Flying to Fail:
Costly Signals and Air Power in Crisis Bargaining*

Abigail S. Post

This draft: September 24, 2017

Abstract

Theories of crisis bargaining suggest that military mobilizations act as costly signals of resolve, increasing the credibility of coercive threats. In this paper, I argue that air mobilizations, as a subset of military signals, demonstrate lack of resolve during coercive bargaining for four reasons: They cost less in terms of human and financial resources (sunk costs), generate lower political costs (hand-tying), do not raise the risks of engagement (manipulation of risk), and do not significantly shift the balance of power—all compared with other military signals. Using new data that disaggregates military demonstrations into air, naval, and land signals during 210 cases of compellence, this paper presents systematic evidence that air signals decrease the probability of coercive threat success compared with the alternatives. This finding holds important implications for theoretical and policy debates regarding the role of costly signals in international bargaining.

*I would like to thank Julia Macdonald, Phil Potter, and Megan Stewart for helpful suggestions and comments on earlier versions of this paper.

Introduction

> "Days after North Korea conducted its most powerful nuclear test yet, two U.S. bombers flew over South Korea in a display of force—a warning to Pyongyang and reassurance to Seoul."[1]
>
> "Buzzing an airplane in the Berlin corridor does no harm unless the planes collide; they probably will not collide but they *may* and if they do the result is sudden, dramatic, irreversible, and grave enough to make even a small probability a serious one."[2]—Thomas Schelling, *Arms and Influence*

How can states improve the effectiveness of their coercive threats? In bargaining, nations hope to achieve the best deal without having to wage costly wars (Fearon 1995). States engage in efforts to convince the adversary that it is better to back down from their threats than to fight. In the context of coercive diplomacy, military maneuvers—troop mobilizations, shows of force, deployments of military assets, or other military exercises—operate as costly signals, demonstrating a willingness and/or ability to risk war. They sink costs (Fearon 1997), tie the hands of political leaders (Fearon 1994*a*), and/or raise the risks of military engagement (Schelling 1966). In addition, they shift the crisis balance of power in favor of the mobilizing side (Slantchev 2005, 2011). States engage is such costly maneuvers to increase the credibility of their coercive threats and induce higher levels of target compliance. But do all shows of force act as costly signals of resolve? While land, sea, and air power often operate in tandem during many military operations, observers debate the relative importance of each type of military force.[3] What is the role of these different types of military signals during international crisis bargaining?

This paper examines the impact of air power, compared with land and sea power, on the success of compellent threats. States engage in *military signaling* in and out of dispute bargaining. Russian jets buzz U.S. ships and planes or fly over the Crimea to remind international observers of the risks of engagement. The United States sends aircraft carriers

[1] Domonoske (2016)

[2] Schelling (1966: 91)

[3] For example, Farley (2014) argues that the United States Air Force (USAF) should be subsumed by the other branches because it lacks independent utility, even while the United States relies heavily on air power in a host of situations.

to the South China Sea, and the Chinese condemn such actions as "aggressive posturing." The U.S. deploys tanks and trunks to Eastern Europe to deter Russian aggression. South Korea and the United States regularly perform joint exercises involving air, ground, and maritime deployments, and these military drills often spark a surge in tensions with North Korea. According to rationalist bargaining theory, these diverse military demonstrations function as costly signals of resolve and capabilities (Fearon 1995). These actions increase the probability that international audiences will correctly interpret the opponent's resolve and consequently make concessions over the disputed issue(s).

Partly due to the rarity of international crises, we have little empirical evidence regarding military signals during international crisis bargaining. However, historical examples suggest that *air signals* operate as a less effective subset of military signals (the other options being land and/or sea signals) in the crisis bargaining context. At the start of the 1961 Berlin Crisis, the Soviet Union permitted their fighters to harass Allied aircraft flying through approved access corridors, beginning a cycle of escalating risk. The United States dismissed this initial "probe" as an effort to provoke the U.S. government, rather than as a serious attempt to reveal Soviet interests in Berlin. As the Berlin Crisis wore on, the United States eventually instigated a full military buildup of ground soldiers and tanks throughout Europe—but only after the Soviets continued to escalate the dispute through different venues.[4]

There is additional evidence, albeit outside of the crisis bargaining context, that both domestic and international audiences interpret air power as the less costly option. In recent months, observers have speculated that Russian air demonstrations are cheap attempts to appease domestic audiences: "Perhaps Mr. Putin believes he needs to divert the Russian people's attention away their economic troubles," said one policy analyst (Browne and Sciutto 2016). In a different context, a majority of American citizens supported the air strikes against ISIS in Syria in September 2014,[5] primarily because they were relieved that the government did not approve "boots on the ground" to support a land war (McCarthy 2015). Military analysts however questioned the wisdom of air power in this case, pondering whether

[4]See Kempe (2011) and *Letter From President Kennedy to the Supreme Commander, Allied Powers Europe (Norstad)* (1961).

[5]*The Washington Post.* September 22, 2014.

this apparently half-hearted commitment against ISIS signaled that the United States was actually lacking inherent resolve (Cordesman 2015).[6]

These historical and contemporary anecdotes bring a different question to the fore: Does air power, as a subset of military power, effectively signal resolve during international crises? Or does it function as an ineffective half-way signal of resolve as these examples suggest? In this paper, I argue that air demonstrations—bomber flyovers, the forward deployment of air assets, air mobilizations, etc.—signal *lack of resolve* during coercive diplomacy, compared with land and sea signals. As a result, a target state will be more likely to resist a threat accompanied by an air signal, compared with other military signals. I focus my theoretical and empirical framework on cases of coercive diplomacy prior to war (peacetime coercion), but I also draw on the logic of air power and wartime coercion to make the argument.[7]

How does this logic operate in terms of existing theories of costly signaling? First, air signals sink fewer costs than land or sea signals in both human and material terms, so an air demonstration does not signal resolve as effectively as its costlier alternatives. By utilizing air signals during coercive diplomacy, the challenger signals that it cares about the issue—but not enough to commit substantial land or naval resources. Second, because of the low costs involved, air signals do not engage political ire to the extent that land and naval signals do. In the language of audience costs, political elites risk lower levels of sanctioning when they back down from threats accompanied by air signals. Third, air signals risk lower levels of dispute escalation. Air signals may be a precursor to further military action, but it is much easier for a nation to limit escalation with air power than with land or naval engagements. Finally, air signals do not shift the crisis balance of power as much as land or naval signals. While a movement of land or naval assets to the potential arena of operations prepares the

[6]The name of this operation against ISIS is "Operation Inherent Resolve."

[7]A number of scholars and policy analysts have analyzed the ability of air power to coerce within interstate war (Pape 1996; Byman et al. 1999; Horowitz and Reiter 2001; Byman and Waxman 2002; Allen 2007; Van Creveld 2011; Pietrucha and Renken 2015) and in the context of counterinsurgency (Corum and Johnson 2003; Kocher, Pepinsky and Kalyvas 2011; Johnston and Sarbahi 2016; Lyall 2015; Felter 2014). There is also considerable debate over specific cases, such as the 1999 air campaign over Kosovo (Byman and Waxman 2000; Stigler 2002/03; Pape 2004; Lake 2009), the Persian Gulf War (Press 2001), and the Vietnam counterinsurgency (Clodfelter 1989; Thies 1980). Most academic scholarship suggests that air power lacks general coercive utility within war (Pape 1996).

challenger for war, air demonstrations generally do not generate a new, significant military presence in the area. Air power absent land support also cannot credibly threaten to take and hold territory, so such efforts do not signal the willingness and ability to wage a winning war.

Using new and original data that disaggregates military signals into air, land, and naval signals during 210 cases of interstate compellence (Sechser 2011), this paper is the first to systematically analyze the comparative role of military signals during interstate crisis bargaining across space and time. I combine the logic of costly signaling with theories of coercive air power, paying careful attention to the selection of military signals, to reveal that air power functions as a less costly demonstration of resolve during crisis diplomacy. In consequence, I find that air signals are far less effective at coercing target states when compared with other military demonstrations. Additionally, air power decreases coercive threat effectiveness even when coupled with other military maneuvers. The paper's theoretical synthesis and empirical results hold important implications for existing debates regarding costly signals in crisis bargaining and the use of military force more broadly.

Military Signaling and Air Power

How can states improve the success of their coercive threats? In coercion, the goal of the challenging state is either to deter the target state from some action or compel the target state to take some action. Both place the impetus on the target, making the challenger's goal one of suasion. According to scholars of coercion, coercive success is a function of four elements: "increasing the costs of continued resistance, raising the certainty that these costs will be suffered, lowering the benefits, or reducing the probability of success" (Pape 1996: 16). Within the costly signaling literature, conspicuous military mobilizations or demonstrations of military force improve the credibility of a state's threats during coercive diplomacy. These military signals—in which a state employs a military exercise, troop mobilization, show of force, deployment of military assets, or other military move—operate as costly actions

that reveal new information regarding capabilities and resolve.[8] These maneuvers alleviate information asymmetries so that the target can update its estimate that the challenger will follow through on its threat. There are four pathways through which this might occur.

Fearon (1997) models military mobilization as a *sunk costs* signal. Military maneuvers are financially costly, and spending money on an action communicates that the state values the issue at hand. In models of sunk costs signaling, these actions must take on an "all or nothing" stance in order to effectively demonstrate resolve.[9] Military mobilizations also tie the hands of the political leader, creating political costs should the government back down from the military threat (Fearon 1994*a*, 1997). Gelpi and Griesdorf (2001) show that democracies tend to win international crises more frequently when they demonstrate resolve through military action, potentially through escalatory *hand-tying* mechanisms. The further a state escalates through military action, the more likely it becomes that a (usually democratic) public will punish a leader for backing down in the military crisis (Tomz 2007).[10] Along similar lines, military signaling is an exercise in *brinkmanship*, because it generates a real risk of war (Schelling 1966; Fearon 1995). A deployment of military forces or a demonstration of pre-existing assets increases the likelihood of accident. This risk increases the possibility that the nation will get dragged into all-out war; the willingness to engage in this risky behavior communicates resolve. In a different strain, military signals also reduce the probability of successful target resistance and/or raise the costs of such resistance. Slantchev (2005, 2011) and Tarar (2013) argue that military mobilizations shift the local balance of military power in favor of the mobilizing side. According to this logic, military maneuvers prepare the mobilizing state for war. This preparation demonstrates a willingness to plan for and risk war. It disadvantages the opponent who does not mobilize in turn.

Thus, the general consensus is that military maneuvers act as costly signals of resolve and/or generate new capabilities during crisis diplomacy. These actions increase the credi-

[8]Resolve is an inherently unobservable feature. See Fearon (1994*b*).

[9]Jervis (1970: 93) first theorized that spending money is a plausible capability indice (costly signal) but concluded that "once the money is spent it does not increase the incentive to follow the policy it was spent on since the money is a 'sunk cost' and thus irrelevant to the calculation of gains and losses that alternative policies entail."

[10]This logic also applies to autocratic regimes to varying degrees (Weeks 2008).

bility of coercive threats by communicating a willingness and ability to wage war. As a form of military mobilization, air demonstrations are presumably one way for a state to signal resolve.[11] Nonetheless, the literature regarding costly signals is primarily theoretical and rarely distinguishes among different types of military signals. In the few empirical studies on military signaling during crises, the datasets categorize all military signals into one variable without distinguishing by type.[12] But there are reasons to believe that air mobilizations and demonstrations, compared with the alternatives, will be uniquely ineffective signals according to the four theoretical logics of military mobilization (sunk costs, tying hands, brinkmanship, and balance of power).

First, in the language of sunk costs, air power acts as the cheaper alternative to land and naval power. Air signals, which can take on many forms, often operate in a much more limited role than land or naval demonstrations. They consequently demand fewer government resources and risk fewer military casualties. In the context of interstate war, the United States often utilizes air power since the Gulf War partly because it is less financially costly. Operation Inherent Resolve, which relies almost exclusively on air power, has incurred $8.7 billion in costs since its commencement (Department of Defense 2016), compared with the $79 billion annual cost of Operation Iraqi Freedom, which relied on ground forces (Belasco 2011). We can extrapolate from this example of wartime coercion to situations of peacetime coercion: air signals entail lower sunk costs than land signals. Additionally, while naval and air assets differ less on financial costs, the deployment of naval assets often entails larger *opportunity costs* than the deployment of air assets. Today, there is concern that the forward-based or rotational deployment of 35% of American naval forces leaves U.S. defenses vulnerable around the world.[13]

Importantly, air signals entail lower human costs as well, as they risk fewer military casualties should the demonstration escalate.[14] This distinction intuitively fits an air versus

[11] According to Schelling (1966: 91), a bomber flyover (one example of an air signal) is a perfect example of "the threat that leaves something to chance."

[12] See, for example, Gelpi and Griesdorf (2001) and Lai (2004).

[13] *An Assessment of U.S. Military Power: U.S. Navy* (2017). See also Colby and Solomon (2016).

[14] Although Press (2001) disputes the decisive role of air power in the Gulf War, he provides a good overview of those scholars and policymakers who view air power as the less risky option.

land comparison, as land troops are generally more exposed than airmen, even when used in limited trip-wire measures.[15] The distinction also applies to naval versus air signals, because air demonstrations risk fewer casualties than naval demonstrations. Fighter jets in the USAF generally only have one crew-member (the pilot) and bombers no more than five; close air support vehicles like the AC-130 include five officers and eight enlisted men. In comparison, littoral combat ships, relatively small surface vessels with minimal crew, have approximately 40 crew aboard. Thus, the sinking of one ship risks far more military casualties that the downing of one fighter. As the less costly option in either financial and/or human terms, air signals should be less effective than the alternatives at demonstrating resolve through sunk costs pathways during crisis bargaining.[16]

In part because of lower anticipated casualty rates and financial burdens, air signals generate fewer political costs. According to Pape (1996: 2), "As the American public's willingness to bear military costs declines, the role of air power in overseas conflicts is increasing..." As the costs of air power decrease, public preferences for air power increase. Politicians can accordingly use air signals without generating high levels of "audience costs" (Fearon 1994*a*). Risk-averse domestic publics may also perceive a lowered risk of "mission creep" and getting dragged into a long and drawn out ground campaign with air power.[17] Risk-averse publics should prefer military signals that mitigate the risk to military personnel, and air power is often the answer for politicians who presumably prefer re-election to electoral ousting.

Indeed, while military maneuvers manipulate the risk of war, air signals do not raise the risks of engagement to the same extent as other military options. A government risks an accident when it deploys any military signal, so governments who chance these accidents engage states in brinkmanship and the manipulation of risk. All military signals raise the risk of the coercing state being sucked into a wider war, but air signals decrease this risk compared with land and/or naval signals.[18] States often rely on air signals during coercive diplomacy

[15]Schelling (1966: 47)

[16]Note that this does not have anything to do with the proportion of a government budget assigned to the Air Force. It instead focuses on the costliness of an air signal after the resources are already in place.

[17]This discussion closely relates to increasing American preferences for unmanned rather than manned drones (Schneider and Macdonald 2016).

[18]See Schelling (1960: 199-201); and Schelling (1966: 99-126). This approach is heavily situated

because they entail a lower risk of being dragged into unwanted conflict. Naturally, this argument drawing on the logic of brinkmanship is more relevant to air *demonstrations* such as bomber flyovers rather than to air *mobilizations* that require more substantial military preparations. Still, a land and naval demonstration or a troop mobilization and naval move entail more significant risks of escalation than a bomber flyover or mobilizing air assets, respectively.

This is partly due to the relative casualty ratio of a given accident. Again, compare naval and air demonstrations. The sinking of a ship entails a significant number of casualties and also the deaths of personnel dedicated to maintaining the ships' basic functions and operations. The downing of a plane risks far fewer casualties. For example, in June 2017, a container ship and an American destroyer (USS *Fitzgerald*) collided off the coast of Japan. While only seven of the 350 on board died, this number is likely higher than had a plane been brought down. Reports of the incident state that some crewman, "believing the ship was under attack, hurried to man the guns."[19] Had this been an attack or even an accident instigated by an enemy state, a battle may have ensued. The American public in particular is extremely sensitive to casualties and may demand of the government a "proportional response" after such incidents, which will be much smaller in the case of air power.[20]

Finally, air signals do not shift the local balance of power to the extent that land and/or naval signals do. Changes in the crisis balance of power have been demonstrated to increase coercive threat success (Slantchev 2011; Tarar 2013). Moving military assets to the theater of operations or mobilizing the military prepares the challenger to wage war. Deploying naval forces prepares for the launching of cruise missiles. Such movements thus increase threat effectiveness by *preparing* the challenger for war and thereby demonstrating the *willingness* to wage war. However, air demonstrations do not significantly shift the local balance of power in a crisis. When a state engages in flyovers, those bombers return to the base

in the literature regarding nuclear brinkmanship. See Jervis (1984, 1989) and Powell (1985, 1990, 2015).

[19] *The New York Times*. June 18, 2017.

[20] Granted, the risks may be dependent on the technological innovation of the opponent. There is concern today that American naval forces face increased vulnerability as Chinese attack capabilities increase (Colby and Solomon 2016). On the other hand, relative air vulnerability oscillated among participants of WWII (Biddle 2010: 269).

they originated from. The deployment of land and naval assets represents a slightly more permanent and/or significant change in the local balance of forces. Of course, a state can send air signals beyond merely bomber flyovers. Planes can be moved to forward bases, as was done in the Berlin Blockade (1948-49). Aircraft carriers can be (re)deployed, as was done during the period leading up to the Gulf War (1990). Levels of overall readiness/alert can be raised, as before the Kosovo airstrikes (1999). However, air assets are much more easily recalled. Comparatively, air signals as a class generate a significantly less permanent and significant shift in the local balance of power. Air assets are a flexible alternative to land and naval assets, easier to deploy and remove from a theater of operations.

Relatedly, air signals do not threaten a significant military campaign that can manipulate the balance of power should war eventually break out. Air power does destroy assets, which can alter the expected outcome of war in favor of the attacker.[21] However, the problem with air power is that *on its own* it does not explicitly threaten to take and hold ground (or sea), a crucial step to achieving military victory (Biddle 2010: 68–69). On the other hand, land deployments threaten invasion and naval deployments on their own threaten naval warfare, bringing the fighting force to bear on the situation. In the case of air power, unless there are air bases nearby, this effect is less significant. In turn, demonstrations of air power do little to promise/threaten decisive victory against the target, should the crisis escalate to war. Thus, because air signals imply an ineffective air campaign, a target state might be less likely to be coerced, holding all else constant.

Selecting Military Signals

Given these theoretical weaknesses regarding air power as a costly signal of resolve, why would a rational state use them in this function? I propose one possible explanation for this selection process, and I discuss and test other possible selection mechanisms in Appendix C. In this paper, I conceptualize of air, land, and naval signals as substitutable policy options. My theoretical framework of selection thus closely mirrors that of the foreign policy

[21] I am grateful to an anonymous reviewer for pointing this out.

substitutability literature (Clark and Reed 2005; Clark, Nordstrom and Reed 2008; Palmer and Morgan 2006). First, I assume that states have preferences and make attempts to realize those preferences through their actions. Second, states take actions in which they expect the benefits to outweigh the costs and in which the expected benefits are greater than the opportunity costs. Finally, states select from a menu of available policies the bundle of policies that best suits their goals.

How is it that then that states choose among land, naval, and air signals? For one, states may select land and/or naval signals when they are resolved and air signals when they lack resolve. Assuming that states are somewhat rational actors, this is at least part of the story. However, cases of compellence present us with challengers that possess fairly homogenous preferences. In the dataset utilized here, each challenger issues an explicit compellent demand against the target(s) to change their behavior. Present in each of these instances is a 1) threat to use military force, 2) a demonstration of military force, and/or 3) the use of military force. By making this threat publicly, the challenger increases the probability of audience costs. All of these challengers thus try to communicate resolve through some form of costly action in an effort to induce target compliance.

Following from this, both highly-resolved and "less-resolved" (for lack of a better term) types in the data want to achieve maximum output with minimum effort; the first by revealing its resolve and the latter by bluffing its way through. Still, why use extremely costly signals when cheaper signals *might* do the trick? By utilizing air power, both types are likely trying to coerce with efforts that entail lower costs and lower risks of escalation. For example, Chamberlain (2016) finds that U.S. threats often fail in the post-Cold War period because the U.S. repeatedly tries to coerce "on the cheap." She argues that the United States usually tries to gain the issue at stake with as little risk and cost as possible, even when it is willing to follow through on its threats.

Governments trying to balance multiple foreign policy goals choose among air, land, and naval signals after calculated cost-benefit analysis. Air signals are a "cheaper" method for the government to appease risk-averse domestic audiences, reassure allies, and signal resolve to adversaries—all without significantly increasing the risk of military conflict. Thus, while

governments may be cognizant of the limitations of air power as a signaling option, I hold that governments choose air signals because they provide states with a low-cost option to extract demands from targets (Schneider and Macdonald 2016). Air signals do work *some* of the time. However, because successful compellence is difficult to accomplish (Schelling 1966), such "cheap attempts" should be associated with higher levels of threat failure than more costly efforts in the form of land and naval signals.

Some may argue that it is unfair to judge air signals as costly signals if there is any possibility that they are intended for different audiences as different signals. However, existing models of military mobilization and costly signaling clump air power in with other demonstrations of military power, presumably driving down estimates of signaling effectiveness. This study disaggregates signals that may indeed have different purposes in order to better understand their aggregate impact on coercive threat success.

Hypotheses on Air Power

According to these logics, demonstrations of air power should decrease threat effectiveness when compared with land and/or naval signals. Air signals cost less in terms of human and material resources (sunk costs), generate lower political costs (hand-tying), do not raise the risks of engagement (manipulation of risk), and do not significantly shift the balance of power—all compared with other military signals. For these reasons, a target state will be more likely to resist a coercive threat accompanied by an air signal than a land or naval signal. This discussion lends itself to expression in two hypotheses. First, air power will be less effective as an *independent* signaling device compared with land and naval options, decreasing the probability that a state will comply with the challenger's demands.

Hypothesis 1 (**Independent Effects of Air Signals**). *All else equal, air signals will decrease coercive threat success compared with land and/or naval military signals.*

This paper's theoretical setup analyzes each military option as a distinct signaling attempt. However, even those who doubt the independent utility of air power during wartime

coercion admit the contributions of theater air power to joint military operations. "History shows that land power and theater air power are much more powerful coercive tools than strategic bombing... the combination of theater air power and land power can often defeat or coerce an opponent..." (Pape 1997*a*: 193). In a similar manner, one might argue that air power contributes to combined signaling efforts. Air power may be used in a complementary fashion with land and/or naval power in several ways: air coverage can help land troops move forward and take ground; naval carriers can get air assets closer to the arena of operations; etc.

However, the logics of costly signaling offer a less optimistic view of air power's contributions to joint signaling efforts. Air power still adds little extra cost or risk (human, financial, or political) to a military signal. Air power can destroy assets, but it cannot hold that new balance of power without considerable contributions from other military branches. Accordingly, adding air demonstrations to a coercive demand does not generate a more substantial military threat. For the same reasons outlined above, air power should contribute little to coercive threat success when used in conjunction with other military operations. Military demonstrations during the crisis will add weight and costliness to a coercive threat, but air power will add little to the threat's effectiveness.

Indeed, if air power is truly signaling lack of resolve, we should see it *decrease* threat effectiveness even when combined with other signaling efforts. Air power signals a desire to mitigate costs. Thus, combined demonstrations with air power should fare poorly compared with demonstrations of land and/or naval power. Even if a state is willing to escalate with land and/or naval power, air power communicates a desire to mitigate risk during a signaling operation.

Hypothesis 2 (**Combined Effects of Air Signals**)**.** *All else equal, air signals will decrease threat success when combined with other military demonstrations, compared with military demonstrations absent air power.*

Research Design and Data

This paper tests the effects of air and other military signals on the success of coercive threats. It utilizes the Militarized Compellent Threats (MCT) dataset, which contains information on 210 interstate compellent threats among 242 dyads issued between 1918 and 2001 (Sechser 2011). The dataset defines a compellent threat as "an explicit demand by one state (the challenger) that another state (the target) alter the status quo in some material way, backed by a threat of military force if the target does not comply" (Sechser 2011: 380). While the demands must be made verbally, the threats to use force may be communicated verbally or through conspicuous demonstrations of military force.

According to the coding rules of the dataset, the challenger issues a threat and the target must in turn choose to resist or concede the demand. Thus, the challenger engages in *compellence*, a category of coercion in which the challenger attempts to alter the status quo, i.e. change another actor's behavior. This contrasts with *deterrence*, in which the challenger seeks to make another actor refrain from certain behavior.[22] Of the 210 compellent threats in the data set, challenging states (or coalitions) achieved success by some definition approximately 47 percent of the time, in ninety-eight cases.

One potential limitation of the MCT dataset is its inclusion of only explicit compellent demands, necessarily excluding coercive threats in which a demand was implied rather than stated. However, this exclusion has some benefits for a study of costly signaling. For one, the category of "success" is clearly defined and operationalized. Within these compellence cases, it is clear what the demand is and whether the target state has conceded. In deterrence cases, it is more difficult to assess whether the state refrained from a particular action due to the challenger's deterrent threat or some pre-existing intention. Also, this dataset is particularly appropriate for comparing the effect of military signals to the non-use of military signals on coercive threats. Threats accompanied by military demonstrations are very public, so the appropriate comparison group is those explicit, conspicuous threats unaccompanied by

[22]Schelling (1966) argues that compellence is easier to demonstrate but more difficult to achieve than deterrence.

military demonstrations. This dataset allows us to hold the baseline public aspect of the crises constant across cases.

Dependent Variable

The dependent variable in the analysis is target COMPLIANCE. The dataset categorizes the target's response to the challenger's demand as *non-compliance*, cases which entail no target concessions to challenger demands; *partial compliance*, cases in which the target acquiesces to some but not all of the challenger's demands; and *full compliance*, cases in which the target complies fully (or nearly so) with the challenger's demands. In the analysis below, COMPLIANCE is coded 1 if the target voluntarily complies with any—as opposed to all—of the demands of the challenger and the challenger does not employ military force that results in more than 100 target fatalities. It is coded 0 otherwise.[23]

Independent Variables

The aggregate measure of military signaling comes directly from the MCT dataset. The variable MILITARY SIGNAL indicates whether the challenger employed a military exercise, troop mobilization, show of force, deployment of military assets, or other military signal to bolster its verbal demand. Since the MCT dataset does not distinguish between different types of military signals, I conduct new research on military signaling for more fine-grained data. I collected data on the type of signal employed in each MCT episode, coding three variables to indicate whether the challenger(s) engaged in LAND, NAVAL, and/or AIR signaling. If no military demonstration was used in the crisis, the case was categorized as a purely VERBAL threat. The independent variable of interest for this paper is AIR SIGNAL, a dichotomous variable that takes on a value of 1 if the challenger used a visible air demonstration, air mobi-

[23]Coding the dependent variable to exclude cases of partial compliance or using a trichotomous measure of compliance changes neither the substance nor significance of the results that follow. Robustness checks using these different specifications are reported in Appendix B.

lization, deployment of air assets, or other show of air power during a crisis, and 0 otherwise. There are 72 observations of air signals across 62 cases in the dataset. That is, seventy-two states engaged in air signaling efforts, but some of these signaling states operated within a coalition of states during one case of coercion. To illustrate: if three states used air signals as part of one coalition, that would result in three observations of air signaling but one case of compellence. Given that the research question is whether air power is an effective tool compared with other available options, I account for LAND SIGNALS (152 observations) and NAVAL SIGNALS (68 observations) for comparison. There can be overlap among the LAND, NAVAL, and/or AIR signaling categories. Additionally, the variable that codes for VERBAL threats (59 observations) forms the base comparison group throughout the analyses.

The second set of results compares the combined effect of military signals with and without air support. Instead of coding by individual signal (land, naval, or air), this variable codes whether some combination of signals was used in a given case. The variable JOINT AIR SIGNAL codes all cases in which air power was used in combination with another military demonstration. An observation takes on the value of 1 if air power was used at any point during the crisis with other military support and 0 otherwise.[24] NO AIR SIGNAL accounts for all observations in which air power was not used but other military demonstrations (land and/or naval power) were. This variable takes on a value of 1 if air power is not used at any point in the crisis but some other military signal is and 0 otherwise. The base category is cases in which no military demonstrations were employed. There is no overlap among the JOINT AIR SIGNAL, NO AIR SIGNAL, and VERBAL categories.

One limitation of this coding method is that it does not disaggregate the MILITARY SIGNAL variable even further. It does not identify exactly what aerial actions are being taken, to see if some kinds of air signals—like forward deploying air assets, missile tests, bomber overflights, or exercises involving assets already present in the conflict zone—are more effective than others. During the coding process, I found much of this information to be unavailable for the early observations in this dataset. The historical literature frequently

[24]There are only five cases in which states relied *only* on air power as a military demonstration of resolve. Two of the five cases succeeded. These cases are excluded from the second set of results but included as a robustness check in Appendix B.

mentions the military tool (air, land, naval) at least in passing, but it provides only vague details as to how those assets were used. Especially when these different military signals were combined, it is sometimes unclear in what ways air power was used to support the signaling effort, even though an air demonstration clearly occurred during the case. This fine-grained coding might be possible with a subset of the data or with a different dataset including more contemporary cases of coercion and would be an important next step to build on this study.

Notwithstanding these limitations, this is the first dataset, to my knowledge, to disaggregate costly signals in any way by type to understand their comparative impact during crisis bargaining.[25] The new data used here also points to revisions in theories of costly signaling during coercive bargaining. Research that examines military mobilizations as signals of resolve either studies it in a purely theoretical sense or collapses all military demonstrations into an effectively homogenous signaling device. However, there are many ways to cut the data, even besides what I have done here: the size of the signal, the geographic location of the military demonstration, whether the target responds with a military signal of its own, etc. As this study aptly demonstrations, the area of costly signaling is rich ground for empirical analysis.

Control Variables

In addition to the primary independent variables of interest, this analysis includes several additional factors that are shown to influence coercive threat success. Drawing on the Correlate's of War Composite Index of National Capability (CINC) score (v. 4), CAPABILITY RATIO captures the *ex ante* balance of power within the dispute dyad,[26] the proportion of material capabilities controlled by the initiator in each dyad (Singer 1987). As an alternative measure of capabilities, some models include indicator variables for whether the dispute dyad included major and/or minor powers. Additionally, NUCLEAR CHALLENGER accounts

[25] See Appendix A for further details regarding the coding of this data. For a new dataset regarding air power during interstate wars, see Allen and Martinez Machain (2017).

[26] For different interpretations of how the *ex ante* balance of capabilities influences interstate crises, see Huth and Russett (1988) and Fearon (1994*b*).

for whether the challenger possesses nuclear capabilities; and NUCLEAR TARGET accounts for whether the target has nuclear capabilities.

Much of the research on signaling emphasizes the importance of issue type as a determinant of compellent threat success (e.g., George and Simons 1994; Art and Cronin 2003). The MCT dataset classifies each compellent threat according to what is being demanded. The challenger can demand TERRITORY or the removal of military assets from the disputed area; territorial disputes make up 143 observations (59 percent) in the dataset. Monetary REPARATIONS entail demands for compensation of perceived or actual injustices; there are 17 demands (7 percent) for reparations. The challenger may demand the removal of an individual within the target government's LEADERSHIP. There are 28 instances (12 percent) of demands for leadership change. Finally, the challenger may demand changes in the target's national POLICY that do not entail territorial concessions. Demands for policy change make up 110 observations (47 percent). Finally, the base category OTHER includes all other demands that do not fall naturally into one of the four above categories. A compellent threat can include demands made in one or several of these categories.

Additionally, a challenger's military demonstrations may be highly correlated with the contiguity status of the target. If two states are contiguous, land demonstrations will be easier and potentially more effective at gaining concessions. Alternatively, contiguous states are also at a higher risk of entering a crisis together in the first place, so neighboring target states may be more likely to resist demands.[27] The Correlates of War (COW) Direct Contiguity Data (Stinnett et al. 2002) includes information on the contiguity relationship for dyads separated by a land or river border. This CONTIGUITY variable is equal to 1 if the states in the crisis dyad share a land border and 0 otherwise.[28] Also, a threat issued by multiple challengers may more likely to succeed, so the COALITION variable accounts for whether the threat is jointly issued. Within this dataset, states within the same coalition may use the same military signals or employ different demonstrations of force.

[27]For the importance of territorial contiguity in international disputes, see Bremer (1992); Diehl (1985).

[28]I can also code this variable to be equal to 1 if the states in the crisis dyad are separated by less than 150 miles of water and 0 otherwise (Schultz 1999). The findings do not change.

Finally, the literature on audience costs and democratic credibility argues that democracies can incur higher levels of audience costs or have institutional constraints that increase the credibility of their threats.[29] The models include two variables to account for the effects of regime type, DEMOCRATIC CHALLENGER and DEMOCRATIC TARGET. I include measures of democracy from the Polity IV project (Marshall and Jaggers 2002), using the 21-point Polity scale to denote states whose overall regime score is 16 or above.

Empirical Analysis

How do different types of military signals impact coercive threat success? The analysis below presents two sets of logistic regressions to evaluate the effect of AIR SIGNAL on TARGET COMPLIANCE. Put another way, do air signals improve the success of compellent threats compared with land and/or naval signals? The first set of results evaluates the *independent* effect of air signals, and the second assesses the contribution of air demonstrations to *joint* military maneuvers. Throughout the body of the text, I report the marginal changes in probability or predicted probabilities rather than coefficients, as these allow for a more natural interpretation of the effects. I report robust standard errors, clustered around the dyad, to account for possible interdependence among cases between the same dyad members that might deflate the standard errors.

Table 1 presents the results of several regressions, comparing the effects of AIR and LAND and NAVAL signals on compellent threat success. Across Models 1-4, the variable AIR SIGNAL has no statistically distinguishable effect on target compliance, while other military signals have a strong, positive association with target capitulation. Naval signals increase the probability of compliance by 18.81 percentage points (p-*value*=0.007) and land by 13.40 (p-*value*=0.043) on average.[30] When controlling for other factors such as the presence of naval and/or land signaling devices in Model 2, air power decreases the probability of tar-

[29]Schultz (2001) and Gelpi and Griesdorf (2001); but see Downes and Sechser (2012).

[30]The baseline predicted probability in this model is 28.25 for just verbal threats, and the average success rate for all threats is 43.39.

	1 *No Controls*	*2* *Main Model*	*3* *Nuclear Possession*	*4* *Power Status*
Air Signal	-0.596*	-0.518	-0.403	-0.454
	(0.326)	(0.369)	(0.403)	(0.373)
Land Signal	0.874***	0.673**	0.745**	0.644*
	(0.289)	(0.342)	(0.356)	(0.340)
Naval Signal	0.723**	0.933***	0.907**	0.996***
	(0.361)	(0.353)	(0.364)	(0.358)
Capability Ratio		-1.046*	-0.682	
		(0.606)	(0.638)	
Territory		-0.406	-0.511	-0.474
		(0.308)	(0.313)	(0.311)
Reparations		-0.157	-0.177	-0.117
		(0.493)	(0.473)	(0.503)
Leadership		2.236***	2.153***	2.228***
		(0.565)	(0.547)	(0.531)
Policy		-0.024	0.003	-0.099
		(0.289)	(0.294)	(0.305)
Contiguous		0.130	0.041	0.119
		(0.357)	(0.363)	(0.378)
Coalition		-0.751*	-0.588	-0.607
		(0.412)	(0.411)	(0.407)
Democratic Initiator		0.338	0.342	0.201
		(0.357)	(0.352)	(0.333)
Democratic Target		0.972***	0.941**	0.992***
		(0.365)	(0.367)	(0.343)
Nuclear Challenger			-0.934*	
			(0.541)	
Nuclear Target			2.295***	
			(0.890)	
Major Power Initiator				0.671
Major Power Target				(0.568)
Major Power Initiator				-0.384
Minor Power Target				(0.361)
Minor Power Initiator				1.470*
Major Power Target				(0.815)
Constant	-0.857***	-0.348	-0.504	-0.930*
	(0.227)	(0.629)	(0.636)	(0.514)
N	242	242	242	242
Pseudo R^2	0.042	0.151	0.175	0.165

note: Robust standard errors in parentheses.
* p <0.10, ** p <0.05, *** p <0.01.

Table 1. *Independent Effect of Air Signals on Target Compliance.*

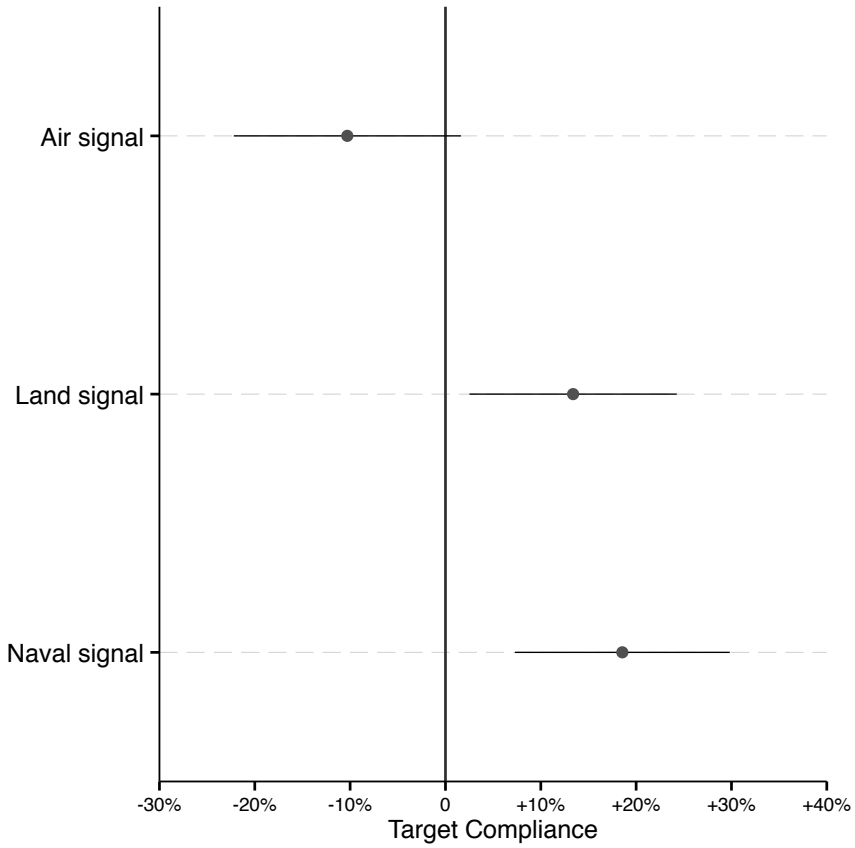

Figure 1. *Marginal effect of* AIR SIGNAL *on target compliance*

get compliance by 10.11 percentage points; but the effect is not statistically significant (*p-value*=0.156). In short, naval and land signals increase threat effectiveness far more than air signals. Figure 1 illustrates this finding, demonstrating that the confidence intervals for AIR signal overlap with zero, but the confidence intervals for LAND and NAVAL signals do not. Theories of costly signaling predict the positive impact of costly military signals on target compliance during coercion, but this finding indicates that only military signals *other* than air signals have this positive effect.

The difference between the effects of air and naval/land signals on threat effectiveness is quite striking. In all models, there is large difference between the effects of air signals and other military signals on threat success, with land and naval signals being far more effective at inducing target compliance. In Model 2, the predicted probability regarding compliance for air signals is 36.41, land is 48.38, and naval is 57.17. The differences between land and

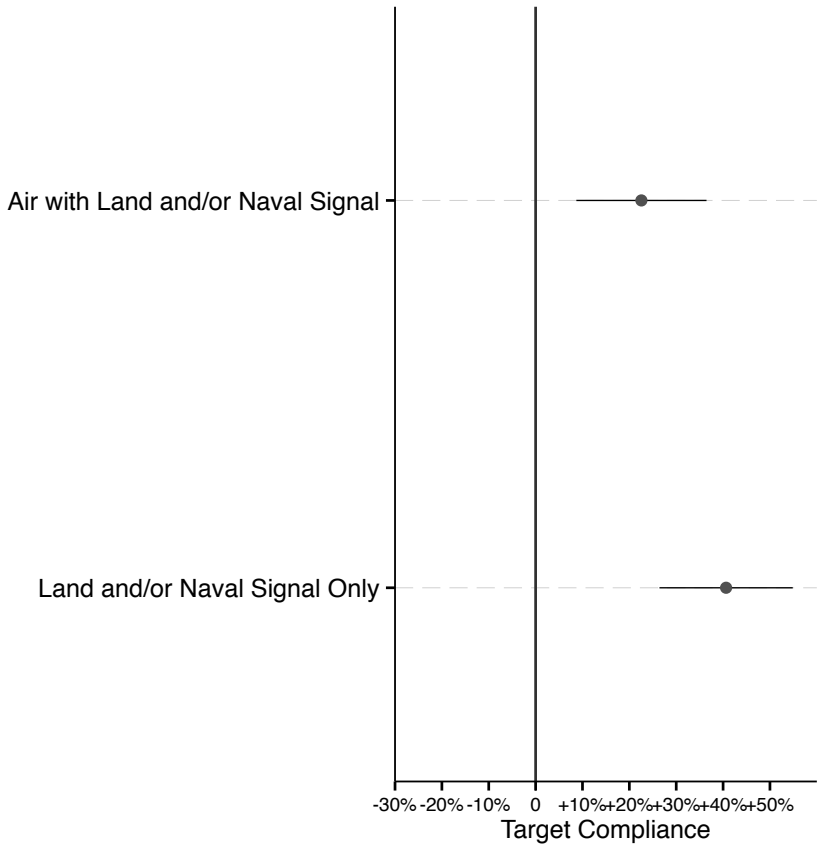

Figure 2. *Marginal effect of* JOINT AIR SIGNAL *on target compliance*

air signals (11.97 points, *p-value*=0.041) and between naval and air signals (20.76 points, *p-value*=0.013) are both substantively large and statistically significant. This comparison provides strong evidence in support of Hypothesis 1, which predicts that air power will decrease compellent threat success compared with other demonstrations of force.

However, these models do not offer an adequate test of Hypothesis 2, which argues that air power decreases the effectiveness of *combined* military signaling efforts. The regressions in Table 2 provide a cleaner test of this second prediction about the impact of joint military signaling efforts. These models compare cases of joint air signals (military signaling that entails air signals joined by land and/or naval demonstrations) to cases with no air demonstrations (military signaling that entails land and/or naval signaling only).[31] In this analysis, threats

[31]The five cases in which states relied *only* on air power as a military demonstration of resolve are excluded from the analysis. Appendix B reports results including them. Including/excluding them has no substantive impact on the results.

	5 *No Controls*	*6* *Main Model*	*7* *Nuclear Possession*	*8* *Power Status*
Air with Land and/or Naval Signal	1.196***	1.356***	1.551***	1.383***
	(0.389)	(0.450)	(0.534)	(0.445)
Land and/or Naval Signal Only	1.972***	2.213***	2.254***	2.162***
	(0.392)	(0.483)	(0.513)	(0.458)
Capability Ratio		-0.625	-0.291	
		(0.601)	(0.639)	
Territory		-0.513	-0.639*	-0.553
		(0.343)	(0.345)	(0.344)
Reparations		-0.105	-0.151	-0.075
		(0.530)	(0.496)	(0.576)
Leadership		2.279***	2.230***	2.311***
		(0.618)	(0.599)	(0.581)
Policy		0.062	0.084	0.020
		(0.328)	(0.327)	(0.331)
Contiguous		-0.122	-0.222	-0.141
		(0.348)	(0.359)	(0.384)
Coalition		-0.411	-0.281	-0.313
		(0.444)	(0.459)	(0.439)
Democratic Initiator		0.502	0.542	0.409
		(0.354)	(0.355)	(0.338)
Democratic Target		1.128***	1.105***	1.125***
		(0.393)	(0.388)	(0.366)
Nuclear Challenger			-1.021*	
			(0.600)	
Nuclear Target			2.191**	
			(0.859)	
Major Power Initiator Major Power Target				0.598
				(0.514)
Major Power Initiator Minor Power Target				-0.326
				(0.391)
Minor Power Initiator Major Power Target				1.069
				(0.867)
Constant	-1.589***	-1.589**	-1.698**	-1.880***
	(0.333)	(0.706)	(0.727)	(0.610)
N	237	237	237	237
Pseudo R^2	0.094	0.199	0.221	0.210

note: Robust standard errors in parentheses.
* p <0.10, ** p <0.05, *** p <0.01.

Table 2. *Combined Effect of Air Signals on Target Compliance.*

issued without air support do better than threats coupled with air demonstrations. In Model 6, the variable JOINT AIR SIGNAL (labelled "Air With Land and/or Naval Signal") increases the probability of target compliance by 22.57 percentage points (p-*value*=0.001). The variable NO AIR SIGNAL (labelled "Land and/or Naval Signal Only") increases the probability of target compliance by 40.65 percentage points (p-*value*=0.000). Although the confidence intervals for these two variables in Figure 2 overlap, the difference between the two groups of 18.08 percentage points is statistically significant (p-*value*=0.019). This result demonstrates that air power does not give a significant bump to but rather decreases the success of military operations, lending support to Hypothesis 2. This finding goes against the common supposition that air power can, at a minimum, be used effectively in combination with land and naval forces to improve coercion outcomes.

Taken together, the results reported in Table 1 and Table 2 provide evidence that air signals decrease coercive threat success when compared with other military signals. Land and naval signals increase threat effectiveness far more than air signals do in these models. Even when used in combination with other military forces, air power decreases the probability of target capitulation. In sum, these results provide strong support for the theory that air power signals lack of resolve when compared with other military options. These tests demonstrate that compellent threats are less effective when governments try to avoid the costs and risks associated with signaling resolve.

Robustness Checks

Table 3 reports the results of a variety of additional regressions. These tests are meant to assess the possibility that specific modeling choices, coding decisions, unique aspects of the dataset, or extreme outlying cases might be driving the results. First, it is possible that cases with multiple challengers—coalitions of challenging states—are overrepresented in the sample. If coalitions of states are more likely to engage in demonstrations of air power, those failures will bias the results against air power. For example, several states as part

of one coalition engaged in air mobilizations preceding the 1999 war against Kosovo, but that compellent threat failed. Model 9 excludes subordinate partners from the regression, retaining only the state coded by the MCT dataset as the "primary" challenger in the crisis. With this coding, each case of compellence only shows up in the dataset one time. Excluding these cases, however, has no impact on the results surrounding the military demonstrations variables. Coalitional threats do not appear to drive the results.

Second, the choice set of what signals to send varies across challengers in the MCT dataset. That is, not all states have an equal capability to use air, land, and naval signals, especially in the early years of the dataset. To partially alleviate this problem, I constrain the statistical model to a subset of the sample in Model 10, running the model on a truncated dataset including only major power challengers. These major power states did possess considerable air power during all dyad-years of the MCT dataset and were capable in every case of engaging in air demonstrations to buttress their compellent threats.[32] In Model 11, among states that can send air and other military signals—and send them effectively—air signals are still far less effective than land signals at inducing target compliance. However, the variable for NAVAL signals no longer reaches statistical significance (*p-value*=0.113). Consequently, the substantive difference of 11.61 points between naval and air signals in this model is only significant at *p-value*=0.182, probably due to the small number of observations (120).

Third, some have noted that the U.S. compellence record is quite poor (e.g., Blechman and Kaplan 1978; George and Simons 1994; Chamberlain 2016). Since the U.S. is prone to rely on air power during international disputes, the failure of air power in these models may simply be a result of U.S. compellence difficulties. Model 11 excludes all cases of U.S. compellent threats (21 observations) to check this possibility. However, the key variables regarding military signals remain essentially the same, suggesting that the U.S. cases probably are not driving the findings.

Finally, the likelihood that a challenger will choose a certain military signal may be

[32]All major powers in the dataset obtained an air force, either independently or as a subsidiary of another military branch, by their first appearance in the dataset. The United States (1918-2001) by 1907; France (1918-1940) by 1909; Germany (1925-1945) by 1910; Italy (1918-1943) by 1923; U.S.S.R. (1922-1945) by 1917; China (1950-2001) by 1949; Japan (1918-1945) by 1912.

	9 *Exclude Coalitions*	*10* *Major Powers*	*11* *Exclude USA*	*12* *Non-Contiguous*
Air Signal	-0.640*	-0.186	-0.609	-0.459
	(0.385)	(0.407)	(0.418)	(0.409)
Land Signal	0.960**	0.885**	0.768**	1.009**
	(0.379)	(0.439)	(0.367)	(0.465)
Naval Signal	0.826**	0.733	0.871**	0.860*
	(0.375)	(0.463)	(0.392)	(0.469)
Capability Ratio	-0.911	-0.117	-0.995	-1.642**
	(0.624)	(0.755)	(0.626)	(0.807)
Territory	-0.371	-0.133	-0.535	0.127
	(0.354)	(0.385)	(0.329)	(0.384)
Reparations	-0.241	0.320	-0.148	0.218
	(0.546)	(0.815)	(0.487)	(0.702)
Leadership	2.036***	1.849**	2.473***	0.753
	(0.601)	(0.728)	(0.541)	(0.860)
Policy	0.151	-0.160	-0.071	-0.474
	(0.318)	(0.439)	(0.304)	(0.364)
Contiguous	-0.067	0.431	0.092	
	(0.356)	(0.479)	(0.368)	
Coalition	-0.878	-1.611**	-0.825**	-1.417**
	(0.573)	(0.728)	(0.415)	(0.615)
Democratic Initiator	0.458	1.231**	0.348	1.559***
	(0.388)	(0.573)	(0.343)	(0.603)
Democratic Target	1.055***	0.826	1.024***	1.634***
	(0.389)	(0.596)	(0.379)	(0.484)
Constant	-0.607	-1.600**	-0.316	-0.838
	(0.642)	(0.699)	(0.640)	(0.624)
N	210	120	221	124
Pseudo R^2	0.144	0.161	0.153	0.226

note: Robust standard errors in parentheses.
* p <0.10, ** p <0.05, *** p <0.01.

Table 3. *Independent Effect of Air Signals on Target Compliance.*

highly correlated with whether it shares a border with the target state. Challengers within contiguous dyads are probably less likely to use naval or air signals, while challengers in non-contiguous dyads inevitably rely on a host of options to signal resolve. In Model 12, I constrain the dataset to non-contiguous dyads. This model thus includes challengers that might weigh the three options more equally than challengers making demands from a contiguous target. Among noncontiguous dyads (those who do not share a land border), the results are essentially the same: Land signals are somewhat more effective, increasing the probability of target compliance by 18.06 percentage points (p-*value*=0.025), and naval signals are somewhat less effective, increasing threat effectiveness by 15.08 percentage points (p-*value*=0.073). Air signals continue to have an effect indistinguishable from zero.

Discussion, Limitations, and Conclusions

Scholars continue to dispute the coercive role of air power in single cases and across space and time. This paper examines coercive air power from a new angle, assessing the role of air power as a costly signal during international crisis bargaining. It argues that air demonstrations communicate a lack of resolve during coercive bargaining. Using new data on air, naval, and land demonstrations during interstate crises, this paper finds that air signals do not increase the success of compellent threats. Air demonstrations are far less effective at inducing target compliance when compared with other military demonstrations.

Of course, this is not to say that air power is ineffective in all problem areas. This paper does not assess the usefulness of air power in a host of other situations, including reconnaissance, force application, and force enhancement processes, to name a few. However, the findings here jive with the growing consensus that air power has limited utility for coercion. While Pape's conclusions concerning air power prompted a litany of criticisms and responses from multiple angles (Watts 1997; Warden 1997; Pape 1997*a*,*b*; Mueller 1998), coercive air power (particularly when used in *punishment campaigns*) has been found to have limited utility within war (e.g., Horowitz and Reiter 2001).[33] This paper finds comparable

[33]The debate over the role of air power in Kosovo also sheds light on the nuances of this debate

patterns in the period leading up to war as well.

One limitation of the study and data is its focus on the challenger state over the target state. The theory here does not delve into the goals/interests of the state being coerced, and the MCT dataset does not include much information on the target state. However, this limitation is unlikely to undermine the findings of this paper regarding the success of compellent threats. If a state is the recipient of a compellent threat, it is likely that the state has already begun to undertake an action that is undesirable and is relatively highly resolved to achieve its goals. Given the target's high resolve, the coercing state(s) will need to employ especially costly signals to bring about a desired behavior. Air power is not a particularly costly signal, so we should expect air mobilizations and demonstrations to be less effective in these cases. However, air signals may be more effective for deterrent purposes, in which the target state's resolve is weaker and the stakes are lower. That is, if the target state has relatively low stakes in a crisis, air mobilizations might actually serve as a credible and effective signal of resolve. A study of deterrence may be better able to address variation in target interests and resolve and find different effects for various military signals.

This paper suggests a significant revision to our understanding of military mobilization and costly signals. To further evaluate this paper's central claim—that audiences interpret air demonstrations as less costly signaling attempts and respond accordingly—we should move to models and datasets that better assess the scope of the signal employed. We also need to focus on the intent behind the signal. Of course, intent is much more difficult to assess in practice than in theory. State leaders often "intend" a policy to communicate multiple meanings to different audiences and may revise their own understanding based on its effects. Future research might look at how leaders (especially presidents) choose among the signaling options in the coercive bargaining context. This may require extensive archival research to understand why leaders rely on certain types of military power. For example, those with military experience may prefer certain types of signals compared with civilian presidents, just

(Pape 2004). Stigler (2002/03) argues that the strategic bombing campaign in Kosovo coerced Milošević's surrender. On the other hand, Byman and Waxman (2000) argue that air power played a role but NATO's threat of a ground invasion ultimately ended the conflict. Lake (2009) challenges all of these conclusions over Kosovo, arguing that strategic bombing in conjunction with economic and political sanctioning caused political destabilization in Milošević's regime.

as military background can influence the willingness to use force more generally (Horowitz and Stam 2014; Horowitz, Stam and Ellis 2015; Sechser 2004).

At a minimum, these findings suggest that air power communicates lack of resolve across a dataset of compellence threats. According to the theoretical framework provided here, air signals reveal important limitations to a state's cost tolerance. In so doing, such acts signal a degree of restraint (perhaps unintentionally) and come across as half-hearted signals. Instead of being costly enough to demonstrate resolve, air signals are costless enough to demonstrate lack of resolve. Recent scholarship on air power notes that "...even when military professionals doubt the effectiveness of air power in certain situations, the siren song of airpower continues to appeal to civilians committed to doing *something*" (Farley 2014: 189). These results lead to the frustrating conclusion that succumbing to the demand to just "do something, anything" may be an efficient way to accomplish little during cases of international bargaining.

References

Allen, Susan Hannah. 2007. "Time Bombs Estimating the Duration of Coercive Bombing Campaigns." *Journal of Conflict Resolution* 51(1):112–133.

Allen, Susan Hannah and Carla Martinez Machain. 2017. "Understanding the Impact of Air Power." *Conflict Management and Peace Science* .

An Assessment of U.S. Military Power: U.S. Navy. 2017. In *2015 Index of U.S. Military Strength*, ed. Dakota L.. Wood. Washington: The Heritage Foundation.

URL: *http://index.heritage.org/military/2017/assessments/us-military-power/u-s-navy/*

Art, Robert J. and Patrick M. Cronin. 2003. *The United States and Coercive Diplomacy*. Washington, DC: U.S. Institute of Peace.

Belasco, Amy. 2011. "The Cost of Iraq, Afghanistan, and Other Global War on Terror Operations Since 9/11.".

URL: *http://fpc.state.gov/documents/organization/150178.pdf*

Biddle, Stephen. 2010. *Military Power: Explaining Victory and Defeat in Modern Battle*. Princeton, NJ: Princeton University Press.

Blechman, Barry M. and Stephen S. Kaplan. 1978. *Force without War: U.S. Armed Forces as a Political Instrument*. Washington, D.C.: Brookings Institution.

Bremer, Stuart A. 1992. "Dangerous Dyads: Conditions Affecting the Likelihood of Interstate War, 1816-1965." *Journal of Conflict Resolution* 36(2):309–341.

Browne, Ryan and Jim Sciutto. 2016. "Russian jets keep buzzing U.S. ships and planes. What can the U.S. do?" *CNN* .

URL: *http://www.cnn.com/2016/04/18/politics/russia-jets-buzz-u-s-ship-rules-of-engagement/*

Byman, Daniel, John G McGinn, Keith Crane, Seth G Jones, Rollie Lal and Ian O Lesser. 1999. *Air Power as a Coercive Instrument*. Santa Monica, CA: RAND.

Byman, Daniel L and Matthew C Waxman. 2000. "Kosovo and the Great Air Power Debate." *International Security* 24(4):5–38.

Byman, Daniel and Matthew Waxman. 2002. *The Dynamics of Coercion: American Foreign*

Policy and the Limits of Military Might. New York; Cambridge University Press.

Chamberlain, Dianne Pfundstein. 2016. *Cheap Threats: Why the United States Struggles to Coerce Weak States.* Washington, D.C.: Georgetown University Press.

Clark, David H, Timothy Nordstrom and William Reed. 2008. "Substitution is in the Variance: Resources and Foreign Policy Choice." *American Journal of Political Science* 52(4):763–773.

Clark, David H and William Reed. 2005. "The Strategic Sources of Foreign Policy Substitution." *American Journal of Political Science* 49(3):609–624.

Clodfelter, Mark. 1989. *Limits of Air Power.* New York: The Free Press.

Colby, Elbridge and Jonathan F Solomon. 2016. "Avoiding Becoming a Paper Tiger." *Joint Forces Quarterly* 82(3):24–32.

Cordesman, Anthony H. 2015. "Iraq and Syria: The Problem of Strategy." *Center for Strategic and International Studies* .

URL: *https://www.csis.org/analysis/iraq-and-syria-problem-strategy*

Corum, James S and Wray R Johnson. 2003. *Airpower in Small Wars: Fighting Insurgents and Terrorists.* Laurence, KA: University Press of Kansas.

Craighill, Peyton M. 2014. "Why Americans' support for bombing ISIS may not last." *Washington Post* .

URL: *https://www.washingtonpost.com/news/storyline/wp/2014/09/22/why-americans-support-for-bombing-isis-may-not-last/*

Diehl, Paul F. 1985. "Contiguity and Military Escalation in Major Power Rivalries, 1816-1980." *The Journal of Politics* 47(4):1203–1211.

Domonoske, Camila. 2016. "U.S. Flies Bombers Over South Korea In Show Of Force To North Korea.".

URL: *http://www.npr.org/sections/thetwo-way/2016/09/13/493729428/u-s-flies-bombers-over-south-korea-in-show-of-force-to-north-korea/*

Downes, Alexander B. and Todd S. Sechser. 2012. "The Illusion of Democratic Credibility." *International Organization* 66(3):457–89.

Farley, Robert M. 2014. *Grounded: The Case for Abolishing the United States Air Force.*

University Press of Kentucky.

Fearon, James D. 1994*a*. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3):577–592.

Fearon, James D. 1994*b*. "Signaling Versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model." *Journal of Conflict Resolution* 38(2):236–269.

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

Fearon, James D. 1997. "Signaling Foreign Policy Interests: Tying Hands Versus Sinking Costs." *Journal of Conflict Resolution* 41(1):68–90.

Felter, Joe. 2014. "Airstrikes Can Only Do So Much to Combat ISIS." *New York Times* . **URL:** *http://www.nytimes.com/roomfordebate/2014/08/13/the-strengths-and-limits-of-air-power-6/airstrikes-can-only-do-so-much-to-combat-isis*

Gelpi, Christopher and Michael Griesdorf. 2001. "Winners or Losers? Democracies in International Crisis, 1918–94." *American Political Science Review* 95(3):633–47.

George, Alexander L. and William E. Simons. 1994. *The Limits of Coercive Diplomacy, 2nd ed.* Boulder, CO: Westview.

Horowitz, Michael C and Allan C Stam. 2014. "How Prior Military Experience Influences the Future Militarized Behavior of Leaders." *International Organization* 68(03):527–559.

Horowitz, Michael C, Allan C Stam and Cali M Ellis. 2015. *Why Leaders Fight*. Cambridge University Press.

Horowitz, Michael and Dan Reiter. 2001. "When Does Aerial Bombing Work? Quantitative Empirical Tests, 1917-1999." *Journal of Conflict Resolution* 45(2):147–173.

Huth, Paul and Bruce Russett. 1988. "Deterrence Failure and Crisis Escalation." *International Studies Quarterly* 32(1):29–45.

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

Jervis, Robert. 1984. "The Illogic of American Nuclear Strategy.".

Jervis, Robert. 1989. *The Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon*. Ithaca, NY: Cornell University Press.

Johnston, Patrick B and Anoop K Sarbahi. 2016. "The Impact of US Drone Strikes on Terrorism in Pakistan." *International Studies Quarterly (forthcoming)* .

Kempe, Frederick. 2011. *Berlin 1961: Kennedy, Khrushchev, and the Most Dangerous Place on Earth*. New York: Putnam.

Kocher, Matthew Adam, Thomas B Pepinsky and Stathis N Kalyvas. 2011. "Aerial Bombing and Counterinsurgency in the Vietnam War." *American Journal of Political Science* 55(2):201–218.

Lai, Brian. 2004. "The Effects of Different Types of Military Mobilization on the Outcome of International Crises." *Journal of Conflict Resolution* 48(2):211–29.

Lake, Daniel R. 2009. "The Limits of Coercive Airpower: NATO's 'Victory' in Kosovo Revisited." *International Security* 34(1):83–112.

Letter From President Kennedy to the Supreme Commander, Allied Powers Europe (Norstad). 1961. In *Foreign Relations of the United States, 1961–1963, Volume XIV, Berlin Crisis, 1961-1962*, ed. Charles S. Sampson. Washington, D.C.: Government Printing Office (1993).

URL: *https://history.state.gov/historicaldocuments/frus1961-63v14/d185*

Lyall, Jason. 2015. "Bombing to Lose? Airpower and the Dynamics of Violence in Counterinsurgency Wars." *Unpublished Manuscript* .

Marshall, Monty and Keith Jaggers. 2002. *Polity IV Project: Political Regime Characteristics and Transitions, 1800-2002*. College Park: University of Maryland.

McCarthy, Justin. 2015. "In U.S., 53% Oppose Sending Ground Troops to Fight Militants." *Gallup* .

URL: *http://www.gallup.com/poll/186590/oppose-sending-ground-troops-fight-militants.aspx*

Mueller, Karl. 1998. "Strategies of Coercion: Denial, Punishment, and the Future of Air Power." *Security Studies* 7(3):182–228.

Palmer, Glenn and T Clifton Morgan. 2006. *A Theory of Foreign Policy*. Princeton, NJ: Princeton University Press.

Pape, Robert A. 1996. *Bombing to win: Air Power and Coercion in War*. Ithaca: Cornell

University Press.

Pape, Robert A. 1997*a*. "The Air Force Strikes Back: A Reply to Barry Watts and John Warden." *Security Studies* 7(2):191–214.

Pape, Robert A. 1997*b*. "The Limits of Precision-guided Air Power." *Security Studies* 7(2):93–114.

Pape, Robert A. 2004. "The True Worth of Air Power." *Foreign Affairs* 83:116–130.

Pietrucha, Mike and Jeremy Renken. 2015. "Airpower may not win wars, but it sure doesn't lose them." *War on the rocks* .

URL: *http://warontherocks.com/2015/08/airpower-may-not-win-wars-but-it-sure-doesnt-lose-them/*

Powell, Robert. 1985. "The Theoretical Foundations of Strategic Nuclear Deterrence." *Political Science Quarterly* 100(1):75–96.

Powell, Robert. 1990. *Nuclear Deterrence Theory: The Search for Credibility*. New York: Cambridge University Press.

Powell, Robert. 2015. "Nuclear Brinkmanship, Limited War, and Military Power." *International Organization* 69(03):589–626.

Press, Daryl G. 2001. "The Myth of Air Power in the Persian Gulf War and the Future of Warfare." *International Security* 26(2):5–44.

Schelling, Thomas C. 1960. *The Strategy of Conflict*. Cambridge, Mass.: Harvard University Press.

Schelling, Thomas C. 1966. *Arms and Influence*. New Haven: Yale University Press.

Schneider, Jacquelyn and Julia Macdonald. 2016. "U.S. Public Support for Drone Strikes: Why Do Americans Prefer Unmanned over Manned Platforms?" *Center for New American Security* .

Schultz, Kenneth A. 1999. "Do Democratic Institutions Constrain or Inform? Contrasting Two Institutional Perspectives on Democracy and War." *International Organization* 53(2):233–266.

Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy*. New York: Cambridge University Press.

Sechser, Todd S. 2004. "Are Soldiers Less War-Prone than Statesmen?" *Journal of Conflict Resolution* 48(5):746–774.

Sechser, Todd S. 2011. "Militarized Compellent Threats, 1918–2001." *Conflict Management and Peace Science* 28(4):377–401.

Shane, Scott. 2017. "Sleeping Sailors on U.S.S. Fitzgerald Awoke to a Calamity at Sea." *The New York Times* .

URL: *https://www.nytimes.com/2017/06/18/world/asia/navy-uss-fitzgerald-japan.html*

Singer, J. David. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816–1985." *International Interactions* 14(2):115–132.

Slantchev, Branislav L. 2005. "Military Coercion in Interstate Crises." *American Political Science Review* 99(4):533–47.

Slantchev, Branislav L. 2011. *Military Threats: The Costs of Coercion and the Price of Peace*. New York: Cambridge University Press.

Stigler, Andrew L. 2002/03. "A Clear Victory for Air Power: NATO's Empty Threat to Invade Kosovo." *International Security* 27:124–57.

Stinnett, Douglas M, Jaroslav Tir, Paul F Diehl, Philip Schafer and Charles Gochman. 2002. "The Correlates of War (COW) project direct contiguity data, version 3.0." *Conflict Management and Peace Science* 19(2):59–67.

Tarar, Ahmer. 2013. "Military Mobilization and Commitment Problems." *International Interactions* 39(3):343–66.

Thies, Wallace J. 1980. *When Governments Collide: Coercion and Diplomacy in the Vietnam Conflict, 1964-1968*. Berkeley: University of California Press.

Tomz, Michael. 2007. "Domestic Audience Costs in International Relations: An Experimental Approach." *International Organization* 61(4):821–840.

U.S. Department of Defense. 2016. "Operation Inherent Resolve: Targeted Operations Against ISIL Terrorists.".

URL: *http://www.defense.gov/News/Special-Reports/*

Van Creveld, Martin. 2011. *The Age of Airpower*. New York: Public Affairs.

Warden, John A III. 1997. "Success in Modern War: A Response to Robert Pape's Bombing

to Win." *Security Studies* 7(2):172–190.

Watts, Barry D. 1997. "Ignoring Reality: Problems of Theory and Evidence in Security Studies." *Security Studies* 7(2):115–171.

Weeks, Jessica L. 2008. "Autocratic Audience Costs: Regime Type and Signaling Resolve." *International Organization* 62(01):35–64.

Appendices for "Flying to Fail"

Abigail S. Post
University of Virginia
asn9dp@virginia.edu

Updated September 24, 2017

Table of Contents

This document contains supporting information for the manuscript "Flying to Fail."

Appendix A: Description of Coding and Data

I research news reports and secondary sources to code military demonstrations in each case of compellence (210) between 1918 and 2001 in the Militarized Compellent Threats (MCT) dataset (Sechser 2011). In the documents, I search for evidence of a military demonstration and then research further to determine whether air, land, and/or naval power was used as part of the demonstration. According to the coding rules of the MCT dataset, MILITARY SIGNAL is a "dichotomous variable denoting episodes in which the challenger employed demonstrations or shows of force or conspicuous military mobilizations in conjunction with the compellent demand." The variable MILITARY SIGNAL indicates whether the challenger employed a military exercise, troop mobilization, show of force, deployment of military assets, or other military signal to bolster its verbal demand.

As stated in the text of this paper, I collected data on the *type* of signal employed in each MCT episode, coding three variables to indicate whether the challenger(s) engaged in LAND, NAVAL, and/or AIR signaling. If no military demonstration was used during the episode, the case was categorized as a purely VERBAL threat. AIR SIGNAL is a dichotomous variable that takes on a value of 1 if the challenger used a visible air demonstration, air mobilization, deployment of air assets, or other show of air power during an episode, and 0 otherwise. There are 72 observations of air signals across 62 cases in the dataset. That is, sixty-two states engaged in air signaling efforts, but some of these signaling states operated within a coalition of states during one case of coercion. To illustrate: if three states used air signals as part of one coalition, that would result in three observations of air signaling but one case of compellence. In the same manner, I code LAND SIGNAL as 1 if a land signal is present and 0 otherwise. There are 152 observations of land signals across 143 cases. Finally, I code NAVAL SIGNAL if a naval demonstration was present and 0 otherwise. There are 68 observations of naval signals across 60 cases. There can be overlap among the LAND, NAVAL, and/or AIR signaling categories. Additionally, the variable that codes for VERBAL threats (59 observations) is by extension coded 1 when no military demonstration is present.

A quintessential example of an air demonstration occurred in the case of Greece and Cyprus in 1964 (MCT-130). After Turkish air raids over Cypriot villages, Cypriot President Makarios demanded that the Turkish air raids cease, warning Turkey that Greek Cypriots would begin indiscriminate attacks on Turkish Cypriot villages if they continued. Greece threatened to support the Greek Cypriots with military force if Turkey did not comply; and on August 10 the Greek Air Force fighters flew over south Cyprus as a show of force against Turkey. This case is coded 1 for AIR SIGNAL and 0 for all else, since Greece did not display any other forces. There are five (5) observations of only air signals being used in the MCT dataset.

We see examples of challenging states employing just land forces in eighty-one (81) cases. A few instances occur during territorial disputes between Pakistan and India. For example, on April 12, 1965, Indian Prime Minister Lal Bahadur Shastri demanded the removal of Pakistani troops at Kanjarkot, promising "appropriate action" if Pakistan refused (MCT-131). The *New York Times* reported concentrations of Indian and Pakistani troops between the Indian state of Gujarat and the Sind region of West Pakistan. This case is coded 1 for LAND SIGNAL and 0 for all else, since India did not display any other forces.

There are a number of cases of "gunboat diplomacy" by the United States and other great powers in the dataset. There are fourteen (14) cases in which only naval power is displayed. A clean-cut example of a naval demonstration occurred during the Cienfuegos crisis of 1970 (MCT-146). On September 16, U-2 photographs revealed that the Soviets were constructing a submarine base in violation of a 1962 agreement reached following the Cuban Missile Crisis. President Nixon and Secretary of State Kissinger met with Soviet Ambassador Anatoly Dobrynin on September 25 and demanded that the Soviet Union dismantle the submarine base and withdraw its naval flotilla from Cuba. President Nixon subsequently dispatched a U.S. destroyer to Cienfuegos harbor to "emphasize [the] warning." This case is coded 1 for NAVAL SIGNAL and 0 for all else, since the United States did not display any other forces.

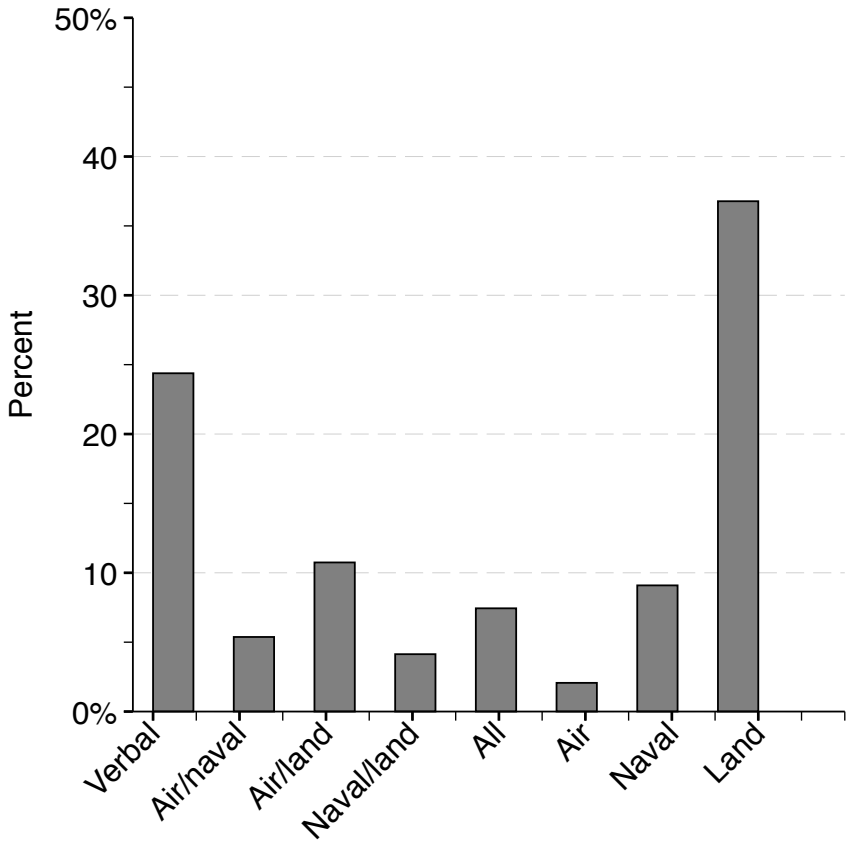

Figure SI-1. *Percent of observations by category*

There is some ambiguity in the coding with aircraft carriers. One example is MCT-199. In November 1997, the UN Security Council passed a resolution against Iraq that would impose new sanctions if Iraq continued to obstruct weapons inspectors. Iraq responded by expelling American weapons inspectors on November 13. The next day, President Clinton dispatched an aircraft carrier to the Persian Gulf near Iraq. This case is coded both naval *and* air, because an aircraft carrier blurs the distinction between these two categories.

Additionally, deploying naval forces if often a precursor to the landing of military forces and jet overflights. For example, on November 18, 1961, the U.S. demanded that Rafael Trujillo resign as dictator of the Dominican Republic (MCT-123). The U.S. stationed an estimated eight vessels and 1,800 marines off of Santo Domingo and flew naval jet fighters overhead twice. In this case, the naval deployment hugely shifted the local balance of power,

Signal	*Success Rate*
Verbal Threat	16.95% (10/59)
Air Signal	40.28% (29/72)
Naval Signal	51.47% (35/68)
Land Signal	50.00% (76/152)

Table SI-1. *Raw averages for threat effectiveness.*

bringing along the threat of land invasion and strategic bombing. Thus, NAVAL SIGNAL is coded 1 for the eight vessels, AIR SIGNAL is labelled 1 due to the overflights, and LAND SIGNAL is coded 1 because land troops were just offshore.

While many air demonstrations include flyovers, there are cases in which air forces are merely deployed. In 1998, Iran deployed around 70,000 Iranian troops to the border of Afghanistan and held large military demonstrations, using both "air, land, and volunteer forces" (MCT-202). Iran engaged in this demonstration after the Taliban attacked the Iranian consulate in Mazar-e Sharif and took 11 Iranian diplomats and dozens of other Iranians hostage. Although no indication was given whether there were actual flyovers, AIR SIGNAL would be coded 1 because air assets were included as part of the demonstration.

A typical example of a full-on combined demonstration includes the 1990 buildup against Iraq in which President Bush ordered more than 150,000 additional American ground, sea and air forces to the Gulf region. In this case, LAND SIGNAL, NAVAL SIGNAL, and AIR SIGNAL would all be coded as 1. Note that I do not code by the branch (air force, navy, army, marines) but rather by the types of force used, irrespective of military branch. A full breakdown of how many signals fall into each category is detailed in Figure SI-1.

I provide some raw baseline statistics regarding the relationship between military signals

and threat success in Table SI-1. At first glance, air signals are far less successful at inducing target rate compliance than land or naval signals, only succeeding in 29 of 72 cases or approximately 40% of the time, although they do improve coercive outcomes compared to no military signal at all. Note also that this estimate does not control for the presence of other (land and/or naval) signals in the cases. These raw numbers lend additional support to the models reported in the main text of the paper.

Appendix B: Additional Robustness Checks

In Table SI-2, I replicate the models reported in Table 3 (Models 9-12) as robustness checks using the combined version of the variables (JOINT AIR SIGNAL and NO AIR SIGNAL, used to test Hypothesis 2). I also replicate Model 6 including the observations in which only air power was used in the JOINT AIR SIGNAL variable, as Table 2 excludes these observations. The results remain consistent regardless of the model specification.

	A1 All Air Signals	*A2 Exclude Coalitions*	*A3 Major Powers*	*A4 Exclude USA*	*A5 Non-Contiguous*
AIR WITH LAND AND/OR NAVAL SIGNAL	1.410***	1.682***	2.540***	1.285**	3.897***
	(0.437)	(0.571)	(0.623)	(0.511)	(1.095)
LAND AND/OR NAVAL SIGNAL ONLY	2.187***	2.371***	3.061***	2.220***	5.040***
	(0.480)	(0.607)	(0.756)	(0.517)	(1.096)
CAPABILITY RATIO	-0.746	-0.607	-0.016	-0.568	-1.623
	(0.601)	(0.709)	(0.737)	(0.628)	(1.041)
TERRITORY	-0.578*	-0.484	-0.313	-0.685*	-0.243
	(0.327)	(0.425)	(0.472)	(0.368)	(0.483)
REPARATIONS	-0.116	-0.502	0.563	-0.121	0.279
	(0.519)	(0.590)	(0.932)	(0.518)	(0.944)
LEADERSHIP	2.240***	1.975***	2.317**	2.546***	0.985
	(0.609)	(0.708)	(0.922)	(0.606)	(1.091)
POLICY	-0.003	0.521	-0.216	-0.006	-0.573
	(0.321)	(0.401)	(0.506)	(0.345)	(0.450)
CONTIGUOUS	-0.048	-0.210	0.282	-0.157	
	(0.349)	(0.398)	(0.484)	(0.364)	
COALITION	-0.420		-0.921	-0.532	-0.425
	(0.409)		(0.753)	(0.452)	(0.697)
DEMOCRATIC CHALLENGER	0.426	0.278	1.212*	0.516	1.808**
	(0.349)	(0.429)	(0.643)	(0.342)	(0.774)
DEMOCRATIC TARGET	1.154***	1.187***	1.305*	1.178***	2.709***
	(0.392)	(0.459)	(0.729)	(0.406)	(0.841)
CONSTANT	-1.445**	-1.886**	-3.330***	-1.470**	-4.145***
	(0.697)	(0.870)	(0.881)	(0.718)	(0.975)
N	242	186	119	217	121
Pseudo R^2	0.199	0.186	0.237	0.205	0.367

NOTE: Robust standard errors in parentheses.
† p <0.10, * p <0.05, ** p <0.01, *** p <0.001.

Table SI-2. *Combined Effect of Air Signals on Target Compliance.*

In Table SI-3, I replicate the main results (Models 2 and 6) using different specifications of the dependent variable. Models A6 and A8 exclude partial successes when operationalizing compellent threat success. Cases are coded as successful if the challenger accomplishes all, not just some, of its demands. I check the results with this coding because it may be the case that air power contributes less significantly than the alternatives to partial successes in the data. If this is the case, classifying these partially successful cases as "successes" might bias the results against air power. This does not appear to be the case. Using this alternative coding, the substantive results regarding air power do not change: the coefficient on AIR SIGNAL remains negative and statistically indistinguishable from zero. The coefficient on LAND SIGNAL in Model A6 is somewhat less significant (p-*value*=0.058) than in the main model, but the difference between land and air signals is still somewhat significant (p-*value*=0.097).

Models A7 and A9 report the results using a trichotomous measure of success, in which the variable is coded 0 for no compliance, 1 for partial compliance, and 2 for full target compliance. The results remain strong (indeed, they are somewhat stronger) with this specification of the dependent variable as well, adding further credence to the findings in the main body of the paper.

	A6 *Partial Successes*	*A7* *Trichotomous measure of Success*	*A8* *Partial Successes*	*A9* *Trichotomous measure of Success*
Air Signal	-0.421	-0.504		
	(0.445)	(0.374)		
Land Signal	0.709⋆	0.695⋆⋆		
	(0.374)	(0.354)		
Naval Signal	0.806⋆⋆	0.908⋆⋆⋆		
	(0.359)	(0.332)		
Air with Land and/or Naval Signal			1.476⋆⋆⋆	1.398⋆⋆⋆
			(0.521)	(0.462)
Land and/or Naval Signal Only			2.258⋆⋆⋆	2.230⋆⋆⋆
			(0.550)	(0.488)
Capability Ratio	-1.245⋆	-1.133⋆	-0.800	-0.708
	(0.663)	(0.629)	(0.673)	(0.627)
Territory	-0.141	-0.327	-0.201	-0.437
	(0.358)	(0.322)	(0.398)	(0.361)
Reparations	-0.717	-0.378	-0.734	-0.322
	(0.504)	(0.440)	(0.505)	(0.467)
Leadership	2.669⋆⋆⋆	2.482⋆⋆⋆	2.757⋆⋆⋆	2.542⋆⋆⋆
	(0.655)	(0.601)	(0.705)	(0.648)
Policy	0.054	-0.003	0.150	0.059
	(0.313)	(0.299)	(0.349)	(0.336)
Contiguity	0.006	0.067	-0.246	-0.180
	(0.426)	(0.379)	(0.399)	(0.359)
Coalition	-0.950⋆⋆	-0.808⋆	-0.646	-0.486
	(0.474)	(0.416)	(0.509)	(0.432)
Democratic Initiator	0.182	0.273	0.310	0.401
	(0.397)	(0.365)	(0.406)	(0.362)
Democratic Target	1.102⋆⋆⋆	1.065⋆⋆⋆	1.244⋆⋆⋆	1.251⋆⋆⋆
	(0.400)	(0.385)	(0.430)	(0.419)
Constant	-0.563		-1.896⋆⋆	
	(0.717)		(0.841)	
Cut 1		0.300		1.532⋆⋆
		(0.666)		(0.741)
Cut 2		0.560		1.817⋆⋆
		(0.667)		(0.748)
N	242	242	237	237
Pseudo R^2	0.186	0.142	0.235	0.183

note: Robust standard errors in parentheses.
† p <0.10, ⋆ p <0.05, ⋆⋆ p <0.01, ⋆⋆⋆ p <0.001.

Table SI-3. *Independent Effect of Air Signals on Target Compliance.*

Appendix C: Selection Effects

It is important to further consider whether these findings can be explained by selection effects in the data. While I propose one explanation for military signal selection in the body of the text, there are other forms of selection effects that could potentially undermine the findings. I consider several of them here.

First, it may be the case that powerful nations use air signals more frequently. Great powers are more likely to possess advanced air forces, and powerful challengers are shown to have their threats resisted on a more frequent basis due to the reputational concerns of the target (Sechser 2010). In this case, air signals may fail more frequently because they are employed by states who are predisposed to have their challenges resisted more often. To assess this possibility, I look at whether 1) the dispute-dyad capabilities are more unbalanced in cases of air power than in others or 2) major or minor powers are more likely to use air power.

Using CAPABILITY RATIO as a measure of relative power balance to explore this possibility, I do not find evidence that powerful challengers are more likely to engage in one type of military signal than another. First, the average capability ratio in the dispute dyad, reported in Table SI-4, is nearly the same across the distribution of demonstrations. The average capability ratio for land signals is slightly lower, but the difference between it and air or naval power is not significant.[1]

Additionally, major powers use air signals at a similar rate compared with minor powers, which is not what we would expect if selection effects were at play. The logic of selection effects suggests that great powers, who are more likely to have their threats fail, are more likely to use air signals. Table SI-4 suggests that major and minor powers utilize military signals at a similar rate, although major powers are more likely to utilize naval signals than are minor powers. However, this is the opposite of what we would expect if this version of

[1]None of the relationships relevant to this discussion in Table SI-4 achieve statistical significance at the 95% level using a chi-squared test.

Signal	*Capability Ratio (SD)*	*Major Powers*	*Minor Powers*	*High Stakes Issues*
Verbal Threat	0.74 (0.25)	25.00% (30/120)	23.77% (29/122)	33.90% (20/59)
Air Signal	0.72 (0.27)	30.83% (37/120)	28.69% (35/122)	41.67% (30/72)
Naval Signal	0.70 (0.31)	36.67% (44/120)	19.67% (24/122)	36.76% (25/68)
Land Signal	0.66 (0.32)	60.00% (72/120)	65.57% (80/122)	46.71% (71/152)

Table SI-4. *Selection Effects*

selection effects were at play. Naval signals should fail more, not less, than air signals if great powers are more likely to use naval signals.

A second alternative explanation for the results is that air power fails as a demonstration of resolve because it is simply used in more difficult cases. That is, states choose air demonstrations to accompany high-value demands, selecting land and naval power for low-stakes issues. However, the data reveals no evidence that air signals fail more because of the issue at stake.[2] For one, land signals are more likely to accompany high-value demands than are air signals, which is the opposite of what this version of selection effects would predict. In this explanation, land signals should fail more, not less, because they are more likely to accompany high-value demands.

On the flip side, perhaps air signals fail more often because they are used in low-stakes situations, cases in which the challenger does not care enough to follow through on its threats. The data does not bear out this possibility either. Air signals are used to accompany more high-value demands than naval signals, which is the opposite of what this version of selection effects would predict. According to this logic, naval signals should fail more frequently if they are being used more in cases that the challenger devalues.

In sum, if air power was failing systematically because of the issue at stake, we would expect it to succeed at a rate somewhere between land and naval signals rather than being

[2]High-value demands here are defined as threats over territory or leadership, which are both considered to be more important to target states than issues of policy, ideology, or reparations (George and Simons 1994; Huth and Allee 2002).

less successful than both. Table SI-4 suggests that this version of selection effects is probably not responsible for the relationships we observe in the main empirical analysis.

Heteroskedastic Ordered Model

There are other possible explanations for why leaders use air signals and why they consequently fail on a more regular basis during cases of compellence. Air signals might be intended (and interpreted) as something like a "probe." A coercer first sends air signals to see if the target backs down; it is a preferred option, because sending an air signal is cheaper than sending a ground or perhaps naval signal. If that initial air signal fails, then a low-resolved challenger backs off, and a high-resolved challenger starts issuing more credible signals, like ground and naval signals. In the first case, air signal is "cheap talk," and the use of air power is essentially a proxy for a lack of resolve on the challenger's part. In the latter case, the crisis will play out in a typical escalatory fashion, moving from air power to land and/or naval assets. In this explanation, air power acts as a pooling signal, in which it does not distinguish resolved and non-resolved types until the target state resists, prompting the resolved type to escalate and the non-resolved type to back down.

How can we distinguish between these two types in the data? The coding does not distinguish the order in which these signals are employed. However, the use of air power potentially increases the variance around target compliance, as the target may not know what type of leader they are dealing with. In some cases of air signaling, they might rightly diagnose the challenger as having low resolve/weak preferences over the demand and refuse to accede to the challenger's demands. Alternatively, the use of land and/or naval assets may reduce variance in the outcome by sending a clearer signal to the target.[3]

To assess this possibility, I run a heteroskedastic ordered model on the data (Allison 1999; Williams 2009, 2010) and model AIR SIGNAL as heteroskedastic in the model. The model includes AIR SIGNAL in the variance equation, thus allowing residual variability to differ by

[3]I am grateful to an anonymous reviewer for pointing this out.

air signals. I obtain the following estimates in Table SI-5 using the OGLM command in Stata 14. The OGLM model says that the standard deviation of the residuals is $\exp(\gamma) = \exp(1.06) = 2.88$ times larger for cases in which an air signal was used than when it was not. This heteroskedasticity is significant at the 0.10 level.

Thus, there is (weak) evidence of heteroskedasticity in air signals. However, modeling this heteroskedasticity does not change the substance or the significance of the results. The variable AIR SIGNAL continues to have a negative but statistically insignificant effect on threat success, while both NAVAL SIGNAL and LAND SIGNAL have a strong, positive association with threat success. The difference between air signals and land and naval signals also remains strong and statistically significant at *p-value*=0.051 and *p-value*=0.035, respectively. Thus, this evidence provides some support for the proposed story of the selection process, adding nuance to our understanding of signal selection. It does not, however, undermine the results reported in the body of the paper.

	A10 Main Model	*A11 Nuclear Possession*	*A12 Power Status*
Air Signal	-1.471	-0.967	-1.311
	(1.049)	(0.897)	(1.056)
Land Signal	1.203**	1.108**	1.415**
	(0.537)	(0.544)	(0.582)
Naval Signal	1.101**	1.110**	1.053**
	(0.448)	(0.466)	(0.466)
Capability Ratio	-1.641**	-1.228*	
	(0.729)	(0.742)	
Territory	-0.575	-0.680*	-0.628
	(0.389)	(0.379)	(0.391)
Reparations	-0.412	-0.436	-0.283
	(0.540)	(0.530)	(0.570)
Leadership	2.895***	2.946***	2.675***
	(0.900)	(0.793)	(0.803)
Policy	0.064	0.107	-0.065
	(0.362)	(0.359)	(0.378)
Contiguous	0.215	0.171	0.174
	(0.406)	(0.415)	(0.452)
Coalition	-0.848	-0.829	-0.551
	(0.640)	(0.608)	(0.629)
Democratic Challenger	0.664	0.626	0.471
	(0.487)	(0.487)	(0.462)
Democratic Target	0.958**	1.037**	1.000**
	(0.434)	(0.427)	(0.397)
Nuclear Challenger		-2.205	
		(1.449)	
Nuclear Target		3.445**	
		(1.671)	
Major Power Initiator Major Power Target			0.576
			(0.648)
Major Power Initiator Minor Power Target			-0.642
			(0.469)
Minor Power Initiator Major Power Target			1.518*
			(0.908)
Air Signal (lnsigma)	1.063*	1.105*	0.893
	(0.610)	(0.613)	(0.652)
Cut 1	0.358	0.468	1.199**
	(0.690)	(0.714)	(0.598)
N	242	242	242
Pseudo R^2	0.166	0.195	0.174

note: Robust standard errors in parentheses.
* p <0.10, ** p <0.05, *** p <0.01.

Table SI-5. *Independent Effect of Air Signals on Target Compliance.*

Two-stage Least-squares Regression

I also check the findings with a two-stage least-squares regression (2SLS) model looking at the selection of air signals and their subsequent impact on compellent threat success. I instrument air power with four variables. First, theories of audience costs indicate that democracies might be less likely to engage in a coercion strategy that is more likely to fail and consequently might be less likely to use air signals (Fearon 1994*a*; Schultz 2001). Alternatively, democracies are often seen as more risk averse, so they may use air signals more often to mitigate costs (Reiter and Stam 2002). I include the variable DEMOCRATIC CHALLENGER to proxy these competing predictions. Second, countries with smaller air forces and/or no nuclear arsenal are probably less likely to use air signals. In Model A13, I include CINC SCORE to proxy capabilities; in Model A14 NUCLEAR CHALLENGER; in Model A15 POWER STATUS. Additionally, air signals are more likely to be used in more recent cases, with the diffusion of air forces across the globe, so I include a dummy variable POST-1945 to account for this temporal shift.

The findings only get stronger in these models, i.e. the negative coefficient on AIR SIGNAL reaches statistical significance and the positive coefficients on both NAVAL SIGNAL and LAND SIGNAL remain statistically significant in all models in Table SI-6. However, such models must be assessed carefully, as few, if any, instruments exist that affect the decision to use an air signal but do not also plausibly influence coercive threat success.

	A13 Main Model	*A14 Nuclear Possession*	*A15 Power Status*
Air Signal	-0.534**	-0.541**	-0.557**
	(0.233)	(0.223)	(0.254)
Land Signal	0.266***	0.295***	0.272***
	(0.084)	(0.080)	(0.081)
Naval Signal	0.340**	0.308**	0.341**
	(0.134)	(0.131)	(0.147)
Target's Capabilities (CINC)	-0.538		
	(0.769)		
Coalition	-0.047	-0.007	-0.029
	(0.085)	(0.085)	(0.086)
Democratic Target	0.188**	0.166**	0.176**
	(0.081)	(0.080)	(0.080)
Nuclear Target		0.434***	
		(0.142)	
Major Power Target			0.093
			(0.108)
Constant	0.313***	0.268***	0.287***
	(0.069)	(0.065)	(0.070)
N	242	242	242
Pseudo R^2			

note: Robust standard errors in parentheses.
* p <0.10, ** p <0.05, *** p <0.01.

Table SI-6. *Independent Effect of Air Signals on Target Compliance.*

Matching Design

As a final check on the results, I use a matching design for the data to assess whether air power is just a proxy for lack of resolve on the part of the coercer. In this manner, I test my theory on cases that are comparable on observable factors so that the variation in coercive tools across cases can be isolated from other potential differences. This is an imperfect strategy, because crises are about the revelation of unobservable factors (Fearon 1994*b*). However, this strategy does allow me to match on other factors that have been shown to affect resolve prior to the onset of a crisis: regime type (e.g., Schultz 2001), capabilities (e.g., Huth and Russett 1988), contiguity as a proxy for long-standing relations (e.g., Bremer 1992; Diehl 1985), and the interests at stake (e.g., George and Simons 1994; Art and Cronin 2003). I use a nearest neighbors matching design such that the difference between air signal, naval signal, and land signal is calculated by comparing nations, which apart from the military signal they used, are very similar.

Rather than just assume nations randomly choose air/land/naval power, the matching design calculates the difference in performance between nations that choose air/land/naval signals and nations that are similar to them. The matching is based on the control variables described in the body of the paper. I match on both challenger and target regime type and capabilities, since the challenger is likely to utilize both pieces of information to calculate actions prior to entering the crisis. I used the four closest matches and implemented the algorithm using Stata 14's NNMATCH (Abadie, Herr and Imbens 2004). The results reported in Table SI-7 lend further support to my theory, as the findings regarding land, naval, and air signals maintain their significance levels and substantive interpretations.

Instrumented Variable	*Main Model*	*Nuclear Status*	*Power Status*
Air Signal	–0.077 (0.078)	–0.080 (0.077)	–0.054 (0.076)
Land Signal	0.172 (0.062)***	0.186 (0.069)***	0.161 (0.065)**
Naval Signal	0.197 (0.076)***	0.183 (0.082)**	0.178 (0.080)**

NOTE: Robust standard errors in parentheses.
* p <0.10, ** p <0.05, *** p <0.01.

Table SI-7. *Average Treatment Effect of Air Signal, Naval Signal, and Land Signal with Nearest Neighbor Matching Based on Capabilities, Regime Type, and Interests.*

References

Abadie, Alberto, Jane L. Herr and Guido W. Imbens. 2004. "Implementing matching estimators for average treatment effects in Stata." *Stata Journal* 4(3):290–3111.

Allison, Paul D. 1999. "Comparing Logit and Probit Coefficients Across Groups." *Sociological Methods & Research* 28(2):186–208.

Art, Robert J. and Patrick M. Cronin. 2003. *The United States and Coercive Diplomacy.* Washington, DC: U.S. Institute of Peace.

Bremer, Stuart A. 1992. "Dangerous Dyads: Conditions Affecting the Likelihood of Interstate War, 1816-1965." *Journal of Conflict Resolution* 36(2):309–341.

Diehl, Paul F. 1985. "Contiguity and Military Escalation in Major Power Rivalries, 1816-1980." *The Journal of Politics* 47(4):1203–1211.

Fearon, James D. 1994*a*. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88(3):577–592.

Fearon, James D. 1994*b*. "Signaling Versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model." *Journal of Conflict Resolution* 38(2):236–269.

George, Alexander L. and William E. Simons. 1994. *The Limits of Coercive Diplomacy, 2nd ed.* Boulder, CO: Westview.

Huth, Paul and Bruce Russett. 1988. "Deterrence Failure and Crisis Escalation." *International Studies Quarterly* 32(1):29–45.

Huth, Paul K and Todd L Allee. 2002. *The Democratic Peace and Territorial Conflict in the Twentieth Century.* Vol. 82 New York: Cambridge University Press.

Reiter, Dan and Allan C Stam. 2002. *Democracies at War.* Princeton, NJ: Princeton University Press.

Schultz, Kenneth A. 2001. *Democracy and Coercive Diplomacy.* New York: Cambridge University Press.

Sechser, Todd S. 2010. "Goliath's curse: coercive threats and asymmetric power." *International Organization* 64(04):627–660.

Sechser, Todd S. 2011. "Militarized Compellent Threats, 1918–2001." *Conflict Management and Peace Science* 28(4):377–401.

Williams, Richard. 2009. "Using Heterogeneous Choice Models to Compare Logit and Probit Coefficients Across Groups." *Sociological Methods & Research* 37(4):531–559.

Williams, Richard. 2010. "Fitting heterogeneous choice models with oglm." *Stata Journal* 10(4):540–567.