for sunk-cost signaling are not guaranteed in practice and suggest that communication of resolve is particularly difficult. An additional extension shows that these incentives persist when sunk costs constitute arming and have a hand-tying effect.

Word Count: 10,416

[*]I would like to thank Alexandre Debs, Matias Iaryczower, Kathy Ingram, James Morrow, and Kris Ramsay for their helpful discussions and suggestions. Additional thanks to audiences and participants at the University of Minnesota's IR Colloquium, Vanderbilt's Conflict Workshop, Formal Models of Conflict Conference 2024, ISA 2023, Eurasian Peace Science 2023, APSA 2022, MPSA 2022 and 2025, and SPSA 2021 who provided feedback on the project.

[†]Yale University. Assistant Professor. Email: noam.reich@yale.edu

Costly signaling is the primary means by which states are believed to communicate their willingness or ability to fight in international crises. According to costly signaling theory, high-quality challengers can distinguish themselves by taking costly actions that they would not take if they were unwilling to fight. It might seem natural to assume that high-quality challengers should be willing to invest a great deal in costly signals because they are those who would be most likely to obtain a peaceful concession if they could credibly communicate their type. Indeed, in Fearon's (1997) seminal paper on the costly signaling literature, he demonstrated that states could successfully demonstrate their willingness to fight with sunk costs.

However, I argue that existing work on sunk cost signaling relies on an often ignored assumption regarding the *type* of private information states have Arena 2013. The canonical winner-takes-all costly lottery represents the payoff to fighting with three parameters: (1) v or the country's *valuation* for the good or issue under dispute; (2) s the country's *strength* that determines its probability of winning the lottery and obtaining the good; and (3) c a *cost of fighting*, henceforth referred to as *resolve*, that must be paid regardless of the conflict outcome.[1] That work which has showed that sunk costly signaling is viable has focused on the case where states have private information regarding their *valuation* of the good or issue under dispute (Fearon 1997; Slantchev 2005).

This is consequential because *valuation* is the only type of uncertainty where high-quality states would be willing to pay sunk costs and low-quality states would not. This is because high valuation states are more willing to pay to be sure that they obtain the good under dispute. By contrast, states with low valuations would not bother to invest in obtaining the good peacefully or through war. Conversely, when states have private information regarding their *strength* or *resolve*, it is lower-quality types who are most willing to pay sunk costs because they are willing to go greater lengths to increase their probability of getting the good without having to resort to war. By contrast, stronger or more resolved types have less

[1]The words *valuation*, *strength*, and *cost of fighting* or *resolve* will be italicized whenever they represent variable names.

to lose from fighting and are less willing to pay to avoid it.

Does this mean that states cannot convey *strength* or *resolve* with sunk cost signals? If so, under what conditions is signaling possible? To answer these questions, I re-analyze Fearon's (1997) classical signaling model with different types of uncertainty. In particular I adopt a "budgeting" approach which asks how a states' willingness to pay sunk costs to "purchase" an increase in the probability that their rival concedes depends on their private information. Signaling requires that higher-quality types who the receiver would like to concede to be willing to outspend their lower-quality counterparts that the receiver would be willing to fight.

My analysis produces three key results. First, I demonstrate that states' budget for signaling with sunk costs will be strictly in their *valuation*, but strictly decreasing in their *strength* and *resolve*.[2] Second, I show that this implies that states can only use sunk costs to convey *strength* or *resolve* under a more restrictive set of assumptions than those required to communicate *valuation*. Specifically, strong and resolved states require a signaling advantage that allows them to circumvent their budget problems and signal without being mimicked by lower-quality types.

I identify two common forms of signaling advantage: differential costs and index signals. The former occurs when strong or resolved states require less effort to produce a signal, effectively receiving a discount. In this case signals are informative because they are a testament to the sender's capacity to produce signals effectively or their ability to shoulder the burden of signaling. Index signals refers to the case where strong types have access to unique technologies that enable them to issue signals but weaker types do not. Originally developed in theoretical biology, such "index signals" are informative because the capacity to send the signal reveals something about the technology itself or is a broader indicator of the signaler's capacity (Maynard-Smith et al. 2003). Such signals derive are informative because weak types do not have the technology to imitate them, not because they require a

[2]Formally, a state's budget for signaling will increase with its cost of fighting.

great deal of effort to produce.

The third set of key results extend the logic of my previous arguments to the case where sunk costs constitute arming, increasing countries strength and having an additional hand-tying effect (Slantchev 2005). In an extension, I show that this additional benefit to sunk cost investments does not fundamentally change state's underlying incentives - as a country becomes stronger or more resolved, its willingness to devote effort to obtaining a peaceful concession still decreases. However, these budgetary constraints no longer prevents signaling. Because the sender benefits from the boost in strength associated with paying sunk costs, they can benefit from doing so even as arming reveals that the sender may have been *initially* weak or unresolved. The weaker or more unresolved a state is ex-ante, the more they benefit from arming and the larger their budget for signaling. The receiver in turn is more likely to concede upon observing higher levels of armament. By contrast the strongest and most resolved types of the sender do not arm at all. Upon observing a sender who does not arm, the receiver can conclude they are facing a type who is confident in their abilities and is willing to fight a war unarmed. Consequently, they still concede with positive probability though a smaller probability than they do versus armed types.

In conjunction, these results have several important implications. First, is that signaling theory is not as robust as is often believed. Specifically, states' ability to communicate their private information using sunk costs may depend on the type of private information they possess. This provides theoretical support for an empirical literature that has become somewhat skeptical of the efficacy of sunk cost signaling (A. Post 2019; Altman and Quek 2024; A. S. Post and Sechser 2024). Second, is that despite this skepticism, the use of sunk cost signaling to communicate *strength* or *resolve* is still possible under restrictive assumptions. A discussion that draws on historical illustrations demonstrates that (1) that the signaling advantages enjoyed by these higher-quality types can sometimes cause the cost of these signals to be quite small; (2) that it is difficult to assess systematically how prevalent signaling advantages are since they rely on unobservable assumptions relating to a country's cost function for signaling; (3) that the assumptions are more amenable to the strategic communication of *strength* via sunk cost signals than *resolve*.

Subsequently, the third and most important implication is that the results suggest that of three types of private information a state could have, *resolve* is the most difficult to communicate and subsequently the most detrimental to peace. States have the easiest time communicating private information regarding their *valuation* - the assumptions required for signaling with sunk costs are the weakest and it has been shown elsewhere that states may even be able to communicate *valuation* through cheap talk (Joseph 2021). It follows that if signaling advantages are more prevalent for communicating strength, then information about *resolve* is the most difficult to convey and consequently the most likely type of private information to cause war. This result goes against the current conventional wisdom established by Fey and Ramsay (2011) that private information regarding strength is the most detrimental to peace. However, their results only consider an environment in which states can convey information by bargaining or through costless messages (cheap talk).[3] By contrast, my results contribute to a growing literature suggesting that states may have more avenues to strategically communicate strength via costly signals (Montgomery 2020; Reich 2022).

The results in this paper builds upon two closely related papers. Arena (2013) was the first paper to identify that sunk cost signaling was sensitive to the type of private information the signaler had. Assuming a binary type space, Arena showed focused on the case where sunk costs constituted arming and showed that communication failed when a signaler had private information regarding their "martial effectiveness," a component of a contest function.[4] I generalize these results to a continuous type space and characterize conditions under which communication via sunk cost signaling is possible anyway. Carroll and Pond (2021)

[3]Fey and Ramsay (2011) use a mechanism design approach to show that when countries can send costless messages and bargain, then in any crisis bargaining game there always exists an equilibrium that always leads to peace. By contrast if countries have private information regarding their strength, then there exists sections of the parameter space in which war must occur with positive probability in any equilibrium.

[4]Formally, Arena (2013) modeled martial effectiveness as a force multiplier e on the amount that the signaler armed $\frac{e*m}{e+m+c}$ where c is the challenger's strength and m is the amount the signaler armed

present a similar signaling model in which an autocrat has private information regarding their strength and must decide whether to issue a show of force (i.e. sink costs) or make transfers in an attempt to deter civilians from rebellion. They show that shows of force can function as signals when the ruler has differential costs in a model with two types. One can therefore consider results 5 and 6 below as an extension of these results to a continuous type space and more generalized signaling function.

The paper proceeds as follows. First, I define signaling theory, review the relevant literature, and situate my arguments within it. Second, I present Fearon's (1997) model which I replicate with three different types of private information. Third, I introduce a "budgeting" approach to signaling, demonstrating that as a country's *valuation* and *cost of fighting* (strength) increase (decreases) its willingness to spend on sunk cost signaling decreases. I then show that this implies that sunk cost signaling is not viable when states have private information regarding *strength* or *resolve*. Fourth, I introduce signaling advantages and formally demonstrate the conditions under which they work, providing historical examples throughout. Fifth, I demonstrate that when sunk costs are an investment in arming, it is the lowest-quality types that arm. I conclude with a discussion on the implications of the results.

Signaling Theory and its Critics

What is costly signaling? Gartzke et al. (2017, p. 2) define signaling as "the purposive and strategic revealing of information about intent, resolve, and/or capabilities by an actor A to alter the decisions of another actor B to improve the chances that an outcome desired by A is reached when the desired outcomes of A and B are dissimilar." Morrow (1997, p. 87) offers a similar definition but adds that this requires that types separate, i.e., when states with different underlying types take different actions thereby allowing an observer to infer the type of state that they are facing.

However, when considering signaling with audience costs, sunk costs, diplomatic offers or other signals in a crisis setting most theoretical papers or empirical papers typically seem to work on a more precise definition (Fearon 1994; Fearon 1997; Reich 2022). First, is that higher-quality challengers will produce larger signals, i.e. spend more on sunk costs. In the case of sunk cost signaling, that would imply that to that states that as a state's valuation, strength, or resolve increases, then the amount of the costly action it engages in increases. Second, that the receiver of the signal infers that a higher action is more likely to have been sent by a higher type. Third, is that, in response to these beliefs, more costly signaling actions induce higher rates of concession from the receiver. Thus costly signaling theory is built around a comparative static prediction wherein higher-quality challengers should be associated with larger signals and concessions, and vice-versa.

Sunk cost signals have formed a core of costly signaling theory since its inception. In Fearon's (1997) original conceptualization of sunk costs, they were defined by two key properties defined sunk cost signals:(1) that they were irrevocable once spent and (2) that they did not alter the sender's payoffs in any direct way, only benefiting the sender by altering the receiver's belief that the sender is willing to fight. Military demonstrations, such as military exercises or the mobilization of troops in a crisis, are the paradigmatic example of the first requirement - regardless of the outcome of the dispute, any resources spent on these demonstrations cannot be recovered. Slantchev (2005) introduced an important modification and argued that the sunk cost signals often failed the second requirement since military mobilizations and demonstrations are irrevocable outlays that are often designed to strengthen the sender and improve their probability of winning a potential war. This argument has received empirical support (A. Post 2019; Altman and Quek 2024; A. S. Post and Sechser 2024). However, regardless of whether one conceives as sunk cost signals as purely "burnt money" or as "arming" the comparative static predictions and logic at the heart of signaling theory remained the same - higher-quality senders both burnt more money and armed more.

Since its introduction to the study of international crises, scholars have pointed to several problems with signaling theory. First, a number of papers expand the strategic environment by introducing additional actors or actions that produce novel and important trade-offs that cut against state's incentive to signal. For example, Slantchev (2010) demonstrates that it can be beneficial for states to hide their strength by making lower demands so as to lull their rival into a false sense of security that leads the rival to under-invest in arming. Similarly, Kurizaki (2007) argues that countries may refrain from producing audience costs if a public crisis would cause their rival to accumulate audience costs as well that would subsequently prevent them from conceding. Wolford (2014) studies a model with multiple audiences in which a state mobilizing against an enemy may make it more difficult for an ally to come to its aid. Though these considerations complicate the signaling story, they do not unravel the core comparative static at the heart of costly signaling theory - across these papers higher-quality challengers are still more likely to take the more "aggressive" action and adopt larger signals. Second, some papers challenge costly signaling theory by identifying features of the strategic environment that are necessary for sunk cost signaling that may be taken for granted. For example, Wolton (2024) combines sunk cost signaling and bargaining in a single model and demonstrates that signaling can break down if the sender does not also have proposal power because the proposer simply offers the sender their expected utility for fighting.

The argument that I advance in this paper is a more fundamental problem for costly signaling theory because it challenges its core comparative static predictions. Following Arena (2013), I show that merely switching the type of private information a states possess can cause signaling to unravel. This diminishes the theory's robustness. Additional assumptions can both restore signaling and the comparative static wherein stronger states signal more and are more likely to receive concessions, but the universality of the theory is reduced to cases where these assumptions hold (Carroll and Pond 2021).

If the assumptions do not hold, then one of two situations is possible. First, if sunk cost signals constitute arming, then states can employ sunk cost signals but the comparative statics of traditional signaling theory won't hold - strong or resolved states will pay fewer sunk costs, arm themselves less, and be less likely to obtain a concession. Alternatively, if the assumptions necessary for costly signaling theory do not hold, then other theories may provide a better understanding of crisis behavior. For example, Reich (2025) presents a theory of dynamic screening in which sunk costs are an investment in diplomacy - a cost that must be paid while states delay war to try and settle a dispute peacefully. Consequently, sunk costs are associated with with an increased probability of obtaining a concession but also, eventually, with less resolve. In either scenario, though more sunk costs are associated with a higher probability of concession, they are not associated with a stronger or more resolved signaler.

The Signaling Game

The basic model analyzed here is identical to that in Fearon (1997). There are two countries who are vying for control of an indivisible good. The Defender can try to deter the Challenger from initiating a challenge by paying a sunk-cost signal m. The Challenger observes the signal and must then decide whether it will challenge the Defender for the good or not. If the Challenger decides to issue a challenge, then the Defender must decide whether to concede the good to the Challenger or to fight. Figure 1 depicts the game tree.

War is modeled as a costly lottery where the winner receives the good (Fearon 1995). The Defender's probability of winning a war is given by $p(s)$ and is strictly increasing in their strength s. The Challenger has the reciprocal probability of winning $1 - p(s)$. Finally, if the countries go to war they each pay a cost for fighting c_C, c_D respectively. The Challenger and Defender each value the good at v_C and v_D respectively so that each country's expected utility for fighting is given by $v_C(1 - p(s)) - c_C$ and $v_D p(s) - c_D$.

I assume that the Defender has private information regarding one of the three parameters that constitute it's payoff to fighting. Throughout the analysis, I will vary which parameter is the Defender's private information. Formally, I let θ denote the parameter over which the Defender has private information in the set $\Theta = \{v_D, s, c_D\}$ and assume that the other parameters Θ/θ are common knowledge. Regardless, of θ's value, I assume that θ is a random

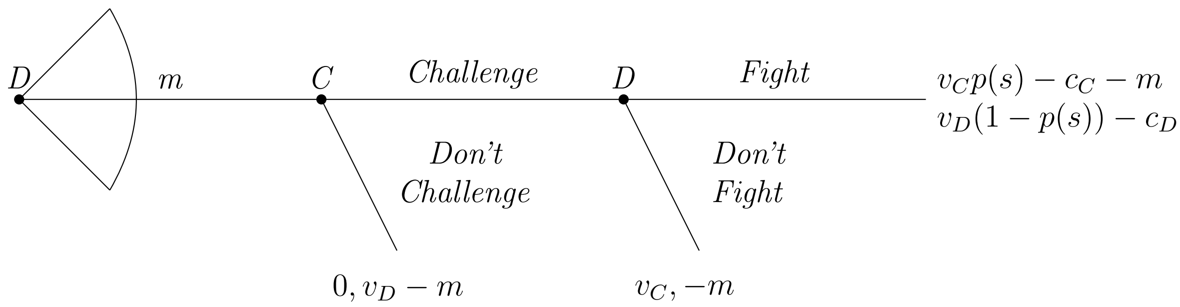

Figure 1: The Game Tree

variable chosen by nature that is distributed according to a common knowledge, continuous, and strictly increasing cumulative distribution function F over support $\theta \in [\underline{\theta}, \bar{\theta}]$. A strategy for the Defender is therefore a mapping from its type θ to a signal $m \in \mathbb{R}^+$ and, conditional on a challenge a binary decision whether to fight or not $\omega \in \{0, 1\}$. Formally, the Defender's strategy is given by $\sigma : \theta \to \mathbb{R}^+ \times \{0, 1\}$. The Challenger's strategy is a mapping from the Defender's observed signal m to a probability with which it will stand firm in response, $\psi : m \to [0, 1]$.

Throughout the paper, I will use Perfect Bayesian Equilibrium as a solution concept. This requires that each type of the Defender first select an optimal strategy given their type, and second, conditional on the Challenger issuing a challenge, decide whether to fight or not so as to maximize its expected utility. Conversely, it requires that the Challenger update its beliefs regarding the Defender's type using Bayes' Rule whenever possible and then select an optimal strategy conditional on those beliefs. I will use the function $G(\theta|m)$ to denote the Defender's posterior beliefs. A Perfect Bayesian Equilibrium will consist of a triple (σ, ψ, G) that satisfies the specified criteria.

Convenience Assumptions

To keep matters interesting, I will assume that the Challenger won't be deterred from a challenge absent any communication, but that there are types of the Defender who could deter the Challenger from issuing a challenge if they could successfully communicate their private information. If the Defender has private information regarding their *valuation* or *resolve* , then this requires that I assume that the Challenger has a negative expected utility from fighting

$$v_C(1 - p(s)) - c_C < 0,$$

so that the Challenger would only be willing to challenge if it expected the Defender to concede. In addition, I assume that some types of the Defender have a positive expected utility to fighting and would fight if challenged while some have a negative expected utility to fighting and will concede in response to a challenge. Let $\tilde{\theta}$ denote the type of Defender that is indifferent between fighting and conceding. In the case where $\theta = v_c$, to ensure that the Challenger would be willing to challenge in equilibrium without any communication, I assume that

$$F(\tilde{v}_D)v_c + (1 - F(\tilde{v}_D))(v_C(1 - p(s)) - c_C) > 0. \quad (1)$$

When the $\theta = c_C$, I assume the following similar assumption holds

$$[1 - F(\tilde{c}_D)]v_c + F(\tilde{c}_D)(v_C(1 - p(s)) - c_C) > 0 \quad (2)$$

Each of these assumptions imply that the Challenger's prior belief that the Defender will concede is large enough to make issuing a challenge worthwhile.

Similarly, I assume that the Challenger won't be deterred from issuing a challenge absent communication whenever the Defender has private information regarding their *strength*. In this case let $\tilde{s}$ denote the type of Defender who is indifferent between fighting the Challenger and not.[5] To ensure that the Challenger would be willing to challenge in equilibrium without

[5]Note that it is possible that $\underline{s} > \tilde{s}$. However, it is without loss of generality and convenient to assume that $\tilde{s} > \underline{s}$.

any communication, I assume that

$$F(\tilde{s})v_C + \int^{\bar{s}_{\tilde{s}}[v_C(1-p(s))-c_C]f(s)ds>0}(3)$$

Let ρ denote the type of the Defender such that if the Challenger had posterior beliefs

$$G(s) = \begin{cases} 0 & \text{if } s < \rho \\ \frac{f(s)}{1-F(\rho)} & \text{otherwise} \end{cases}$$

it would be indifferent between challenging and not. Note that that ρ can either be defined as the value of s for which the following holds

$$\int_{\rho}^{\bar{s}} v_C(1-p(s))f(s)ds - C_C = 0 \tag{4}$$

or the value of s for which the following holds

$$[F(\tilde{s}) - F(\rho)]v_C + \int_{\tilde{s}}^{\bar{s}} v_C(1-p(s))f(s)ds - C_C = 0 \tag{5}$$

In the former instance the Challenger has a positive expected utility for fighting a sufficiently large set $[\tilde{s}, \bar{s}]$, such that it is willing to risk a challenge even if it were certain that the Defender were going to fight. In this case deterrence requires the strongest types $s \in [\rho, \bar{s}]$ to distinguish themselves via a costly signal from weaker types who were also willing to fight $[\tilde{s}, \rho]$. In the latter case, the Challenger is deterred by a sufficiently large set of types of the Defender who have a positive payoff for fighting that for it to be indifferent between challenging and not, it must be sure that it is also challenging a sufficient set of types $s \in [\rho, \tilde{s}]$ who will concede if challenged. Without loss of generality and for ease of exposition, I will assume that equation (4) defines ρ throughout.

Baseline Results

In this section I demonstrate that sunk cost signaling is viable when the Defender has private information regarding their *valuation*, but not when they have private information regarding their *strength* or *resolve*. To do so, I use a "budgeting" approach that asks how much a state will be willing to pay to achieve a specific probability of obtaining a peaceful concession when the alternative is being challenged. My results show that as a country's *valuation* increases, its willing to pay for a concession increases because it cares more about the issue at stake. This makes sunk cost signaling possible. Conversely, as a country's *strength* or *resolve* increases, its willingness to pay for a concession decreases because outside option improves while its valuation for a concession stays the same. As a result, lower-quality types will always be willing to "match" any investment in a sunk cost signal that a high-quality type would be willing to make, thereby preventing communication via sunk cost signals.

A simple example illustrates the benefits of the budgeting approach. Suppose that the Defender had private information regarding their *resolve* ($\theta = c_D$). Conjecture that there were an equilibrium in which one of two signals were sent: $\hat{m} > 0$ in response to which the Challenger did not challenge, and $m = 0$ in response to which the Challenger issued a challenge. A resolved type of the Defender ($c_D \leq \widetilde{c}_D$) would only be willing to send signal $\hat{m}$ so long as it wasn't more costly than fighting

$$v_D - \hat{m} \geq v_D p(s) - c_D$$

It follows that the most that a resolved type of the Defender would be willing to spend on the signal $\hat{m}$ is

$$\hat{m}(c_D) := v_D - v_D p(s) + c_D$$

The right-hand side of the equation makes it clear that the maximum amount that the

Defender is willing to invest on sunk cost signaling is increasing in their cost of fighting c_D. For unresolved types of the Defender who would prefer to concede than fight ($c_D > \tilde{c}_D$), mimicking such a signal is worthwhile whenever

$$v_D - \hat{m} > 0$$

which must hold if $\hat{m}(c_D)$ is positive for any type $c_D < \tilde{c}_D$. This implies that any signal that a resolved type of the Defender ($c_D < \tilde{c}_D$) would be willing to send that obtains a concession with certainty, an unresolved type of the Defender ($c_D > \tilde{c}_D$) would be willing to send as well. It follows that there the conjectured strategies cannot form an equilibrium - the Challenger's posterior belief in response to a signal $\hat{m}$ must always include all types $c_D \in [\underline{c}_D, \tilde{c}_D]$ and per the assumption in equation (2) they must challenge in response.

I now extend the above example and provide a formal definition for the Defender's budget for signaling. Let $B(m|\theta)$ denote the maximum amount a Defender of type θ is willing to spend on a given sunk cost signal m when the alternative is being challenged. The Defender's budget for sunk cost signaling $B(m|\theta)$ is defined as the difference between a type's expected utility for issuing sunk cost signal m and their expected payoff from best responding to a challenge (either fighting or conceding outright).[6]

$$B(m|\theta) := v_D\psi(m) + (1 - \psi(m)) \max\{0, [v_D p(s) - c_D]\} - \max\{0, v_D p(s) - c_D\} \quad (6)$$

If the Defender has private information $\theta \in \{v_D, s\}$, then for a type of the Defender who intends to concede if challenged, this budget simplifies to

$$B(m|\theta < \tilde{\theta}) = v_D\psi(m) \quad (7)$$

[6]In the appendix I demonstrate that in any PBE, there must always be types of the Defender who send signal $m = 0$ and that the Challenger best responds to $m = 0$ by issuing a challenge.

and to

$$B(m|\theta \geq \tilde{\theta}) = \psi(m)[v_D(1 - p(s)) + c_D] \tag{8}$$

for a type of the Defender who intends to fight if challenged. If the Defender has privcate information regarding their resolve $\theta = c_D$, then expression (7) and (8) still apply except that it is types $c_D > \tilde{c}_D$ who will concede if challenged and types $c_D \leq \tilde{c}_D$ who will fight if challenged.

As the illustrative example suggests, the budget function $B(m|\theta)$ satisfies certain monotonicity properties. The following proposition formalizes this claim (all proofs are provided in the appendix):

Proposition 1

The Defender's budget for sunk cost signaling $B(m|\theta)$ is:

(i) Strictly increasing in their valuation v_D

(ii) Strictly increasing in their cost-of-fighting c_D for types $c_D \leq \tilde{c}_D$ and constant for types $c_D > \tilde{c}_D$.

(iii) Strictly decreasing in their strength s for types $s \geq \tilde{s}$ and constant for types $s < \tilde{s}$.

The intuition for this result is straightforward and follows directly from equations (7) and (8). First, as the Defender's *valuation* increases, they benefit more from a peaceful resolution of the dispute and are less willing to risk resorting to war to determine ownership of the good. Consequently, the Defender's willingness to pay sunk cost signals to achieve a peaceful settlement of a dispute goes up. This is true regardless of whether the Defender ultimately intends to fight or concede if challenged. By contrast, if the Defender is strong or resolved enough to fight ($s > \tilde{s}$ or $c_D < \tilde{c}_D$), then its willingness to spend on sunk cost signaling decreases as its strength or cost of fighting increases (an increase in resolve implies a lower cost of fighting c_D). This is because it stands less to lose from fighting and so has less incentive to avoid being challenged. Conversely, if the Defender is too weak or unresolved to fight, then its willingness to pay sunk costs becomes independent of its exact type since it will concede if challenged.

A direct consequence of the budget's monotonicity properties is that signaling will fail whenever the Defender has private information regarding their *strength* or *resolve*. Any signal that a strong or resolved type would be willing to send, a weak or unresolved type would be willing to send as well. The following Proposition provides a formal statement of this result.

Proposition 2

If the Defender has private information regarding their strength *or* resolve*, then there exists no PBE in which the Defender sends a signal* $m > 0$*.*

The intuition underlying these results follows directly from equations (7) and (8). Observe that an unresolved or weak type of the Defender will be willing to signal so long as

$$v_D \psi(m) \geq m \tag{9}$$

or whenever the probability of a concession and the value of the good are large enough to merit paying the sunk costs. By comparison for a strong or resolved type of the Defender to be willing to pay sunk costs, it must be the case that

$$\psi(m)[v_D - (V_D p(s) - c_D)] \geq m \tag{10}$$

This condition is more restrictive than the one above. Not only must the value of the prize and the probability of a concession be large, but the difference in value between a concession and going to war must be sufficiently large. As a result of this more restrictive condition, strong or resolved types cannot distinguish themselves from their lower-quality counterparts.

When the Challenger has Private Information Regarding *Valuation* ($\theta = v_D$)

Proposition 1 shows that when countries have private information regarding their *valuation*, their budget for signaling as their valuation increases. This implies that types with higher valuation are willing to outspend their low-quality counterparts in the hopes of obtaining a concession. While this is sufficient to show that there exists a PBE under which signaling occurs ($\sigma(v_D) \neq 0$ for some v_D), I can make a stronger claim.[7]

Let m_1 and m_2 denote two signals satisfying $m_1 > m_2$ and $\psi(m_1) > \psi(m_2)$. The following is a generalized budget equation denoting the difference in utility for a type v_D between a signal m_1 and m_2

$$\begin{aligned} B^{v_D}_{SC}(m_1, m_2, v_D) = \psi(m_1)v_D + (1 - \psi(m_1)) \max\{0, v_D p(s) - c_D\} - m \\ -\psi(m_1)v_D - (1 - \psi(m_1)) \max\{0, v_D p(s) - c_D\} + m_2 \end{aligned} \tag{11}$$

The budget equation (6) is a special case of the generalized budget equation in which $m_2 = 0$. As with equation (6), the generalized budget equation (11) is monotonically increasing in v_D

Proposition 3

The generalized budget equation $B^{v_D}_{SC}$ is strictly increasing in the Defender's valuation v_D

The intuition underlying this result is similar to that underlying proposition 1. This strengthened result is important because it demonstrates that the Defender's expected utility function satisfies increasing-differences and therefore single crossing (Ashworth and Bueno de Mesquita 2006). This property is useful in ruling out alternate equilibria (Ramey 1996).

Let $\check{v}_D$ denote the boundary type of the Defender who the Challenger would be indifferent if they were to face all types with valuation $\check{v}_D$ or higher

$$[F(\check{v}_D) - F(\tilde{v}_D)]v_C + [1 - F(\tilde{v}_D)][v_C(1 - p(s)) - c_C] = 0 \tag{12}$$

The following Proposition formalizes the set of viable PBE.

[7] A formal proof of this sufficiency argument is provided in the online appendix.

Proposition 4

Let v'_D denote a type $v'_D \in [\check{v}_D, \tilde{v}_D]$. Subject to the D1 criterion, any PBE must have the countries play according to

$$\sigma^*(v_D) = \begin{cases} v'_D & \text{if } v_D \geq v'_D \\ 0 & \text{otherwise} \end{cases} ; \; G(m|v) = \begin{cases} \frac{f(s)}{1-F(v'_D)} & \text{if } m \geq v'_D \\ \frac{f(s)}{F(\hat{v}_D)} & \text{otherwise} \end{cases} ; \; \psi(m) = \begin{cases} 1 & \text{if } m \geq v'_D \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

This simple equilibrium replicates the classic result in Fearon (1997): the sufficiently resolved types of the Defender separate by sending signal $m = v'_D$. In response the Challenger can infer that the signal implies that the Defender has valuation v'_D or higher. Consequently the Defender will fight with a sufficiently high probability to induce the Challenger to play $\psi(v'_D) = 1$. All remaining types of the Defender send signal $m = 0$, in response to which the Challenger issues a challenge and the Defender concedes. Three additional features of the equilibrium are worthy of note. First, the signal selected is large enough to make type v'_D is indifferent between signaling and obtaining the good for sure and not signaling at all. Second, the equilibrium is not unique - the threshold type v'_D can vary so long as it comes from the set $v'_D \in [\check{v}_D, \tilde{v}_D]$. Finally, all equilibria are in pure strategies.

Signaling with Differential Costs

The previous section shows that countries will struggle to communicate *strength* or *resolve* with sunk costs because stronger or more resolved states are less willing to spend on attaining peaceful outcomes. In this section, I demonstrate that countries can communicate *strength* or *resolve* when strong or resolved states pay discounted signaling costs. Whenever this assumption holds, strong or resolved countries can produce large signals with minimal effort, thereby circumventing the "budget" issue. I describe conditions under which discounts can replicate the classic sunk-cost signaling result and allow states to successfully communicate *strength* or *resolve*. I conclude this section with a discussion of how differential costs work in practice.

Modeling Differential Costs

In the previous section, the cost of producing a signal of size m was independent of the Defender's type. In this section, we will introduce an effort function $e(m, \theta)$ that determines the cost of producing a signal m for a Defender with type θ. I assume that the Defender must exert more effort to produce a larger signal m, such that $\frac{\partial e(m,\theta)}{\partial m} > 0$ for all types.[8] To allow for higher-quality types to receive a discount, I impose the single crossing assumption on the Defender's effort function (Ashworth and Bueno de Mesquita 2006).

Assumption 1

Let m' and m'' be two signals satisfying $m' > m''$. Similarly, let θ' and θ'' denote two types satisfying $\theta' > \theta''$. If $\theta = s$, then

$$e(m', s'') - e(m'', s'') > e(m', s') - e(m'', s') \tag{14}$$

If $\theta = c_D$, then the inequality is reversed.

Simply, put the assumption states that if there are two types s' and s'' with s' being stronger than s'', then it will take more effort for the weaker types s'' to increase their signal from m'' to m' then it will for the stronger type s'. Whereas the Defender's utility function had the single crossing property arise naturally when the Defender had private information regarding their *valuation* ($\theta = v_D$), assumption 1 is necessary to have the Challenger's expected utility function satisfy the single crossing property when $\theta \in \{s, c_D\}$. Note, that when the Defender has private information regarding their costs of fighting, this relationship is flipped so that a type c'_D with a higher cost of fighting than type c''_D will need to exert more effort to produce larger signals.

[8]In the previous section, this function took the form of $e(m, \theta) = m$

Signaling with Differential Costs

Under the new cost structure on costly signals, we can restate the Defender's generalized budget equation for signaling as

$$\begin{aligned} B_{SC}(m_1, m_2|\theta) :=& v_D\psi(m_1) + (1 - \psi(m)) \times \max\{0, v_D p(s) - c_D\} - e(m_1, \theta) \\ & - v_D\psi(m_2) - (1 - \psi(m_2)) \times \max\{0, v_D p(s) - c_D\} + e(m_2, \theta) \end{aligned} \tag{15}$$

We can restate this updated budget equation as

$$B_{SC}(m|c_D > \tilde{c}_D; s < \tilde{s}) = \psi(m_1)v_D - e(m_1, \theta) - \psi(m_2)v_D + e(m_2, \theta) \tag{16}$$

for a type of the Defender who intends to concede if challenged and as

$$B_{SC}(m|c_D \leq \tilde{c}_D; s \geq \tilde{s}) = \psi(m_1)[v_D(1-p(s)]+c_D)-e(m_1, \theta)-psi(m_2)[v_D(1-p(s)]+c_D)+e(m_2, \theta) \tag{17}$$

for a type of the Defender who intends to fight if challenged.

The following Proposition specifies the conditions required to have the Defender's generalized budget equation be strictly increasing (decreasing) in strength (resolve).

Proposition 5

The Defender's budget for sunk cost signaling $B_{SC}(m|\theta)$ is:

(i) Strictly decreasing for types $c_D > \tilde{c}_D$ and strictly decreasing for types $c_D \leq \tilde{c}_D$ if

$$\frac{\frac{\partial e(m, c_D)}{\partial m}}{v_D(1 - p(s)) + c_D} < \frac{\partial^2 e(m, c_D)}{\partial c_D \partial m} \tag{18}$$

and always strictly decreasing for types $c_D > \tilde{c}_D$.

(ii) Strictly increasing in strength for types $s \geq \tilde{s}$ and $s \geq \tilde{s}$ whenever

$$-\frac{\partial^2 e(m, s)}{\partial m \partial s} > \frac{dp(s)}{ds} v_D \frac{\frac{\partial e(m,s)}{\partial m}}{v_D(1 - p(s)) + c_D} \tag{19}$$

As this proposition demonstrates having the Defender's effort function $e(m, \theta)$ satisfy single-crossing is not a sufficient condition for the Defender's expected utility function to satisfy increasing differences. This is because an increase (decrease) in strength (cost of fighting) has two competing effects for type of the Defender strong (resolved) enough to fight. First, an increase in strength (cost of fighting) makes fighting more (less) attractive. Second, an increase in strength (cost of fighting) reduces (increases) the amount of effort required to signal. For a strong (resolved) type to find spending more on signaling worthwhile as their strength increases (cost of fighting decreases), the discount they receive must exceed their improved payoff from fighting. Formally, the inequality in equation (18) states that for higher cost of fighting types of the Defender to find signaling less worthwhile, the additional costs that they need to spend on signaling (the right-hand side of the inequality) must exceed the marginal benefits of increasing m (the left-hand side of the inequality).[9] Similarly, the inequality in equation (19) states that for a Defender's expected utility for sending larger signals to increase with its strength, the discount that it receives must be larger than its expected increase in probability of obtaining the prize, multiplied by the value of the prize and the marginal cost of signaling.

Note that the generalized budget for signaling is always increasing (decreasing) in the Defender's strength (cost-of-fighting) for types of the Defender who are too weak (unresolved) to fight. Because these choose not to fight, they do not reap any benefit from an improving wartime payoff. In this case, the fact that assumption 1 simply implies that their signaling costs decrease (increase) as their strength (cost of fighting) increases is sufficient to for the increasing differences result.

When the Defender's expected utility function satisfies increasing differences, signaling takes a form akin to that which it takes when the Defender has private information regarding

[9] More specifically, the left-hand side of the inequality is an expression for the marginal costs of increasing m (the additional effort required to signal divided by the difference between having the Defender concede and having to fight). However, the marginal costs and benefits of signaling are equivalent at an optimum. I use the former since the latter are represented by $\psi'(m)$ and are an equilibrium object that can otherwise be difficult to interpret.

their valuation. Let $\hat{c}_D$ denote the boundary type of the Defender who the Challenger would be indifferent to challenge if they believed they were facing all types with cost of fighting $\hat{c}_D$ or lower

$$[F(\hat{c}_D) - F(\tilde{c}_D)]v_C + [F(\tilde{c}_D) - F(\underline{c}_D)][v_C(1 - p(s)) - c_C] = 0 \tag{20}$$

Additionally, let $\tilde{s}_C$ denote the type of Defender that the Challenger is indifferent between challenging and not when $\theta = s$ ($v_C(1-p(\tilde{s}_C)) - c_C = 0$). The following Proposition mirrors Proposition 4:

Proposition 6

Let s' denote a type $s' \in [\rho, \tilde{s}_C]$ and c'_D denote a type $c'_D \in [\tilde{c}_D, \hat{c}_D]$. Subject to the D1 criterion, whenever conditions (18) *and* (19) *are satisfied, any PBE must have*

$$\sigma^*(s) = \begin{cases} e^{-1}(v_D(1 - p(s')) + c_D, \rho) & \text{if } s \geq \rho \\ 0 & \text{otherwise} \end{cases} ; \; \sigma^*(c_D) = \begin{cases} e^{-1}(v_D, c'_D) & \text{if } c_D \leq c'_D \\ 0 & \text{otherwise} \end{cases} ;$$
$$G(s|m, \theta = s) = \begin{cases} \frac{f(s)}{1-F(s')} & \text{if } m \geq e^{-1}(v_D(1 - p(s')) + c_D, \rho) \\ \frac{f(s)}{F(s')} & \text{otherwise} \end{cases} ;$$
$$G(c_D|m, \theta = c_D) = \begin{cases} \frac{f(c_D)}{F(c'_D)} & \text{if } m \geq e^{-1}(v_D, c'_D) \\ \frac{f(c_D)}{1-F(c'_D)} & \text{otherwise} \end{cases} ;$$
$$\psi(m|\theta = s) = \begin{cases} 1 & \text{if } m \geq e^{-1}(v_D(1 - p(s')) + c_D, \rho) \\ 0 & \text{otherwise} \end{cases} ; \; \psi(m|\theta = c_D) = \begin{cases} 1 & \text{if } m \geq e^{-1}(v_D, c'_D) \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

This proposition is very similar to Proposition 4 and so is the intuition underlying it. As such, I will now proceed to discussing how the differential costs assumption works in practice.

Differential Costs in Practice

The differential costs assumption presents a challenge for empirical study. Though the use of sunk costs may be observable and measurable (e.g. A. Post 2019) we do not observe whether a signal's costs are based on the sender's underlying quality. This means that determining whether the differential costs assumption is present or not in a particular case requires that the researcher eliminate other possible explanations for why a state did or did not signal. If no signal was observed, the researcher must rule out that no signal was observed because the sender was weak or unresolved, before attributing it to the absence of differential costs. This is complicated by the fact that strength and resolve are also latent variables (Kertzer 2017).

On the other hand, if a signal is observed, the fact that the sender's signaling cost function is unobserved implies that even the most thorough investigation cannot conclusively determine whether differential costs were present. All that a researcher can do is eliminate or address alternate explanations for the observed signaling. To illustrate how differential costs could work in practice and the difficulties in assessing whether they are present or not in a particular case consider the following example: two weeks after the Russian invasion of Ukraine, French President Macron personally oversaw a nuclear exercise in which Rafaele fighter jets loaded with dummy nuclear bombs practiced bombing an unnamed country (Michaels, Bisserbe, and Gordon 2025). The exercise came days after Russian nuclear sabre rattling and was timed to coincide with the orbit of a Russian spy satellite, to ensure that Russia would be watching.

If differential costs held in this scenario, then the precisely timed exercise was informative because France's underlying resolve enabled it to tolerate the costs of organizing, timing, and conducting the exercise along with any permanent increase in the probability of nuclear escalation. Conversely, had President Macron been unresolved, then he would have found the conduct of the exercise in this manner too onerous. Russia should then have realized that the fact that France was willing to conduct the exercise in such a manner was an indication of its ability to shoulder the associated costs, thereby making it a credible signal.

What else could explain France's signal? First, it is possible that the signal was intended to convey *valuation*, which does not require differential costs for sunk cost signaling, as opposed to *resolve*, which does. This is unlikely. Before the Russian invasion, President Macron had taken a prominent role trying to mediate the crisis and had committed to deploy soldiers to Bulgaria and Romania (Reuters 2022; Weber 2022). France's heavy involvement in the crisis likely conveyed that it cared about the issue enough to devote effort towards seeking a favorable resolution to the dispute – actions that a high-valuation type would be willing to undertake. However, French pre-invasion actions did not convey resolve, i.e. a willingness to tolerate the costs of confronting Russia. Macron's diplomatic outreach included a visit to Moscow, attempts to eschew NATO and keep independent channels of communication open with Russia, and a French position in favor of NATO passivity so as not to provoke Russia – actions which were spurred in part by the French assessment that Russia would not invade Ukraine (Stoltenberg 2025, pp. 342-344).

Second, it is possible that the French military exercise was also designed to increase French strength, in which case it would constitute an instance of "arming'.' This would make it impossible to infer the presence of differential costs from the occurrence of signaling alone as arming is both possible with differential costs and, as will be shown below, without them. However, in this particular instance, the nuclear exercise that Macron oversaw is conducted in some form four times a year (Powis 2025). It is therefore unlikely that this shifted French military strength in any significant way and could constitute an instance of arming.

There may be other potential explanations for the French signal. For example, in the next section, I will contend with whether France's signal should be interpreted as an index signal. Ultimately, if no alternate explanations for the French signal remain, then Propositions 2 and 5 imply that differential costs were present since these are required to signal resolve with sunk costs. However, this style of argument is not as convincing as direct evidence in favor of differential costs would be and in essence leaves the argument that differential costs underwrote the French signal a conjecture. As this example illustrates the unobservable nature of signaling cost functions would likely complicate any systematic attempt to assess their prevalence.

Index Signals

Whereas the signaling advantage described above allows for costly communication by ensuring that weaker types don't want to mimic costly signals, index signals are sunk cost signals that weak or unresolved types cannot mimic because they do not have the capacity to do so (Maynard-Smith et al. 2003). In this section, I demonstrate that communication of *strength* or *resolve* is possible whenever stronger or more resolved types of the Defender have access to a signaling technology that is sufficiently exclusive and cheap to use, thereby circumventing the "budget" issue. I also provide examples that illustrate what such signals look like in practice.

Modeling Index Signals

To introduce index signals, I now assume that there exist a set of types $[s^i, \bar{s}]$ with $s^i \in (\underline{s}, \bar{s})$ or $[\underline{c}_D, c_D^i]$ with $c_D^i \in [\underline{c}_D, \bar{c}_D)$ that can choose whether or not to send a signal $m \in \{0, 1\}$ at a cost e. Types $s \in [\underline{s}, s^i)$ can only send signal $m = 0$ at no cost. The following Proposition demonstrates that there are two conditions required for an equilibrium in which the strongest types issue an index signal. First, the signaling technology must be exclusive enough to be informative of the sender's quality. Second, the signal must be sufficiently cheap for strong types to be willing use it.

Index signals circumvent the problem of stronger or more resolved types having smaller budgets for signaling by reducing the set of potential senders. An inspection of the the

Defender's budget for sending index signals

$$B_{index}(e,\theta) = v - e - vp(s) + c_D \tag{22}$$

reveals that its budget for signaling is once again strictly decreasing in the Defender's strength and resolve (i.e. increasing with their cost of fighting) as in Proposition 1. As a result, if the cost of sending an index signal is high enough, it is possible for the strongest types of the sender to prefer to fight even with access to the signaling technology. However, this does not imply that more moderately strong types cannot make use of the signal. The following Proposition formalizes this result.

Proposition 7

If equation (22) *is positive for type* $\underline{\theta}_i$, *then let* $\hat{s}$ ($\hat{c_D}$) *denote the strongest (most resolved) type of the Defender for whom equation* (22) *is non-negative. Then there exists a PBE in which*

$$\sigma(s|\theta=s) = \begin{cases} 0 & \text{if } s > \hat{s} \\ 1 & \text{if } s_i \leq s \leq \hat{s} \\ 0 & \text{if } s < s^i \end{cases} ; \ \sigma(c_D|\theta=c_D) = \begin{cases} 0 & \text{if } c_D > c_D^i \\ 1 & \text{if } \hat{c}_D \leq c_D \leq c_D^i \\ 0 & \text{if } c_D < \hat{c}_D \end{cases}$$
$$G(s|m,\theta=s) = \begin{cases} \frac{f(s)}{F(\hat{s})-F(s^i)} & \text{if } m=1 \\ \frac{f(s)}{F(s^i)+1-F(\hat{s})} & \text{otherwise} \end{cases} ; \ G(c_D|m,\theta=c_D) = \begin{cases} \frac{f(c_D)}{F(\hat{c}_D)-F(c_D^i)} & \text{if } m=1 \\ \frac{f(c_D)}{F(\hat{c}_D)+1-F(c_D^i)} & \text{otherwise} \end{cases}$$
$$\psi(m) = \begin{cases} 1 & \text{if } m=1 \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

if $c_D^i \leq \tilde{c}_D$ *or* $c_D^i > \tilde{c}_D$ *and satisfies*

$$[F(\tilde{c}_D) - F(c_D^i)]v_C + [F(\hat{c}_D) - F(\tilde{c}_D)][v_C(1-p(s)) - c_C] = 0 \tag{24}$$

whenever $\theta = c_D$. If $\theta = s$, then this is a PBE if $s^i \geq \tilde{S}_C$ or if

$$v_C \int_{s^i}^{\hat{s}} 1 - p(s)ds - c_C \leq 0 \tag{25}$$

whenever $\rho \leq s^i < \tilde{s}_C$

This proposition is straightforward and a proof is therefore omitted. When the index signaling technology is only available (i) to types sufficiently resolved to fight if challenged or (ii) to types that are strong enough to deter the Challenger from issuing a challenge, then the Challenger knows to stand down upon observing the signal m and so a PBE is possible. Alternatively, if the signal is available either to (iii) unresolved types willing to stand firm or (iv) types too weak to deter the Challenger, then the mass of types with access to the technology *and* willing to signal must on the whole deter the Challenger for the signal to be sent. All this is contingent on the signal being cheap enough so to be preferable to fighting for a sufficiently large enough mass of strong or resolved types with access to the technology.

Index Signals in International Relations

Index signals originate in evolutionary biology (Maynard-Smith et al. 2003, pp. 1, 45–67). Scholars studying Scottish red deer observed that during the mating season large males protected their harems by roaring frequently and loudly (Clutton-Brock and Albon 1979). Roaring in this way is costly as males lose a significant portion of their body weight doing so and was found to correlate with a stag's ability to both deter and defeat challenges. Alternatively, tigers mark their territory and attempt to deter rivals from moving in by clawing trees as high as they can (Thapar 1986, p. 43). Since larger tigers can reach higher, this sends an honest advertisement of the tiger's size. In both of these cases this message is considered to be an index signal because it cannot be mimicked by low-quality types - weak deer cannot roar in this way and small tigers cannot reach very high.

Unlike differential costs, the empirical study of index signals is more straightforward and the IR literature already contains several references to them even if the term had not been formalized.[10] Slantchev (2011, p. 78) echoes the concept when he discusses how "military moves can, under certain circumstances reveal one's strength unambiguously." Montgomery (2020) details how states try to advertise their military power through the the employment of military capabilities, public exercises, weapons tests, and exhibitions. For example, between 2016 and 2017, North Korea conducted three successful nuclear tests and two successful tests of ICMBs and boasted that it had now achieved deterrence vis-a-vis the US (Sankaran and Fetter 2022). What makes these demonstrations of capabilities informative is that they cannot be faked - North Korea demonstrated a capability that it could not have if it were a weaker type (Maynard-Smith et al. 2003).

The prevlance of index signals makes it possible to find cases that correspond to different parts of the parameter space. Proposition 7 demonstrates that it is possible for weak types to get away with bluffing by using index signals as long as the pool of types capable of doing so is sufficiently small. As an example of this consider Germany in 1934, when Albert Speer, requested the use of 130 searchlights, then an important component of air-defense, for the use in a party rally for aesthetic purposes (Speer 1970, pp. 100-101). Herman Göring, the future head of the German airforce, balked at the request since this constitued the bulk of Germany's strategic reserve of searchlights. However, Hitler overruled Göring and, explicitly referencing the strategic logic of index signals, argued that "If we use them in such large numbers for a thing like this, other countries will think we're swimming in searchlights." This represents an index signal because Germany could only conduct such a display if it in fact had 130 searchlights to deploy. For a depiction of this usage for the searchlights, see Figure 2.

There are two important limitations to note regarding index signals. First, existing work suggests that index signals are rarely used during crises. Montgomery (2020) reasons that

[10] Jervis (1970) introduced the term indices into IR as a concept distinct from signals. The former meant "observable, unalterable attributes" which were conceived as credible sources of information by virtue being perceived as not being manipulable (Spence 1973, p. 357). However, my focus is on *index signals*, manipulable actions that are observable and can only be indicative of an unalterable attribute.

Figure 2: “The Cathedral of Light:” A photo of a Nazi rally using searchlights in Nuremberg, 1936. Source: wikimedia commons. Searchlights were an important tool of aerial defense and their use for aesthetic purposes was controversial. Hitler approved this usage because it suggested Germany had large military stockpiles.

such signals are mostly used in peacetime to sustain general deterrence, as was the case in the examples provided above. Green and Long (2020) reach a similar conclusion, and further argue that demonstrations of clandestine capabilities during war or crises may both be more difficult to interpret under a noisier environment and may be more likely may increase the risk that enemy countermeasures may mute the benefits of the revealed capability. Second, it is difficult to conceptualize what an index signal of *resolve* would look like. Such a signal would require a signaling technology exclusive to low cost-of-fighting types. However, the military technology associated with the use of index signals is associated with strength rather then resolve. Returning to the French example from the previous section, for the France's exercise to constitute an index signal it would have to have been the case that only a highly resolved type of France could conduct an exercise simulating air-delivered nuclear strikes. Though it is clear why France might not *want* to conduct such an exercise if it were of low resolve, it does not seem that being low resolve would preclude it from being able to conduct an exercise.

Sunk Costs as Investments in Arming

A number of scholars have argued that "pure" sunk cost signals that don't affect the sender's payoff beyond their affect on the Challenger's beliefs may be rare. Instead, these scholars contend that sunk costs are often investments in military readiness, such as troop mobilizations or arming that shift the balance of power in favor of the signaler (Slantchev 2005; A. S. Post and Sechser 2024). Existing work (Slantchev 2005) restricts itself to studying arming when the sender has private information regarding their valuation and shows that there exists a PBE in which (i) the highest-valuation types arm themselves enough to deter the Challenger and (ii) medium-valuation types arm themselves enough to prepare for an optimal fight with the Challenger. This section studies how costly signaling operates when sunk costs represent an investment in arms and the Defender has private information regarding their *strength* or *resolve*.

This section demonstrates that there exists PBE in which the Defender pays sunk costs to arm even absent differential costs or index signals. However, it is the weakest types of the Defender and those with middling costs-of fighting who choose to arm. Though signaling reveals that the Defender was originally weak or unresolved, an initially unresolved or weak type of the Defender arms enough to deter the Challenger. Conversely, the initially strongest and most resolved types continue to have a smaller budget for signaling and subsequently do not arm at all. However, the Challenger recognizes that an opponent would only remain unarmed if they were a formidable opponent to begin with and so is still deterred from Challenging with positive probability.

Modeling Sunk Costs as an Investment in Arming

Formally, let the Challenger's probability of winning a war now be a concave function of its level of signal $p_1(s_1, m)$ so that $\frac{\partial p(s,m)}{\partial m} > 0$ and $\frac{\partial^2 p(s,m)}{\partial^2 m} < 0$. Moreover, assume that no type enjoys any advantage in arming so that $\frac{\partial^2 p_1(s_1,m)}{\partial s_1 \partial m} = 0$. The Challenger has the following expected utility

$$U(m|\theta) = \psi(m)v_D + (1 - \psi(m)) \max\{0, v_D p(s,m) - c_D\} - m \tag{26}$$

and has a generalized budget for signaling given

$$\begin{aligned} B_a(m_1, m_2|\theta) = \psi(m_1)v_D + (1 - \psi(m_1)) \max\{0, v_D p(s, m_1) - c_D)\} - m_1 \\ -\psi(m_2)v_D - (1 - \psi(m_2)) \max\{0, v_D p(s, m_2) - c_D\} + m_2 \end{aligned} \tag{27}$$

for a signal $m_1 > m_2$. The following proposition shows that arming does not affect the monotonicity properties first explored in previous sections, with higher valuation types willing to pay more to obtain the good and stronger and more resolved types willing to pay less.

Proposition 8

The Defender's budget for sunk cost investments in arming $B_a(m_1, m_2|\theta)$ ($m_1 > m_2$) is:

(i) Strictly increasing in its valuation v_D

(ii) Strictly increasing in their cost-of-fighting c_D for types $c_D \leq \tilde{c}_D(m_1)$ and constant for types $c_D > \tilde{c}_D(m_1)$ for $\tilde{c}_D(m_1) := m_1 - p(s, m_1)$.

(iii) Strictly decreasing in their strength s for types $s \geq \tilde{s}(m_1)$ for higher probabilities of concession ($\psi(m_1) > \psi(m_2)$) and constant for types $s < \tilde{s}(m_1)$ for $\tilde{s}(m_1) = p^{-1}(\frac{m_1+c_D}{v_D}, m_1)$

The intuition for this result is similar to that underlying Propositions1 and 3 except for two minor differences. First is that arming now increases lowers the bar for the threshold for types of Defender who are strong or resolved enough to fight. In doing so it makes the threshold between when budgets start decreasing for sufficiently resolved or strong enough types to fight endogenous to the level of arming. Second, is that the Defender's budget may be constant in s when $\psi(m_1) = \psi(m_2)$. This does not constitute a part of the PBE defined below and occurs because I have assumed that the cross-partial for $p(m, s)$ satisfies $\frac{\partial^2 p(m,s)}{\partial m \partial s} = 0$. For a constant probability of concession all types who are sufficiently strong enough to fight ($s > \tilde{s}(m)$), arming becomes a simple cost-benefit analysis that weighs the benefits of improved wartime payoff versus the costs of arming.

Unlike the baseline case, signaling is still possible despite the Defender's budget for signaling decreasing in strength or resolve whenever it is possible for the weakest or lowest resolve types to arm enough to deter the Defender outright. Let $\check{c}_D$ be defined as the type of the Defender that satisfies the following

$$F(\tilde{c}_D)(v_C(1 - p(s, 0)) - c_C) + (1 - F(\check{c}))v_C = 0$$

when $\theta = c_D$. The following Proposition formalizes the stated result:

Proposition 9

Let $\theta = s$ and assume that $v_D - p^{-1}\left(\underline{s}, 1 - \frac{c_C}{v_C}\right) > 0$. Then there exist a PBE in which the weakest types separate by each adopting a unique level of arms and the strongest types do not arm at all

$$\sigma(s|\theta = s) = \begin{cases} p^{-1}\left(s, 1 - \frac{c_C}{v_C}\right) & \text{if } s < \rho \\ 0 & \text{otherwise} \end{cases}$$
$$G(s|m, \theta = s) = \begin{cases} 1 & \text{if } m = p^{-1}\left(s, 1 - \frac{c_C}{v_C}\right) \\ \frac{f(s)}{1 - F(\rho)} & \text{if } m = 0 \end{cases} \tag{28}$$

and the Challenger responds by conceding with higher probability in response to higher levels *of arms*

$$\psi(m|\theta = s) = \begin{cases} 1 \text{ if } m = p^{-1}(\underline{s}, \chi) \\ \ell^*(m) \text{ if } m \in \left(p^{-1}(\underline{s}, \chi), p^{-1}(\rho, \chi)\right) \\ \frac{\psi\left(p^{-1}(\rho,\chi)\right)v_D + \left(1-\psi\left(p^{-1}(\rho,\chi)\right)\right)\left(v_D p^{-1}(\rho,\chi) - c_C\right) - p^{-1}(\rho,\chi) - v_D p(\rho,0) - c_D}{v_D - v_D p(\rho,0) - c_D} \text{ if } m = 0 \end{cases} \tag{29}$$

where $\chi = 1 - \frac{c_C}{v_C}$ *and* $\ell^*(m)$ *is given by the solution to the following differential equation*

$$\frac{\psi'(m)}{1-\psi(m)} = \left(\frac{1}{1-\psi(m)} - v_D \frac{\partial p(s,m)}{\partial m}\right) \cdot \left(\frac{1}{v_D(1-p(s,m)) + c_D}\right)$$

Let $\theta = c_D$ *and assume that* $v_D - p^{-1}\left(s, \frac{\bar{c}_D}{v_D}\right) > 0$*. Then there exist a PBE in which the middling cost-of-fighting types separate by each adopting a unique level of arms and the highest and lowest cost-of-fighting types do not arm at all.*

$$\sigma(c_D|\theta = c_D) = \begin{cases} 0 & \text{if } c_D \leq \tilde{c}_D \\ p^{-1}\left(s, \frac{c_D}{v_D}\right) & \text{if } \tilde{c}_D < c_D < \check{c}_D \\ 0 & \text{otherwise} \end{cases} \tag{30}$$

$$G(c_D|m, \theta = c_D) = \begin{cases} 1 & \text{if } m = p^{-1}\left(s, \frac{c_D}{v_D}\right) \\ \frac{f(c_D)}{1-F(\check{c}_D)+F(\tilde{c}_D)} & \text{if } m = 0 \end{cases}$$

and the Challenger responds by conceding with higher probability in response to higher levels of arming

$$\psi(m) = 1 - \frac{p^{-1}\left(s, \frac{\check{c}_D}{v_D}\right) - m}{v_D} \tag{31}$$

If challenged a type of the Defender who selects an $m > 0$ *concedes with probability*

$$\phi(m) = \frac{-(v_C(1-p(s,m)) - c_C)}{v_C p(s,m) + c_C} \tag{32}$$

The equilibrium described in the case where $\theta = s$ is unique when off-path beliefs are subject to the D1 refinement. When $\theta = c_D$ the equilibrium described delivers all types of the Defender with their highest expected utility of all equilibria that survive D1.

The following is the intuition underlying the result. When $\theta = s$, then condition (i) states that arming will be cheap enough such that the (initially) weakest type of the Defender $\underline{s}$ will be willing to arm themselves to a level that deters the Challenger from challenging. Each type of the Defender weaker than ρ then arms themselves to a level that leaves the Challenger indifferent between arming and not. The indifferent Challenger then selects a probability of issuing a challenge that ensures that the Defender's strategy is incentive compatible, a quantity given by equation $\ell^*(m)$. Proposition 8 demonstrated that stronger types of the Defender have less to lose from fighting and are therefore less willing to pay sunk costs to avoid war. Equation $\ell^*(m)$ is the hazard rate representing how the Challenger's probability of concession changes as the Defender arms and is designed to ensure that initially weaker types of the Defender are more incentivized to arm. Consequently, the initially weakest types of the Defender ($s < \rho$) perfectly separate. However, there exist a set of types that are strong enough even without arming that the Challenger would rather not fight these types even if they did not arm. These types cannot make the Defender indifferent and therefore cannot arm. Instead, they pool on not arming with a sufficiently large mass of weaker types such that the Challenger is indifferent between arming and not. In response, the Challenger concedes with probability $\psi(0)$ which is designed to keep type ρ indifferent between separating and not arming at all.

Similarly, when $\theta = c_D$ the lowest cost-of-fighting types ($c_D < \tilde{c}_D$) have the least to lose from deterring the Challenger and do not arm. However, this time is the intermediate types $c_D \in (\tilde{c}, \check{c})$ who arm with the goal of deterring the Challenger. These are the types whose costs of fighting were only slightly too high to merit fighting after being challenged while unarmed. Proposition 8 still implies that higher cost-of-fighting types have larger budgets for signaling so it might seem puzzling that it is not the highest cost-of-fighting types who signal.

However, this is because these moderate cost-of-fighting types arm to a level where they, as opposed to the Challenger, are indifferent between fighting and not. The Challenger concedes with probability 1 in response to the highest level of arming (that which would be played by type $\check{c}_D$). In response to lower-levels of arming, the Challenger reduces the probability with which they concede at a rate that ensures that types $c_D \geq \tilde{c}_D(m)$ are indifferent between m or a larger level of arms. Since the Defender's budget is weakly increasing in its cost of fighting, sufficiently low cost-of-fighting types ($c_D < \tilde{c}_D(m)$) strictly prefer lower levels of arming. Observe that at the unique level of arming adopted by each type of the arming types, the Defender is indifferent between fighting the Challenger and conceding, but the Challenger would strictly prefer not to challenge then fight. To ensure that the Challenger is willing to mix between challenging and not, the Defender must mix between conceding and fighting if challenged. Perhaps surprisingly, the more the Defender is armed the more likely they are to concede if challenged.

Conclusion

This paper builds on a burgeoning literature in signaling theory to argue that resolved or strong states have weaker incentives to communicate private information about their willingness to fight in a crisis. Using a formal model, I demonstrate that strong or resolved states' higher payoff to fighting translates implies that they benefit less from getting their rival to concede. Consequently, as a state becomes stronger or more resolved it becomes less willing to invest in paying sunk cost signals to avert a war. This inhibits signaling by preventing separation - any signal that a strong or resolved state would find worthwhile to send, an unresolved or weak state would be willing to send also.

I demonstrate that states may still be able to signal anyway if additional assumptions are present. However, these weaken the applicability of costly signaling theory - differential costs are unobservable and will likely prove difficult to measure or proxy. Index signals are

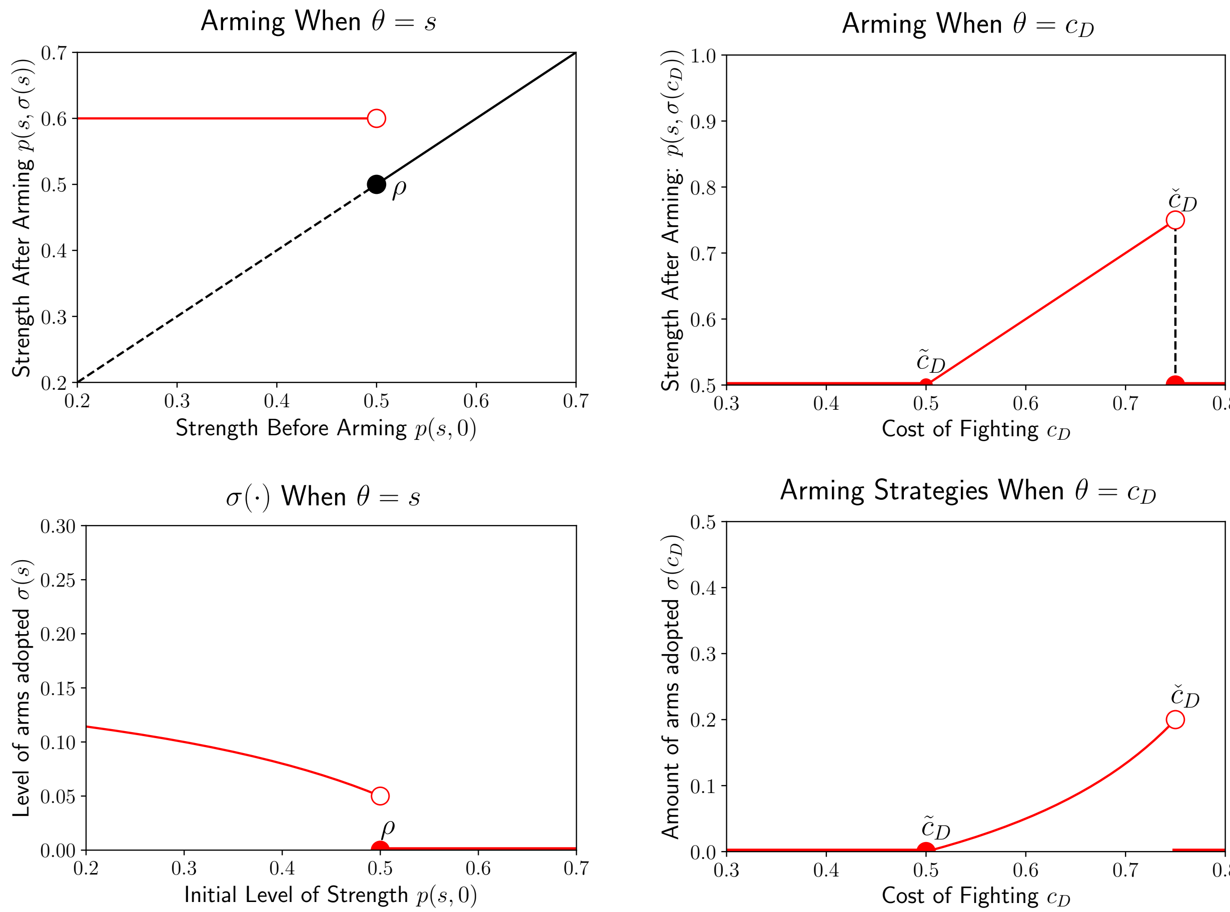

Figure 3: **Arming when** $\theta = s, c_D$: In all figures the x-axis depicts the Defender's type. In the top row the y-axis represents the Defender's probability of winning post-arming. In the bottom row, the y-axis represents the absolute levels of arms m the Defender adopts. When $\theta = s$, sufficiently weak types, $s < \rho$ of the Challenger arm themselves to the level of arms required to make the Challenger indifferent between fighting them and not. Conversely, the Challenger is indifferent when fighting an unarmed Defender when types $[\rho, \bar{s}]$ pool on $m = 0$. When $\theta = c_D$, initially unresolved types of the Defender in the range $(\tilde{c}_D, \check{c}_D)$ each arm themselves just enough to make themselves indifferent between fighting the Challenger and not. The Challenger is indifferent between challenging in response to no arming, when it believes that it is facing types in the range $[\underline{c}_D, \tilde{c}_D]$ which it expects to fight for sure and types in the range $[\check{c}_D, \bar{c}_D]$ which it expects to concede in response to a challenge. In the all figures θ is distributed uniformly and $v_D = 1$. When $\theta = s$ ($\theta = c_D$), $\underline{s} = 0.2$ and $\bar{s} = 0.7$ ($\underline{c}_D = 0.3$ and $\bar{c}_D = 0.8$), $c_D = 0.2$ ($s = 0.5$), and $p(s, m)$ is a contest function with form $p(s, m) = s + (1 - \bar{s})\frac{m}{k+m}$ when $\theta = s$ and ($p(s, m) = s + (1 - s)\frac{m}{k+m}$) when $\theta = c_D$ where k is a constant with value $k = 0.2$ The Challenger's valuation for the good is $v_C = 1$ in both columns and its cost of fighting is equal to $c_D = 0.4$ in the left columns when $\theta = s$ and $c_D = 0.4$ in the right column when $\theta = c_D$

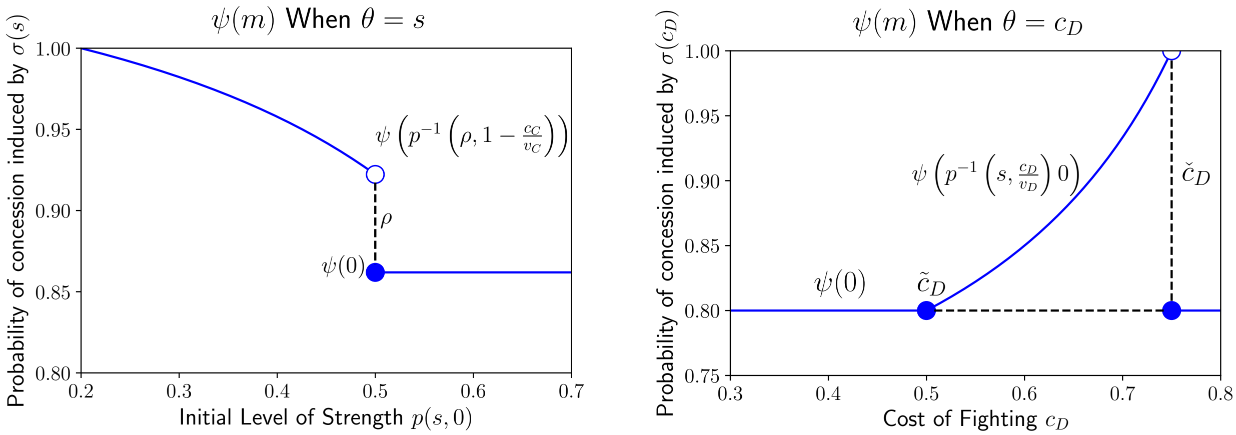

Figure 4: **Challenger's response to arming when $\theta = s, c_D$:** The Challenger responds to more arming with a greater probability of concession. In the left-hand figure, when $\theta = s$, because the weakest types of the Challenger arm themselves the most, they are the most likely to receive a concession. Types with strength ρ and higher do not arm at all and receive a probability of concession $\psi(0)$ designed to make type ρ indifferent between arming and not. Conversely, in the right-hand figure when $\theta = c_D$, it is the initially moderately unresolved types who arm, with higher-cost-of-fighting type arming more, and who are most likely to receive a concession. The most resolved types and the least resolved types do not arm at all. The model parameters are the same as those in Figure 3.

much more prevalent, but are likely used more heavily in peacetime than in crisis settings.

Alternatively, the assumptions required for states to spend on sunk cost signals when these constitute arming are much weaker. I show that in this case, strong or resolved states' reduced incentive to spend on arming persists and they arm less. By contrast, weak or unresolved states will separate and reveal their initial low-quality by arming enough to deter the attacker anyway. Though these set of results suggest that sunk cost signaling, through arming, may be more pervasive than "pure" sunk cost signaling, they complicate costly signaling theory and its implications for empirical scholars. This is because it breaks the comparative static predictions traditionally considered to be a part of costly signaling theory - though larger signals are more likely to be associated with a concession, larger signals may be associated with a reduced payoffs to fighting when the sender has private information regarding their strength or resolve.

References

Altman, Dan and Kai Quek (June 2024). “Do States Really Sink Costs to Signal Resolve?” *Journal of Global Security Studies* 9.2.

Arena, Philip (2013). “Costly Signaling, Resolve, and Martial Effectiveness, Unpublished Manuscript, University of Buffalo”.

Ashworth, Scott and Ethan Bueno de Mesquita (2006). “Monotone Comparative Statics for Models of Politics”. *American Journal of Political Science* 50.1, pp. 214–231.

Carroll, Robert and Amy Pond (July 2021). “Costly signaling in autocracy”. *International Interactions* 47.4, pp. 612–632.

Clutton-Brock, T. H. and S. D. Albon (Jan. 1979). “The Roaring of Red Deer and the Evolution of Honest Advertisement”. *Behaviour* 69.3-4, pp. 145–170.

Fearon, James D. (1994). “Domestic Political Audiences and the Escalation of International Disputes”. *American Political Science Review* 88.3, pp. 577–592.

— (1995). “Rationalist Explanations for War”. *International Organization* 49.3, pp. 379–414.

— (1997). “Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs”. *Journal of Conflict Resolution* 41.1, pp. 68–90.

Fey, Mark and Kristopher W. Ramsay (2011). “Uncertainty and Incentives in Crisis Bargaining: Game-Free Analysis of International Conflict”. *American Journal of Political Science* 55.1, pp. 149–169.

Gartzke, Erik A. et al. (2017). “Signaling in Foreign Policy”. In: *Oxford Research Encyclopedia of Politics*. Oxford University Press.

Green, Brendan Rittenhouse and Austin Long (Jan. 2020). “Conceal or Reveal? Managing Clandestine Military Capabilities in Peacetime Competition”. *International Security* 44.3, pp. 48–83.

Jervis, Robert (1970). *The Logic of Images in International Relations*. Princeton University Press.

Joseph, Michael F. (Jan. 2021). "A Little Bit of Cheap Talk Is a Dangerous Thing: States Can Communicate Intentions Persuasively and Raise the Risk of War". *The Journal of Politics* 83.1, pp. 166–181.

Kertzer, Joshua D. (Jan. 2017). "Microfoundations in International Relations". *Conflict Management and Peace Science* 34.1, pp. 81–97.

Kurizaki, Shuhei (2007). "Efficient Secrecy: Public versus Private Threats in Crisis Diplomacy". *American Political Science Review* 101.3, pp. 543–558.

Maynard-Smith, Professor Emeritus School of Biological Sciences John et al. (Nov. 2003). *Animal Signals*. Oxford University Press.

Michaels, Daniel, Noemie Bisserbe, and Michael R. Gordon (Mar. 2025). *Trump Prompts European Calls for a Homegrown Nuclear Umbrella*. URL: https://www.wsj.com/world/europe/europe-nuclear-umbrella-france-britain-nato-10c78d15 (visited on 01/06/2026).

Montgomery, Evan Braden (Feb. 2020). "Signals of strength: Capability demonstrations and perceptions of military power". *Journal of Strategic Studies* 43.2, pp. 309–330.

Morrow, James D. (1997). "The Strategic Setting of Choices: Signaling, Commitment, and Negotiation in International Politics". In: *Strategic Choice and International Relations*. Princeton University Press.

Post, Abigail (Apr. 2019). "Flying to Fail: Costly Signals and Air Power in Crisis Bargaining". *Journal of Conflict Resolution* 63.4, pp. 869–895.

Post, Abigail S and Todd S Sechser (June 2024). "Hand-Tying through Military Signals in Crisis Bargaining". *International Studies Quarterly* 68.2, sqae028.

Powis, Gaétan (Mar. 2025). *Poker 2025-01 : Rarissime Exercice Aérien Nucléaire Majeur sur la France aux Premières Lueurs du Jour*. URL: https://air-cosmos.com/article/poker-2025-01-rarissime-exercice-aerien-nucleaire-majeur-sur-la-france-aux-premieres-lueurs-du-jour-70065 (visited on 01/07/2026).

Ramey, Garey (May 1996). “D1 Signaling Equilibria with Multiple Signals and a Continuum of Types”. *Journal of Economic Theory* 69.2, pp. 508–531.

Reich, Noam (Aug. 2022). “Signaling Strength with Handicaps”. *Journal of Conflict Resolution* 66.7-8, pp. 1481–1513.

— (July 2025). “Dynamic Screening in International Crises”. *The Journal of Politics* 87.3, pp. 1079–1093.

Reuters (Mar. 2022). *France’s Macron: Russia’s Putin Alone Chose War in Ukraine*. URL: https://www.reuters.com/world/europe/frances-macron-russias-putin-alone-chose-war-ukraine-2022-03-02/ (visited on 01/07/2026).

Sankaran, Jaganath and Steve Fetter (Feb. 2022). “Defending the United States: Revisiting National Missile Defense against North Korea”. *International Security* 46.3, pp. 51–86.

Slantchev, Branislav L. (2005). “Military Coercion in Interstate Crises”. *American Political Science Review* 99.4, pp. 533–547.

— (July 2010). “Feigning Weakness”. *International Organization* 64.3, pp. 357–388.

— (Feb. 2011). *Military Threats: The Costs of Coercion and the Price of Peace*. Cambridge University Press.

Speer, Albert (1970). *Inside the Third Reich: Memoirs*. New York: Macmillan.

Spence, Michael (Aug. 1973). “Job Market Signaling”. *The Quarterly Journal of Economics* 87.3, pp. 355–374.

Stoltenberg, Jens (2025). *On My Watch: Leading NATO in a Time of War*. W. W. Norton & Company.

Thapar, Valmik (1986). *Tiger: Portrait of a Predator*. Facts on File Publications.

Weber, Gesine (Mar. 2022). *A Case for French Leadership on Ukraine*. URL: https://warontherocks.com/2022/03/a-case-for-french-leadership-on-ukraine/ (visited on 01/07/2026).

Wolford, Scott (2014). “Showing Restraint, Signaling Resolve: Coalitions, Cooperation, and Crisis Bargaining”. *American Journal of Political Science* 58.1, pp. 144–156.

Wolton, Stephane (Jan. 2024). “Decentralised Information Transmission in the Shadow of Conflict”. *Journal of Theoretical Politics* 36.1, pp. 64–82.

Online Supplemental Appendix

When Can States Signal with Sunk Costs?

Contents

1 Preliminary Lemmas

It is useful to begin with a number of preliminary results.

1.1 Lemma A. 1

Lemma A. 1
In any PBE,

(i) there exists a subset of the Defender types who send signal $m = 0$ with positive probability.

(ii) the Challenger always challenges in response to signal $m = 0$.

Proof of Lemma A. 1:
First we want to demonstrate that any PBE must include an on-the-equilibrium-path signal m_0 that induces the Challenger to respond by challenging with probability 1. Suppose not, that is suppose that a PBE existed in which there was no signal m_0 sent. This would imply that the Challenger (weakly) preferred not to challenge in response to any given signal m'. Let $\theta = v_D$ so that

$$v_C \int_{v_D|\sigma(v_D)=m';v_D<\tilde{v}_D} f(v_D)dv_D + [v_c(1-p(s)-c_C] \int_{v_D|\sigma(v_D)=m';v_D\geq\tilde{v}_D} f(v_D)dv_D \leq 0$$

for any given signal m'. Summing up these inequalities for all on-path signals, we obtain that the Challenger's expected prior to any signaling is

$$v_C F(\tilde{v}_D) + [1-F(\tilde{v}_D)][v_c(1-p(s)-c_C] \leq 0$$

a contradiction of the assumption imposed in equation (1). Therefore, there must exist a signal m_0 in response to which the Challenger challenges with probability 1. The proof for this claim when $\theta \in \{c_D, s\}$ is isomorphic.

Second, we want to show that there can be no PBE in which the Defender does not send signal $m = 0$. Suppose not. Per the previous step there must be a signal m_0 which causes the Defender to challenge with probability 1. Given that signal $m = 0$ is not an on-the-equilibrium-path signal, $m_0 > 0$. Observe that any type sending signal $m_0 > 0$ could strictly increase their expected payoff by deviating from sending signal m_1 to m_0 regardless of the Challenger's off-path belief. Therefore, every PBE must have some types of the Defender sending signal $m = 0$.

Finally, we want to show that the Challenger must issue a challenge in response to signal $m = 0$. If signal $m = 0$ is the only signal sent in equilibrium, then per the assumptions imposed in equations (1), (2), and (3), no information is conveyed and the Challenger's best response must be to challenge. Alternatively, suppose that there exists a PBE in which the Challenger sends two or more signals, that per the previous step signal $m = 0$ is sent, and that the Challenger's best response to $m = 0$ is to challenge with positive probability. Per the first step of this proof there exists a signal $m_0 > 0$ in response to which the Challenger issues a challenge with probability 1. It follows, that any type sending signal m_0 has a strictly profitable deviation to sending signal $m = 0$ instead. A contradiction. ■

1.2 Lemma A. 2

Lemma A. 2
The most a Defender of type θ will spend on a signal m is given by equation (6)*.*

Proof of Lemma A. 2 Per Lemma A. 1, in any PBE a subset of the Defender's types must send signal $m = 0$ with positive probability and have the Challenger challenge in response. It follows that the Defender's can never do worse than sending $m = 0$ and having an expected utility of $\max\{0, v_D p(s) - c_D\}$.

In response to any signal $m > 0$, the Challenger may concede with positive probability. Even if the Challenger issues a challenge in response to a signal $m > 0$, the Defender retains the option to concede or fight. Therefore, the Defender's expected utility for sending any signal $m > 0$ is given by

$$\psi(m) + (1 - \psi(m)) \times \max\{0, v_D p(s) - c_D\} - m \qquad \text{(A. 1)}$$

It follows that a Defender would prefer sending signal $m > 0$ to $m = 0$ whenever,

$$\psi(m) + (1 - \psi(m)) \times \max\{0, v_D p(s) - c_D\} - m \geq \max\{0, v_D p(s) - c_D\}$$

and that if this inequality does not hold, then the Defender is strictly better off sending signal $m = 0$ instead.

For any given fixed value of $\psi(\cdot)$, the highest value of m at which the Defender would find signaling at least as good as sending $m' = 0$ is the value of m for which the above inequality holds with equality, or

$$m(\theta) = psi(m) + (1 - \psi(m)) \times \max 0, v_D p(s) - c_D - \max 0, v_D p(s) - c_D$$

This is the formal definition of the budget equation provided in equation (6). ■

1.3 Lemma A. 3

Lemma A. 3
$\psi(m)$ must be strictly increasing for all m and all θ.

Proof of Lemma A. 3:

Suppose that there exists a PBE in which $\psi(m_1) \leq \psi(m_2)$ for $m_1 > m_2$ and that both these signals are in the support of σ. Then all types of the Defender would strictly prefer to send signal m_1. To see this observe that for any type of the Defender who would fight if challenged

$$v_D\psi(m_1) + (1 - \psi(m_1))(v_D p(s) - c_D) - m_1 < v_D\psi(m_2) + (1 - \psi(m_2))(v_D p(s) - c_D) - m_2$$

This is easier to see when the equation is rearranged to be

$$\psi(m_1)(v_D(1 - p(s)) + c_D) - m_1 < \psi(m_2)(v_D(1 - p(s)) + c_D) - m_2$$

as $\psi(m_2) < \psi(m_1)$ and $m_1 > m_2$ by assumption. Similarly for any type of the Defender who would concede if challenged $\psi(m_1) - m_1 < \psi(m_2) - m_2$ by assumption. Therefore, m_1 cannot be in the support of σ. ■

2 Proof of Baseline Results

This section provides proofs of results from the section of the paper titled "Baseline Results"

2.1 Proof of Proposition 1

We will handle each possible of θ in turn. First, let $\theta = v_D$. In this case, it is sufficient to note that both equations (7) and (8) are strictly increasing in v_D. Second, let $\theta = c_D$. In this case, equation (8) is strictly increasing in c_D. This implies that the budget is increasing for each type that prefers to fight rather than concede if challenged ($c_D < \theta$). Conversely, observe that (7) is constant in c_D, implying that the budget for signaling is independent of type for types of the Defender willing to concede. Finally observe that type $\tilde{c}_D$ is indifferent between fighting and not. This implies that the budget for signaling is strictly increasing in c_D for all types $c_D < \tilde{c}_D$ and constant for all types $c_D \geq c_D$. Isomorphic arguments establish the claim for $\theta = s$. ■

2.2 Proof of Additional Claim from Main Text

In the main text, I state that Proposition 1 is sufficient to show that there exist a PBE in which $\theta = v_D$ and signaling occurs (there exists a type v_D for which $\sigma(v_D) \neq 0$).

To see this assume let v'_D denote a type $v'_D \in [\check{v}_D, \tilde{v}_D]$. If the Challenger plays according to σ^* as defined in equation (13), then the Challenger has no incentive to deviate. Additionally, no type of the Defender $v_D < v'_D$ has an incentive to deviate since $B(v'_D|v_D = v'_D) = 0$ and Proposition 1 implies that $B(v'_D)$ must be negative for any additional type. Similarly, no type with a higher valuation has an incentive to deviate to any other signal. This is sufficient to prove the claim.

2.3 Proof of Proposition 2

Let $\theta = c_D$. Following the arguments in the main text, types $c_D \geq \tilde{c}_D$ will be willing to send any signal satisfying condition in equation (9). Similarly, types $c_D < \tilde{c}_D$ will be willing to send any signal satisfying condition (10). Observe that

$$v_D\psi(m) > \psi(m)[v_D - (v_D p(s) - c_D)]$$

thereby implying that any signal that a type $c_D < \tilde{c}_D$ would be willing to send (as compared to sending $m = 0$ and being challenged for sure) a type $c_D \geq \tilde{c}_D$ would strictly prefer to send as well.

For the Challenger to stand down with positive probability in response to a signal $m > 0$, it must be the case that the set of types sending the signal $m > 0$ satisfy

$$v_C \int_{c_D|c_D > \tilde{c}_D, \sigma(\theta)=m'} f(c_D)dc_D + [v_C(1 - p(s)) - c_C] \int_{c_D|c_D \leq \tilde{c}_D, \sigma(\theta)=m'} f(\theta)d\theta \leq 0$$

Suppose that there existed such a signal. Then by the arguments above, a type $c_D < \tilde{c}_D$ would be willing to send such a signal if and only if all types $c_D < \tilde{c}_D$ would strictly prefer to send that signal to $m = 0$. Recall from Lemma A. 1 that some types of the Defender must be sending $m = 0$ and the Challenger must challenge in response. However, if all types $c_D \geq \tilde{c}_D$ are sending a signal $m > 0$, then the Challenger would strictly prefer to stand down upon observing a signal $m = 0$ since only a type $c_D < \tilde{c}_D$ would be willing to send it. Therefore, this cannot be an equilibrium.

Isomorphic arguments establish the claim for $\theta = s$. ■

2.4 Proof of Proposition 3

To show that $B_{SC}^{v_D}$ is increasing in v_D it is sufficient to show that the Defender's expected utility satisfies the single crossing condition (Ashworth and Bueno de Mesquita 2006). First, observe that $B_{SC}^{v_D}$ is increasing in v_D for types $v_D \geq \tilde{v}_D$. For these high-valuation types who intend to fight if challenged, the Defender's expected utility function is given by

$$V(m|v_D) = v_D\psi(m') + (1 - \psi(m'))(v_D p(s) - c_D) - m'$$

which has the following positive cross partial

$$\frac{\partial^2 V(m|v_D)}{\partial m \partial v_D} = \psi'(m)(1 - p(s)) > 0$$

This is sufficient to show that the Defender's strategy is increasing for high-valuation types.

Similarly, $B_{SC}^{v_D}$ is increasing in v_D for types $v_D \leq \tilde{v}_D$. Again, it is sufficient to show that the Defender's expected utility satisfies the single crossing condition. Observe that the Defender's expected utility from a mixed strategy signal is given by

$$V(m|v_D) = v_D\psi(m') - m'$$

which has the positive cross-partial

$$\frac{V(m|v_D)}{\partial m \partial v_D} = \psi'(m') > 0$$

This is sufficient to show that the Defender's strategy is weakly increasing amongst types who are going to concede.

Finally, to show that $B_{SC}^{v_D}$ is increasing I need to show that there can be no signal m_1 that an unresolved type ($v_D \leq \tilde{v}_D$) would prefer to a signal $m_2 < m_1$ but a resolved type ($v_D \geq \tilde{v}_D$) would not. Suppose not, that is suppose there exists a PBE in which two signals m_1, m_2 ($m_1 > m_2$) are sent, that a type $v_D < \tilde{v}_D$ sends signal m_1, and a type $v_D > \tilde{v}_D$ sends signal m_2.

Recall that type $\tilde{v}_D$, the weakest type that is willing to fight, has a payoff to fighting equal to $\tilde{v}_D p(s) - c_D = 0$. This implies that its expected utility for fighting is equal to its expected utility for down after being challenged. Because there exists a type $v_D > \tilde{v}_D$ playing m_2, type $\tilde{v}_D$ the first part of this proof that establishes the single crossing property for types $v_D \geq \tilde{v_D}$ implies that type $\tilde{v}_D$ must strictly prefer signal m_2 to m_1. However, this would imply that $\tilde{v}_D\psi(m_2) - m_2 > \tilde{v}_D\psi(m_1) - m_1$. But then the second part of this proof establishing the single crossing property for types $v_D \leq \tilde{v}_D$ implies that all types $v_D < \tilde{v}_D$ must also strictly prefer to send signal m_2 to m_1, a contradiction.

2.5 Proof of Proposition 4

It is useful to begin this proof by establishing that in any equilibrium there can be at most one signal sent that leads the Challenger to respond with $\psi(\cdot) = 1$. Suppose not, and there were two such unique signals $m_1 \neq m_2$ such that $\psi(m_1) = 1$ and $\psi(m_2) = 1$. Without loss of generality, let $m_1 > m_2$. Any type v_D sending signal m_1 has a strictly profitable deviation to sending m_2 instead since they can increase their expected utility from $v_D - m_1$ to $v_D - m_2$. This contradicts the claim that m_1 is part of a PBE.

Second, observe that any PBE can have the Defender respond by mixing $\psi(m) \in (0, 1)$ to at most one signal. Suppose not. That is, suppose that there exists a PBE in which the Defender sends two signals m_1, m_2 that satisfy $m_1 > m_2$ and $0 < \psi(m_2) < \psi(m_1) < 1$. The Challenger will only be willing to mix in response to a given signal m_i if

$$v_C \int_{v_D | v_D < \tilde{v}_D, \sigma(v_D) = m_i} f(v_D) dv_D + [v_C(1 - p(s)) - c_C] \int_{v_D | v_D \geq \tilde{v}_D, \sigma(v_D) = m_i} f(v_D) dv_D = 0$$

implying that there exist some types $v_D \geq \tilde{v}_D$ who send m_1 and others who send signal m_2. Similarly, different types $v_D \leq \tilde{v}_D$ send m_1 while others send signal m_2. However, this violates Proposition 3. Therefore this cannot be an equilibrium.

We can now begin to show that the proposed equilibrium is the only that survives D1. The proof has three stages. First, I will show that the equilibrium in which all types pool on $m = 0$ does not survive D1. Second, I will show that an equilibrium in which a signal m is sent in response to which the Defender mixes ($\psi(m) \in (0, 1)$) does not survive D1. Finally, I will show that equilibrium described survives D1.

First, consider the case where all types pool on $m = 0$. The Challenger must believe that any deviation to a signal $m > 0$ must be by type $\bar{v}_D$. To see why simply consider a type $v_D < \bar{v}_D$ and consider the response to a signal by the Defender that would make that type indifferent between signaling and not (assuming that such a response exists)

$$v_D \psi(m) + (1 - \psi(m)) \max\{0, v_D p(s) - c_D\} - m = \max\{0, v_D p(s) - c_D\}$$

Per Proposition 3, type $\bar{v}_D$ must strictly prefer the larger signal, demonstrating that it has a profitable deviation to m from a larger set of the Challenger's possible responses. If the Challenger believes that a deviation to m must be by type $\bar{v}_D$, then it will choose not to challenge. Consequently any type with a valuation $v_D > m$ has a strictly profitable deviation to $m + \epsilon$ where $\epsilon > 0$ is arbitrarily small and so pooling on $m = 0$ cannot be an equilibrium.

Second, consider the PBE in which there is a signal m sent by some set of types $[v'_D, v''_D]$ in response to which the Defender plays $\psi(m) \in (0, 1)$. Observe that in this PBE type v''_D is either indifferent between sending signal m and the signal m'' that induces the Challenger not to challenge with probability 1 or that such a signal is not on the equilibrium path for this particular PBE (in which case $v''_D = \bar{v}_D$). Regardless, the Challenger must believe that any deviation to a signal m' satisfying $m' > m$ or $m < m' < m''$ if m'' exists must be by type v''_D. To see this observe that the payoff to this deviation is given by

$$v_D\psi(m') + (1 - \psi(m')) \max\{0, v_D p(s) - c_D\} - m'$$
$$-v_D\psi(m) + (1 - \psi(m)) \max\{0, v_D p(s) - c_D\} - m$$

for any type pooling on m or

$$v_D\psi(m') + (1 - \psi(m')) \max\{0, v_D p(s) - c_D\} - m'$$
$$-v_D\psi(m'') + (1 - \psi(m'')) \max\{0, v_D p(s) - c_D\} - m''$$

for a type pooling on m'' (if such a signal exists). Observe that per Proposition 3 any response by the Challenger that leaves a type $v_D \in [v'_D, v''_D)$ indifferent between sending the signal m and the deviation m' must leave type v''_D strictly preferring signal m'. Similarly, any response by the Challenger that leaves a type $v_D > v''_D$ indifferent between m'' (if such a signal exists) and the deviation m' must leave type v''_D strictly preferring the signal m'. It follows that the Challenger must believe that the devaiting type is v''_D. Type v''_D will fight if challenged. Consequently the Challenger's best response to this devaition will be not to challenge. Consequently, any type sending signal m' or m'' has a strictly profitable deviation to sending signal $m + \epsilon$ for $\epsilon > 0$ arbitrarily small.

Finally, I will show that the conjectured equilibrium survives D1. Consider a deviation to a signal $m' \in (\hat{v}_D, 0)$. The type which profits from that deviation for the largest possible set of responses is type $\hat{v}_D$. First, recall that $\hat{v}_D$ is indifferent between $m = 0$ and $m = \hat{v}_D$. Second, using isomoprohic arguments to those above, it follows from Proposition 1 that type $\hat{v}_D$ can profit from the deviation for a larger set of possible responses than any lower valuation type. Finally, type $\hat{v}_D$ can profit from the deviation for a larger set of possible responses than any type sending signal $m = \hat{v}_D$. To see consider a response to m' that makes $\hat{v}_D$ indifferent between $m = \hat{v}_D$ and m'. Per the single crossing principle, any type with higher valuation must strictly prefer $m = \hat{v}_D$ to m'. It follows that the Challenger must believe that any deviation to a signal m' must be by type $\hat{v}_D$ who will concede if challenged. Consequently, the Challenger plays $\psi(m') = 0$ such that there can be no profitable deviation from $m = 0$ or $m = \hat{v}_D$ to signal m'. ■

3 Proof of Signaling with Differential Costs Results

This section provides proofs of the results from the section "Signaling with Differential Costs"

3.1 Proof of Proposition 5

As in the proof of Proposition 3, to show that B_{SC} is increasing in s and decreasing in c_D it is only necessary to show that the Defender's expected utility satisfies single crossing.

Consider first the case where $\theta = c_D$. The expected utility for a type of the Defender who lacks the resolve to fight ($c_D > \tilde{c}_D$) is given by

$$V(m|c_D) = \psi(m)v_D - e(m, c_D)$$

The cross-partial of this equation is given by

$$\frac{\partial^2 V(m|c_D)}{\partial m \partial c_D} = -\frac{\partial^2 e(m, c_D)}{\partial m \partial c_D} < 0$$

which is negative by assumption. It follows that the Defender's Budget for signaling must be strictly decreasing for unresolved types $\theta = c_D$. The proof for demonstrating that the Defender's budget for signaling is decreasing in strength ($\theta = s$) for types too weak to fight ($s < \tilde{s}$) is isomorphic.

Continue to consider the case where $\theta = c_D$. The expected utility for a type of the Defender who is sufficiently resolved to fight ($c_D \leq \tilde{c}_D$) is given by

$$V(m|c_D) = \psi(m)v_D + (1 - \psi(m))(v_D p(s) - c_D) - e(m, c_D)$$

The cross-partial of this equation is given by

$$\frac{\partial^2 V(m|c_D)}{\partial m \partial c_D} = \psi'(m) - \frac{\partial^2 e(m, c_D)}{\partial m \partial c_D}$$

This is indeterminate - ψ' is positive while the cross-partial is positive by assumption meaning that the value of this equation is as of yet unknown.

The function ψ is an equilibrium object. Taking the first-order condition for a resolved type ($c_D \leq \tilde{c}_D$) with respect to m, we obtain that

$$\psi'(m)(v_D - p(s) + c_D) - \frac{e(m, c_D)}{\partial m} = 0$$

or that

$$\psi'(m) = \frac{\frac{\partial e(m,c_D)}{\partial m}}{v_D - p(s) + c_D}$$

As the first-order condition, this value of $\psi'(m)$ represents the rate of increase in $\psi'(m)$ at which the marginal costs of signaling equal the marginal benefits. Substituting this value into our previous inequality, we obtain that the inequality is decreasing whenever the inequality in equation (18) holds. This implies that as c_D increases, the additional effort required to marginally increase the signal exceed the marginal benefits.

All that remains to complete the proof for the $\theta = c_D$ case is to show that the first-order condition represents a maximum. This will be true whenever the Defender's expected utility function is concave.

$$\psi''(m)(v_D - p(s) + c_D) - \frac{\partial^2 e(m, c_D)}{\partial^2 m} < 0$$

This inequality holds whenever inequality (18) holds. To see that this is true, differentiate inequality (18) by m and then integrate by c_D. This produces

$$\psi''(m)(c_D) + C < \frac{\partial^2 e(m, c_D)}{\partial^2 m}$$

where C is a constant of integration. This is sufficient to demonstrate that the second order condition does indeed constitute a maximum.

Now consider the case where $\theta = s$. The cross-partial for the expected utility function for a type strong enough to fight ($s > \tilde{s}$) is given by

$$V(m|s) = -\psi'(m)v_D \frac{dp(s)}{ds} - \frac{\partial^2 e(m, s)}{\partial m \partial s}$$

This is again indeterminate as the function ψ' is a positive equilibrium object and the cross-partial of the cost-function e is negative by assumption.

The function ψ is an equilibrium object. Taking the first-order condition for a strong type ($s > \tilde{s}$) with respect to m. Taking the first-order condition for a resolved type ($s \geq \tilde{s}$) with respect to m, we once again obtain that

$$\psi'(m) = \frac{\frac{\partial e(m,s)}{\partial m}}{v_D - p(s) + c_D}$$

As the first-order condition, this value of $\psi'(m)$ represents the rate of increase in $\psi'(m)$ at which the marginal costs of signaling equal the marginal benefits. Substituting this value into our previous inequality we obtain equation (19). This implies that as s increases, budgets will increase in s so long as discount on additional signaling exceeds the marginal costs of additional signaling.

Once again, all that remains to complete the proof is to show that the first-order condition represents a maximum. This will be true whenever the Defender's expected utility function is concave.

$$\psi''(m)(v_D - p(s) + c_D) - \frac{\partial^2 e(m,s)}{\partial^2 m} < 0$$

This inequality holds whenever inequality (19) holds. To see that this is true, differentiate inequality (19) by m and then integrate by s. This produces

$$\psi''(m)p(s)v_D + C > \frac{\partial^2 e(m, c_D)}{\partial^2 m}$$

where C is a constant of integration. This is sufficient to demonstrate that the second order condition does indeed constitute a maximum. ■

3.2 Proof of Proposition 6

Follows isomorprhic arguments to those in Proposition 4. ■

4 Proof of Arming Results

This section presents proofs for the results from the section titled "Sunk Costs as Investments in Arming"

4.1 Lemma when Signaling Constitutes Arming

Before proceeding to the proofs of Propositions 8 and 9, it is useful to reprove Lemma A. 3 when signaling also constitutes arming.

Lemma A. 4
Let $\theta \in \{v_D, c_D\}$. Then $\psi(m)$ must be strictly increasing for all m. Otherwise $\psi(m)$ must be weakly increasing.

Proof: To prove the claim, I will consider each possible value of θ in turn. First, consider the case where $\theta = v_D$. Suppose that $\psi(m_1) \leq \psi(m_2)$ for $m_1 > m_2$ in a PBE. Note that any type with a sufficiently low valuation post-arming ($v_D < \tilde{v}_D(m_1)$) strictly prefers to send signal m_2 because $\psi(m_1)v_D - m_1 < \psi(m_2)v_D - m_2$. It follows that the only type of Defender who would send signal m_1 is a Defender with a valuation high enough to fight if challenged ($v_D \geq \tilde{v}_D(m_1)$). But then, the Challenger would strictly prefer to concede to any type sending m_1, thereby implying that $\psi(m_1) = 1$, a contradiction. The claim for the case where $\theta = c_D$ can be proven with isomorphic arguments.

Consider the case where $\theta = s$. Suppose that $\psi(m_1) \leq \psi(m_2)$ for $m_1 > m_2$ in a PBE. Following the logic for the case where $\theta = v_D$, any type too weak to fight ($s < \tilde{s}(m_1)$), must strictly prefer signal m_2. However, unlike the previous cases (when $\theta = \{v_D, c_D\}$), when $\theta = s$ the Challenger will be willing to challenge a weak Defender who will fight with certainty. Therefore, the arguments for the previous cases cannot be used again.

To prove the claim, I will therefore first establish that the Defender's expected utility function satisfies increasing differences when ψ is decreasing. Recall, that the utility for a type that would rather fight than concede (post-arming) is given by

$$U(m|\theta = s, s \geq \tilde{s}_D(m)) = \psi(m)v_D + (1 - \psi(m))(v_D p(s, m) - c_D) - m$$

which has the following cross partial

$$\frac{\partial^2 U(m|\theta = s, s \geq \tilde{s}(m))}{\partial m \partial s} = -\psi'(m)v_D \frac{\partial^2 p(s, m)}{\partial s} + (1 - \psi(m))v_D \frac{\partial^2 p(s, m)}{\partial s \partial m}$$

I have assumed that the cross-partial $\frac{\partial^2 p(s,m)}{\partial s \partial m} = 0$. Therefore, if $\psi'(m)$ is negative, then the Defender's expected utility function satisfies increasing differences, implying that stronger types of the Defender strictly prefer to send strictly larger signals.

Consider two signals m_1 and m_2 where $m_1 > m_2$. If $\psi(m_1) < \psi(m_2)$, then stronger types of the Defender prefer the larger signal m_1. Next, observe that for $\psi(m_1)$ to be strictly smaller than $\psi(m_2)$, $\psi(m_2)$ must be positive. For $\psi(m_2)$ to be positive, the Challenger's expected utility for issuing a challenge must be negative or zero:

$$Pr(s < \tilde{s}(m)|m = m_2)v_C + Pr(s > \tilde{s}(m)|m = m_2)v_C \mathbb{E}[p(s, m_2)] \leq 0$$

Now consider the Challenger's expected utility for challenging upon observing m_1. Because the Defender's optimal strategy must be weakly increasing in type, types that send signal m_1 must be strictly stronger than any type sending signal m_2 and have armed more. This implies that they will surely fight if challenged and that $\mathbb{E}[p(s, m_1)] > \mathbb{E}[p(s, m_2)]$, thereby implying that the Challenger's expected utility for issuing a challenge must be strictly negative. It follows that $\psi(m_1) = 1$. A contradiction of the premise. This is sufficient to demonstrate that $\psi(m)$ must be weakly increasing.

4.2 Proof of Proposition 8

Following the arguments in Ashworth and Bueno de Mesquita (2006), to check and see whether the Defender's expected utility function satisfies increasing (decreasing) differences, and subsequently single-crossing, all that is necessary is to check and see whether it has a positive (negative) cross-partial whenever $\theta \in \{v_D, c_D\}$ ($\theta = s$).

Let $\theta = v_D$. For a type of the Defender that would rather concede than fight (post-arming), their expected utility for a given signal m is given by

$$U(m|\theta = v_D, v_D < \tilde{v}_D(m)) = \psi(m)v_D - m$$

has the cross partial

$$\frac{\partial^2 U(m|\theta = v_D, v_D < \tilde{v}_D(m))}{\partial v_D \partial m} = \psi'(m)$$

which is positive as desired. Similarly, the expected utility for a type that has a sufficiently high-valuation such that they prefer to fight than concede (post-arming) is given by

$$U(m|\theta = v_D, v_D \geq \tilde{v}_D(m)) = \psi(m)v_D + (1 - \psi(m))(v_D p(s, m) - c_D) - m$$

This has the following cross partial

$$\frac{\partial^2 U(m|\theta = v_D, v_D \geq \tilde{v}_D(m))}{\partial m \partial v_D} = \psi'(m)(1 - p(s, m)) + (1 - \psi(m))\frac{\partial p(s, m)}{\partial m}$$

which is positive given Lemma A. 4 and the assumption that $p(s, m)$ is increasing in m.

For a type that prefers to concede at the low level of arming but fight at the higher level of arming $v_D \in [\tilde{v}_D(m_1), \tilde{v}_D(m_2)]$, the difference in utility between arming and not is given by

$$\begin{aligned} U(m_1, m_2|\theta = v_D; \tilde{v}_D(m_2) > v_D > \tilde{v}_D(m_1)) = \psi(m_1)v_D \\ +(1 - \psi(m_1))(p(s, m)v_D - c_D) - \psi(m_2)v_D + m_2 \end{aligned}$$

Note that this can be rewritten as

$$U(m_1, m_2|\theta = v_D; \tilde{v}_D(m_2) > v_D > \tilde{v}_D(m_1)) = \psi(m_1)v_D + (1 - \psi(m_1))(p(s, m)v_D - c_D)$$
$$-\psi(\tilde{m})v_D + (1 - \psi(\tilde{m}))(p(s, \tilde{m})v_D - c_D) + \psi(\tilde{m})v_D + \tilde{m} - \psi(m_2)v_D + m_2$$

where $\tilde{m}$ is the level of arming that leaves the Defender indifferent between fighting and not. Observe that the first two terms must be increasing in v_D because we have already shown that the utility function satisfies increasing differences for types who are sufficiently high valuation to fight for both level of arms. Similarly, the second two terms are increasing because we have already shown that the utility function satisfies increasing differences for types who would rather concede than fight (if challenged). The sum of two increasing functions is increasing. Therefore it must be the case that the (27) is increasing for types $v_D \in [\tilde{v}_D(m_1), \tilde{v}_D(m_2)]$. This completes part (i) of the proof.

Now let $\theta = c_D$. For an unresolved type of the Defender who would rather concede than fight (post-arming), their expected utility for a given signal m is given by

$$U(m|\theta = c_D, c_D > \tilde{c}_D(m)) = \psi(m)v_D - m$$

which has cross-partial

$$\frac{\partial^2 U(m|\theta = c_D, c_D > \tilde{c}_D(m))}{\partial c_D \partial m} = 0$$

which is neither positive or negative as desired. Similarly, the expected utility for a resolved type that would rather fight than concede (post-arming) is given by

$$U(m|\theta = c_D, c_D \leq \tilde{c}_D(m)) = \psi(m)v_D + (1 - \psi(m))(v_D p(s, m) - c_D) - m$$

which has the following cross partial

$$\frac{\partial^2 U(m|\theta = c_D, c_D \leq \tilde{c}_D(m))}{\partial m \partial c_D} = \psi'(m)$$

This is positive per Lemma A. 4.

For a type that prefers to concede at the low level of arming but fight at the higher level of arming $c_D \in [\tilde{c}_D(m_1), \tilde{c}_D(m_2)]$, the difference in utility between arming and not is given by

$$U(m_1, m_2|\theta = c_D; \tilde{c}_D(m_2) < c_D < \tilde{c}_D(m_1)) = \psi(m_1)v_D \\ +(1-\psi(m_1))(p(s,m)v_D - c_D) - \psi(m_2)v_D + m_2$$

Note that this can be rewritten as

$$U(m_1, m_2|\theta = c_D; \tilde{c}_D(m_2) > c_D > \tilde{c}_D(m_1)) = \psi(m_1)v_D + (1-\psi(m_1))(p(s,m)v_D - c_D) \\ -\psi(\tilde{m})v_D + (1-\psi(\tilde{m}))(p(s,\tilde{m})v_D - c_D) + \psi(\tilde{m})v_D + \tilde{m} - \psi(m_2)v_D + m_2$$

where $\tilde{m}$ is again the level of arming that leaves the Defender indifferent between fighting and not. Observe that the first two terms must be increasing in c_D because we have already shown that the utility function satisfies increasing differences for types who are sufficiently high valuation to fight for both level of arms. The second two terms are constant because we have already shown that the budget equation (27) is constant for types who would rather concede than fight (if challenged). The sum of a constant and an increasing function is increasing. Therefore it must be the case that the (27) is increasing for types $c_D \in [\tilde{c}_D(m_1), \tilde{c}_D(m_2)]$. This completes part (ii) of the proof.

Finally, let $\theta = s$. For a weak type of the Defender who would rather concede than fight (post-arming), their expected utility for a given signal m is given by

$$U(m|\theta = s, s < \tilde{s}(m)) = \psi(m)v_D - m$$

which has the neither negative nor positive cross partial

$$\frac{\partial^2 U(m|\theta = s, s < \tilde{s}(m))}{\partial s \partial m} = 0$$

as desired. Similarly, following the calculations performed in the proof of Lemma A. 4, a type sufficiently strong enough to fight(post-arming) is given by

$$\frac{\partial^2 U(m|\theta = s, s \geq \tilde{s}(m))}{\partial m \partial s} = -\psi'(m)v_D \frac{\partial p(s,m)}{\partial s}$$

Per Lemma A. 4 $\psi'(m)$ is nonnegative. It follows that the Defender's expected utility function satisfies single crossing and is strictly decreasing whenever $\psi'(m)$ is positive and is constant whenver $\psi(m)$ is constant. Finally, the proof that (27) is decreasing for types $s \in [\tilde{s}(m_1), \tilde{s}(m_2)]$ is isomorphic to the $c_D \in [\tilde{c}_D(m_1), \tilde{c}_D(m_2)]$ case when $\theta = c_D$. This completes the proof. ■

■

4.3 Proof of Proposition 9

4.3.1 Part 1: Proving the prescribed strategies are a PBE when $\theta = s$:

Consider the case where $\theta = s$. Observe that the Challenger is indifferent in response to any level of arming chosen by the Defender on the path of play. Therefore, to prove that the postulated strategies constitute a PBE there are three steps. First, it is necessary to show that no type $s \in [\underline{s}, \rho)$ have a profitable deviation. Second, it is necessary to show that $\psi(0) < \psi(\sigma(\rho))$ to ensure that the decreasing differences property establish in Proposition 8 applies. Third, it is necessary to show that for types $s \in [\rho, \bar{s}]$ there is no profitable deviation from pooling on $m = 0$ and that no type sending a signal $m > 0$ has a profitable deviation to $m = 0$.

Step 1: No type $s \in [\underline{s}, \rho)$ has a profitable deviation. To prove the first part it is necessary to solve for the Defender's first-order condition and demonstrate that it constitutes a maximum. The first-order condition is given by

$$\frac{\partial U(m|s \leq \rho, \theta = s)}{\partial s} = \psi'(m)v_D - \psi'(m)(v_D p(s,m)) - c_D) + (1 - \psi(m))\frac{\partial p(s,m)}{\partial m} - 1 = 0$$

which can be rearranged into the strategy $\psi(m)$ as stated in the Proposition. To demonstrates that this constitutes a maximum, it is sufficient to show that the second-order condition is negative. The second-order condition is given by

$$\begin{aligned}\frac{\partial^2 U(m|s \leq \rho, \theta = s)}{\partial^2 m} &= \psi''(m)(v_D - v_D(p(s,m) + c_D) \\ &-2v_D\psi'(m)\frac{\partial p(s,m)}{\partial m} + (1 - \psi(m))\frac{\partial^2 p(s,m)}{\partial^2 m}\end{aligned}$$

we cannot make a determination as to whether this expression is negative or not without more information regarding $\psi''(m)$.

Before proceeding it is useful to establish the following identity. For separating types $s \in [\underline{s}, \rho)$ it is useful to remember that their level of arming must always satisfy.

$$v_C(1 - p(s,m)) - c_C = 0$$

This implies that for all separating types

$$p(\sigma^{-1}(m), m) = 1 - \frac{c_C}{v_C}$$

Taking the derivative of this function with respect to m, we arrive at the following useful identity

$$\frac{\partial p(s,m)}{\partial s}\frac{1}{d\sigma(s)}+\frac{\partial p(s,m)}{\partial m}=0$$

Now observe that to obtain the derivative of $\psi''(m)$ we can retake the derivative of $\psi'(m)$, obtained for the first-order condition while allowing s to be a function of m as given by $\sigma^{-1}(m)$. That is, I am taking the derivative of the strategy function $\psi'(m)$ with respect to m as opposed to the Defender's expected utility function. Doing so we find that

$$\psi'(m)(v_D(1-p(s,m))+c_D)=1-(1-\psi(m))\frac{\partial p(\sigma^{-1}(m),m)}{\partial m}$$

becomes

$$\begin{aligned}\psi''(m)(v_D(1-p(s,m))+c_D)-\psi'(m)v_D\left[\frac{\partial p(s,m)}{\partial s}\frac{1}{d\sigma(s)}+\frac{\partial p(s,m)}{\partial m}\right]\\ =\psi(m)v_D\frac{\partial p(s,m)}{\partial m}-(1-\psi(m))v_D\left[\frac{\partial^2 p(s,m)}{\partial s\partial m}\frac{1}{d\sigma(s)}+\frac{\partial^2 p(s,m)}{\partial^2 m}\right]\end{aligned}$$

Per the identity obtained above, the terms in the square brackets on the left-hand side of the equation are equal to zero. The cross-partial term $\frac{\partial^2 p(s,m)}{\partial s\partial m}$ is also equal to zero by assumption. Therefore, substituting the value of $\psi''(m)$ into the second order condition we find that we are left with

$$\frac{\partial^2 U(m|s\le\rho,\theta=s)}{\partial^2 m}=-\psi'(m)v_D\frac{\partial p(s,m)}{\partial m}$$

which is negative as desired. This demonstrates that no type $s\in[\underline{s},\rho)$ has any incentive to deviate.

Step 2: $\psi(0)<\psi(\sigma(\rho))$: To apply the single-crossing property it is necessary to demonstrate that $\psi(m^*(\rho))>\psi(0)$ where $m^*(\rho):=p^{-1}(\rho,1-\frac{c_C}{v_C})$, the amount that type ρ would arm to make the Challenger indifferent between challenging it and not. This will occur whenever

$$\psi(0)>\frac{\psi(m^*(\rho))(v_D(1-p(\rho,m^*(\rho)))+c_D)+v_D(p(\rho,m^*(\rho))-p(\rho,0))-m}{v_D(1-p(\rho,0))+c_D}$$

which simplifies into the following condition

$$m>(1-\psi(m^*(\rho)))v_D(p(\rho,m^*(\rho))-p(\rho,0))$$

Note that when $m = 0$, both sides of the inequality are equal to zero. Therefore, to show that the inequality holds it is sufficient to show that the derivative of the right-hand side with respect to m is strictly less than 1.

The derivative of the right-hand side of the inequality is given by

$$-\psi'(m)v_D(p(\rho, m) - p(\rho, 0)) + (1 - \psi(m))v_D\frac{\partial p(\rho, m)}{\partial m}$$

Per equation (29),

$$\psi'(m) = \left[1 - v_D\frac{dp(s, m)}{dm}(1 - \psi(m))\right]\left[\frac{1}{v_D(1 - p(\sigma(s), m)) + c_D}\right]$$

It follows that the derivative of the right-hand side of our original inequality is strictly less than 1 whenever

$$1 > -\left[\frac{1 - v_D\frac{dp(s,m)}{dm}(1 - \psi(m))}{v_D(1 - p(\sigma(s), m)) + c_D}\right]v_D(p(\rho, m) - p(\rho, 0)) + (1 - \psi(m))v_D\frac{\partial p(\rho, m)}{\partial m}$$

Rearranging we have

$$1 - (1 - \psi(m))v_D\frac{\partial p(\rho, m)}{\partial m} > -\left[\frac{1 - v_D\frac{dp(s,m)}{dm}(1 - \psi(m))}{v_D(1 - p(\sigma(s), m)) + c_D}\right]v_D(p(\rho, m) - p(\rho, 0))$$

Note that $\frac{\partial p(s,m)}{\partial m}$ is independent of s by assumption. It follows that we can divide both sides by $1 - (1 - \psi(m))v_D\frac{\partial p(\rho,m)}{\partial m}$ and have the terms cancel out.[1] This leaves us with

$$1 > -\frac{v_D(p(\rho, m) - p(\rho, 0))}{v_D(1 - p(\sigma(s), m)) + c_D}$$

Observe that the right-hand side is negative, while the left-hand side is positive. This demonstrates $\psi(m) > \psi(0)$.

Step 3: No type pooling on $m = 0$ has a profitable deviation and no type $s > \rho$ has a profitable deviation to $m = 0$: To complete the proof, recall that $\psi(0)$ has been constructed so as make type ρ indifferent between pooling on $m = 0$ and separating by sending $m = p^{-1}(\rho, 1 - \frac{c_C}{v_C})$. Therefore, type ρ has no incentive to deviate. Then, per Lemma 8, weaker types of the Defender experience decreasing differences for larger signals when these lead to a greater probability of concession. It follows that if type ρ is indifferent between $m = 0$ and $m = p^{-1}(\rho, 1 - \frac{c_C}{v_C})$, then no type $s > \rho$ has an incentive to deviate from pooling on

[1] Note that this quantity is positive as $\psi'(m)$ is positive.

$m = 0$ to any other on-path signal. Similarly, observe that Lemma 8 also implies that no separating type $s < \rho$ has any incentive to deviate from their prescribed strategy to $m = 0$. This suffices to show that the strategies and beliefs described in the Proposition constitute a PBE when $\theta = s$.

4.3.2 Part 2: Proving the prescribed strategies are a PBE when $\theta = c_D$

Observe that an unresolved type of the Defender who intends to concede (post-arming) has an expected utility of signaling given by

$$U(m|c_D > \tilde{c}_D(m), \theta = c_D) = \psi(m)v_D - m$$

This expected utility is independent of the Defender's type so long as it intends to concede if challenged ($c_D > \tilde{c}_D(m)$). Such an unresolved Defender is indifferent between two choices m_1 and m_2 whenever $\psi(m)$ increases at the constant rate

$$\psi'(m) = \frac{1}{v_D}$$

as m increases. It follows that any type $c_D > \tilde{c}_D$ is indifferent with respect to any signal m that satisfies $m \geq p^{-1}\left(s, \frac{c_D}{v_D}\right)$.

Now consider types of the Defender who prefer to fight given a particular level of arming $c_D \leq \tilde{c}_D(m)$. Per Proposition 8, the Defender's expected utility function satisfies the single crossing principle. This implies that if an unresolved type ($c_D > \tilde{c}_D(m)$) is indifferent between two signals m_1 and m_2 satisfying $m_1 > m_2$, then a resolved type ($c_D \leq \tilde{c}_D(m_1)$) must strictly prefer to send signal m_2. It follows that any type $c_D \in (\tilde{c}_D, \bar{c}_D)$ will strictly prefer to send signal $p^{-1}\left(\frac{c_D}{v_D}\right)$ over any larger signal and that types in the range $c_D \in (\underline{c}_D, \tilde{c}_D]$ will strictly prefer to pool on $m = 0$ over any larger signal. This suffices to show that no type of the Defender has any incentive to deviate.

As for the Challenger's response to being challenged, note that the Challenger is indifferent between challenging and not when $m = 0$, because

$$F(\tilde{c}_D)((1 - p(s, m))v_D - c_C) + (1 - F(\hat{c}))v_C = 0$$

To ensure that the Challenger is indifferent between challenging and not when $m > 0$, the Defender must choose not to fight with positive probability. Recall, that any type of the Defender that adopts a signal $m > 0$ is indifferent between fighting and not when challenged.

The Challenger is indifferent between challenging and not whenever

$$\phi v_C + (1-\phi)(v_c(1-p(s,m)) - c_C) = 0$$

which can be rewritten as a function of m

$$\phi(m) := -\frac{v_c(1-p(s,m)) - c_C}{p(s,m) + c_C}$$

It follows that when the Defender plays accordingly, then the Challenger has no incentive to deviate from its prescribed strategy in response to any signal m.

4.3.3 Step 3: The PBE Described in the Proposition is Unique Subject to the D1 Criterion ($\theta = c_D$):

Let $\theta = c_D$. In this case, any alternative PBE must either feature (i) pooling on some signal $m \neq 0$, (ii) pooling by additional types on $m = 0$ or (iii) an alternative equilibrium in which all types who don't pool on $m = 0$ separate. This proof will demonstrate that cases (i) and (ii) do not survive D1 and that in case (iii) of the equilibria that satisfy D1 that described in the proposition delivers the sender with their highest expected utility.

Case (i): Consider an alternative PBE with pooling on some alternative signal $\hat{m} \neq 0$. By the single-crossing property the set of types sending signal $\hat{m}$ must form a connected interval $[c'_D, c''_D]$ (with $c'_D < c''_D$).

The Challenger must respond to $\hat{m}$ with a response $\psi(\hat{m}) > 0$. Suppose not. If the Challenger responded by playing $\psi(\hat{m}) = 0$, then no unresolved type of the Defender ($c_D > \tilde{c}_D(\hat{m})$) would play $\hat{m}$ since they would be strictly better off playing $m = 0$. However, if only resolved types ($c_D \leq \tilde{c}_D(\hat{m})$ send signal $\hat{m}$, then the Challenger would strictly prefer to concede. It follows that the Challenger must either be indifferent between challenging or not or strictly prefer not to challenge upon observing $\hat{m}$. Note that this implies that type c'_D must strictly prefer to fight after paying signaling cost $\hat{m}$.

Let m' denote the supremum of on-the-equilibrium-path signals $m < \hat{m}$ if one exists. Observe that it must either be the case that (i) the boundary type c'_D is indifferent between sending m' and $\hat{m}$ or that (ii) no signal smaller than $\hat{m}$ is sent with positive probability as part of the PBE (m' does not exist). Suppose not. First, suppose not because the boundary type c'_D strictly prefers $\hat{m}$ to m'. Observe that the Defender's expected utility is continuous in c_D. It follows that if the former were true, then there exist some type $c'_D - \epsilon$ for $\epsilon > 0$ arbitrarily small that must also strictly prefer sending signal $\hat{m}$ to m'. But this contradicts the premise that c'_D is the lowest cost-of-fighting type to demand $\hat{m}$. This is sufficient demonstrate that

c_D' must be indifferent between $\hat{m}$ and m' or that m' does not exist.

Recall that D1 requires that the Challenger believe that any deviation to an off-path signal be by the type which benefits from the largest set of the Defender's possible responses. Let m' denote the supremum of signals $m < \hat{m}$ of on-path signals if one exists. Observe that type c_D' or a strictly lower-cost-of-fighting type strictly benefit from more responses to signals $m \in (m', \hat{m})$ if m' exists or $[0, \hat{m})$ if m' does not exist than any higher cost-of-fighting type. To see this simply observe that per the single crossing principle, any response by the Challenger that leaves a type of the Defender $c_D > c_D'$ indifferent between m and $\hat{m}$ must imply that c_D' strictly prefers m'. Now consider lower cost-of-fighting types. If m' does not exist, then $c_D' = \underline{c}_D$ and the proof is complete. On the other hand if, m' exists, then there are two sub-cases to consider. First, is the case where m' is another signal pooled upon by another subset of signalers. In this case, observe that all types in the set $c_D \in [\tilde{c}_D(m'), c_D']$ will concede if challenged in response to m'. Subsequently they all must be indifferent between m' and $\hat{m}$. In this case, the signaling crossing principle determines that it must be type $\tilde{c}_D(m')$ who benefits from the largest possible set of responses to any compared to any other signal. To see this observe that any signal that makes type $\tilde{c}_D(m')$ indifferent between m and $\hat{m}$ must leave any higher cost-of-fighting type strictly preferring $\hat{m}$. Second, it could be that on-path signals in the neighborhood of $(m' - \epsilon, m)$ for $\epsilon > 0$ are each only sent by a single type c_D and are perfectly separating. In this latter case, we know that type c_D' is indifferent between m' and $\hat{m}$ and that all lower cost-of-fighting types must strictly prefer to send some signal $m'' < m'$. Once again the single crossing principle informs us that any response by the Challenger that leaves type c_D' of the Defender indifferent between m' and m must have lower cost-of-fighting types strictly prefer signal m'. It follows that the Challenger must believe that the Defender is type c_D'.

This suffices to show that the Challenger must believe that any deviation to an off-path signal $m \in (m', \hat{m})$ if m' exists or to a signal $m \in (0, \hat{m})$ otherwise must either be by type c_D' or a lower cost-of-fighting type. Now consider a deviation to a signal $\hat{m} - \epsilon$. Recall that type c_D' must strictly prefer to fight after paying signaling cost $\hat{m}$, thereby implying that there exists an $\epsilon > 0$ such that after paying $\hat{m} - \epsilon$, type c_D' will also strictly prefer to fight. In response to such a deviation, the Challenger must believe that the Defender will fight if challenged and so strictly prefers to concede. This is a profitable deviation for any type sending signal $\hat{m}$ regardless of the precise value of $\psi(\hat{m})$. It follows that there can be no PBE in which the Defender pools on a signal $m \neq 0$ that survives D1.

Case (ii): All types pool on $m = 0$. Suppose that all types pooled on $m = 0$. In this case, the Challenger fights for sure and the expected utility for an unresolved type of the Defender is 0 since they will concede if challenged. However, this cannot be an equilibrium.

By assumption $v_D - p^{-1}\left(s, \frac{\underline{c}_D}{v_D}\right) > 0$ so that if the highest cost-of-fighting type armed to the level that would leave it indifferent between fighting and not and the Challenger conceded with certainty, then it would still have positive utility after accounting for the costs of arming. In this case, the PBE where all types cannot be an equilibrium. Any unresolved type of the Defender ($c_D > \tilde{c}_D(0)$) could deviate to $m = p^{-1}\left(s, \frac{\underline{c}_D}{v_D}\right) + \epsilon$ for $\epsilon > 0$ greater than zero and the Challenger would strictly prefer to concede than fight regardless of their belief in the Defender's type. In which case the Defender's payoff from the deviation is $v_D - p^{-1}\left(s, \frac{\underline{c}_D}{v_D}\right) - \epsilon > 0$.

Case (iii): Let $\bar{C}$ denote the set of unresolved types pooling on $m = 0$. While the proposition considers a case where $\bar{C} = [\underline{c}_D, \tilde{c_D}] \cup [\check{c}_D, \bar{c}_D]$, an alternative equilibrium PBE could alter this set of types so long as

$$v_C \int_{c_D \in \bar{C}, c_D > \tilde{c}_D(0)} \iota(c_D) f(c_D) dc_D + \int_{c_D \in \bar{C}, c_D \leq \tilde{c}_D(0)} f(c_D) dc_D [(1 - p(s, 0)) v_C - c_C] = 0 \quad \text{(A. 2)}$$

where ι denotes a mixed strategy function describing the probability with which an unresolved type of the Defender sends $m = 0$ (and sending the signal $m = p^{-1}\left(s, \frac{c_D}{v_D}\right)$ with the reciprocal probability). This ensures that the Challenger is still willing to mix between challenging and not after observing $m = 0$.

Step (i): Characterizing the alternative PBE: There are several facts worth considering in such a PBE. First, absent any pooling any on-the-equilibrium-path level of arming $m > 0$ can only be sent by a single type satisfying $m = p^{-1}\left(s, \frac{c_D}{v_D}\right)$. It follows that the Challenger has no incentive to deviate from its strategy so long as the Defender responds to being challenged by playing $\phi(m)$ as defined in equation (31) after arming to a level $m > 0$. The Defender would always be willing to play $\phi(m)$ since any separating type must be indifferent between fighting and not after selecting a positive level of arming. Second, per the single crossing principle, any type $c_D \leq \tilde{c}_D$ must strictly prefer to pool on $m = 0$ such that no type of the Defender has an incentive to deviate. Third, following the logic of part 2 of the proof in this proposition, any type c_D is indifferent between any level of arming on the equilibrium path satisfying $m \leq p^{-1}\left(s, \frac{c_D}{v_D}\right)$ but strictly prefers this level of arms to any $m > p^{-1}\left(s, \frac{c_D}{v_D}\right)$. This implies that any type $c_D > \tilde{c}_D$ pooling on $m = 0$ nor any type sending a signal $m > 0$ have an incentive to deviate to any other on-path signal.

Step (ii): Such an alternative equilibrium survives D1: If there is no pooling on any signal $m > 0$, than any set of types $\bar{C}$ pooling on $m = 0$ besides that described in the proposition must leave some $m \in \left(0, p^{-1}\left(s, \frac{\check{c}_D}{v_D}\right)\right)$ that is off-the equilibrium path. Consider a deviation to the off-path signal $m' \in \left(0, p^{-1}\left(s, \frac{\check{c}_D}{v_D}\right)\right)$. Let m_i and m_s denote the infimum and supremum of signals respectively in the image of $\sigma(m)$ satisfying $m_s \leq m$ and $m_i \geq m$.

Observe that the Challenger must believe that types $c_D > \tilde{c}_D(m')$ are those that profit from a deviation to the signal m' for the greatest set of possible responses. To see this observe that this set of types is indifferent between sending signal m' and signal $m'' = \max\{m_i, m' + \epsilon\}$ for ϵ arbitrarily small so long as

$$\psi(m') = \psi(m'') + \frac{m' - m''}{v_D}.$$

Decreasing differences implies that any signal that any type with a lower cost of fighting than $\tilde{c}_D(m)$ must strictly prefer signal m''. In response to a deviation m', the Challenger must therefore have posterior beliefs

$$G(m'|\theta = c_D) = \begin{cases} 0 & \text{if } c_D < \tilde{c}_D(m') \\ \frac{f(c_D)}{1-F(\tilde{c}_D(m'))} & \text{otherwise} \end{cases}$$

and an expectation that the Defender will concede if challenged.[2] It therefore chooses to fight with certainty thereby implying that this is not a profitable deviation for any type to such a signal m'.

Step iii: Of the equilibria that survive D1, that described in the Proposition leads to a higher payoff for every type of the Defender. Consider an equilibrium of the type described in step (i) in which $\bar{C}$, the set of types pooling on $m = 0$, is different from that described in the proposition. The largest on-the equilibrium path signal in this equilibrium must be larger than that sent by $\check{c}_D$. This is because sustaining the equality in equation (A. 2) while also having types $c_D \in (\check{c}_D, \bar{c}_D]$ not play $\iota(c_D) = 1$ requires that any type not playing $\iota(c_D) = 1$ separate with positive probability.

Let $\bar{m}$ denote the supremum of on-the-path signals in a particular PBE that survives D1. Observe that the Challenger's concession strategy $\psi(m)$ is constructed such that $\psi(\bar{m}) = 1$ and is then equal to $\psi(m) = 1 - \frac{\bar{m}-m}{v_D}$ for any other signal. Across equilibria, any type separating is sending the same signal $m = p^{-1}\left(s, \frac{c_D}{v_D}\right)$. The probability that such a signal leads to a concession is highest when $\bar{m}$ is small, i.e. when the highest cost signaling type to send a positive signal is minimized. Therefore, across equilibria that survive D1, the equilibrium that delivers the highest expected utility for separating is that described in the Proposition. Similarly, consider the expected utility for pooling on zero. For types $c_D \leq \tilde{c}_D$ who pool on $m = 0$ across all equilibria that survive D1, the equilibria which maximizes their probability of obtaining a concession (and consequently delivers the highest expected utility) is the one described in the proposition.

[2]Technically type $\tilde{c}_D(m')$ can fight if challenged, but has mass zero.

Finally, consider the expected utility for types $c_D > \tilde{c}_D$ who may pool on $m = 0$ or separate with signal $p^{-1}\left(s, \frac{c_D}{v_D}\right)$ in different PBE that survive D1. Such types must either be (i) indifferent across all on-path signals in the PBE or (ii) be indifferent amongst all on-path signals satisfying $m < p^{-1}\left(s, \frac{c_D}{v_D}\right)$ and strictly prefer these signals to those $m > p^{-1}\left(s, \frac{c_D}{v_D}\right)$. It follows that if sending signal $m = 0$ delivers a higher expected utility to such a type in a particular PBE, then that PBE must deliver a higher expected utility to this type regardless of their postulated strategy. But we have already shown that the probability of a concession in response to $m = 0$ is maximized when $\bar{m}$ is minimized. This completes the proof. ■

4.3.4 Step 4: The PBE Described in the Proposition is Unique Subject to the D1 Criterion ($\theta = s$)

: Let $\theta = s$. As in the c_D case, any alternate equilibrium must either feature (i) pooling on some signal $m \neq 0$, (ii) pooling by additional types on $m = 0$, or (iii) an alternative equilibrium in which all types who don't pool on $m = 0$ separate. The proof showing that cases (i) and (ii) do not survive D1 is isomorphic to the c_D case and is therefore omitted. However, the proof regarding alternative equilibria does not necessarily apply. The remainder of this proof is limited to addressing this last case.

Case (iii): there is no alternate equilibrium where all types who do not pool on $m = 0$ separate: No (potentially noncontiguous) subset of types $\tilde{S} \in [\underline{s}, \bar{s}])$ can pool on the $m = 0$ other than that described in the proposition ($[\rho, \tilde{s}_C]$). Suppose not, that is suppose that there were such a subset pooling on $m = 0$ along with the strongest types $[\tilde{s}_C, \bar{s}]$ such that

$$v_C \int_{s \in \tilde{S} \wedge s < \tilde{s}} \iota(s) f(s) ds + \left[\int_{s \in \tilde{S} \wedge s > \tilde{s}} \iota(s) f(s) ds + 1 - F(\tilde{s}_C) \right] (v_C(1 - p(s)) - c_C) = 0$$

where ι can be a mixed strategy variable wherein with some probability a particular type mixes between choosing $m = 0$ with probability $\iota(s)$ and separating with the unique signal $p^{-1}\left(s, 1 - \frac{c_C}{v_C}\right)$ with reciprocal probability $1 - \iota(s)$. Observe that this would require that there exists a type $s' < \tilde{s}_C$ separating by sending signal $m = p^{-1}\left(s', 1 - \frac{c_C}{v_C}\right)$ and another type $s'' < s'$ pooling on $m = 0$. However, this is a violation of the single-crossing property established in Proposition 8. Whereas in Step (iii) of this proof, the It follows that either type s' has a strictly profitable deviation to $m = 0$ or that type s'' has a strictly profitable deviation to $m = p^{-1}\left(s', 1 - \frac{c_C}{v_C}\right)$. It follows that there is no alternative PBE that survives D1 in this case. ■

References

Ashworth, Scott and Ethan Bueno de Mesquita (2006). “Monotone Comparative Statics for Models of Politics”. *American Journal of Political Science* 50.1, pp. 214–231.