Limited Audience Costs in International Crises*

Ahmer Tarar  
Department of Political Science  
Texas A&M University  
4348 TAMU  
College Station, TX 77843-4348  
email: ahmertarar@pols.tamu.edu

Bahar Leventoğlu  
Department of Political Science  
Duke University  
Durham, NC 27708  
email: bahar.leventoglu@duke.edu

June 27, 2012

Abstract

Do audience costs have to be extremely large in order to credibly signal resolve and affect international crises? Existing theoretical work on audience costs suggests an affirmative answer, and recent empirical work on audience costs focuses on whether a leader can generate such large audience costs as to create a commitment to fight where no such commitment previously existed. We analyze a richer crisis bargaining model with audience costs and find that (i) audience costs can have war-reducing effects on incomplete-information crisis bargaining through a non-informative, bargaining-leverage mechanism, and (ii) audience costs can have war-reducing effects even when such large audience costs are not being generated as to create a commitment to fight where no such commitment previously existed. Even more limited audience costs can have war-reducing effects in international crises. We discuss how the bargaining-leverage mechanism is consistent with a number of prominent historical cases.

*Forthcoming in Journal of Conflict Resolution. The name ordering is reverse alphabetical, and does not denote unequal contributions. We gratefully acknowledge the support of the National Science Foundation (SES-0518945 and SES-0518185).
1 Introduction

During international crises, can leaders generate such large audience costs as to create a commitment to fight where no such commitment previously existed? If not, can audience costs have any effect in international crises? Fearon (1994) introduced the audience cost mechanism as a means by which leaders can try to credibly signal their high resolve in international crises, thereby possibly overcoming informational asymmetries that can lead to costly and inefficient war (e.g., Bueno de Mesquita and Lalman 1992; Fearon 1995; Morrow 1989).

Audience cost theory has proven influential, and scholars have applied it to a variety of other substantive settings such as international cooperation (Leeds 1999), alliances (Gaubatz 1996), economic sanctions (Martin 1993), trade (Mansfield, Milner, and Rosendorff 2002), foreign direct investment (Jensen 2003), monetary commitments (Broz 2002), and debt repayment (Schultz and Weingast 2003).

The early work on audience costs was primarily theoretical (e.g., Fearon 1994, 1997; Smith 1998), in part due to the daunting methodological challenges of empirically testing for the existence of audience costs using observational data (Schultz 2001a). More recently, a number of empirical works have avoided these problems with observational data by testing for the existence of audience costs using survey experiments. A central question in this literature is whether leaders can generate large audience costs. Tomz (2007) shows that audience costs...

---

1Schelling (1960) first argued that negotiators can generate bargaining leverage by tying their hands with public statements, but did not address signaling under incomplete information.

exist, and Trager and Vavreck (2011) find that leaders can generate very large audience costs. Levendusky and Horowitz (2012) show that when the leader justifies backing down by citing new information that is endorsed by the military and intelligence agencies, audience costs are of significantly lower magnitude than when he backs down without justifying the decision.

Underlying the literature is the assumption that only large audience costs can affect international crises. For example, Slantchev (2005, 544) writes that the audience cost mechanism “...requires the demanding assumption that leaders incur sufficiently high audience costs; so high, in fact, that peace becomes worse than war.” Trager and Vavreck (2011, 526-7) state that “…if the political costs of conflict are higher than the costs of backing down after a threat, the credibility of threats to use force will be much less than otherwise.” Snyder and Borghard (2011) write that the assumption that “domestic audience costs are substantial, independent of other reputational considerations” (p.439) is one of “four assumptions of audience costs theory that must hold true for domestic audience costs to have a significant effect on crisis outcomes” (p.438).

This assumption that audience costs must be large in order to affect international crises is intuitive and is supported by existing theoretical work on audience costs. In particular, game-theoretic analyses find that audience costs affect international crises by allowing for credible signaling (via semi-separating equilibria) of privately-held information that is relevant for the crisis, and in all of these informative equilibria, at least some types of the informed actor are generating large enough audience costs as to create a commitment to fight (unless the other side backs down) where no such commitment previously existed.\(^3\) However, it should

---

\(^3\)To see this, note that in Fearon (1994) and Schultz (1999, 2001a), each type’s payoff for war is negative, meaning that each type begins the interaction preferring to live without
be noted that none of these analyses have the explicit goal of examining whether audience costs must be large in order to affect international crises; the affirmative answer emerges implicitly from the fact that in all informative equilibria, at least some types are generating large enough audience costs to commit them to war.

In this paper, we theoretically examine whether more limited audience costs can have any effect in international crises. We do so by analyzing a game-theoretic model of crisis bargaining with audience costs in which we make a number of innovations over previous work. We first consider a setting in which each type of the informed actor begins the interaction with a credible commitment to fight unless the status-quo is sufficiently revised in its favor, i.e., each type begins the interaction dissatisfied with the status-quo. Because each type begins the interaction with a credible commitment to fight, no type can use audience costs to create a commitment to fight where no such commitment previously existed; the commitment is already there. However, because we treat the disputed good as fully divisible and allow the leader to endogenously choose how much of the good to publicly demand — a demand that generates audience costs for backing down — audience costs can be used to generate the disputed good (which results in a payoff of 0) rather than fight over it. Hence, for any type that is willing to fight in equilibrium (and there are such types), this willingness arises from generating sufficiently high audience costs as to cause it to become resolved. In Fearon (1997), high-valuation types are resolved for war from the outset, and in the essentially unique equilibrium, the (pooling) audience cost $m^*$ chosen by high-valuation and medium-valuation types is high enough to commit the latter group to standing firm (in fact, it is high enough that it would commit even low-valuation types to war, but low-valuation types choose not to generate audience costs; hence the partial signaling effect; see p.80-81). Similarly, in Schultz (2001b), there is a range of types of the challenger whose payoff for war is positive (i.e., $0 < w_c$), and hence they are resolved for war from the outset, but there is also a range of types who in equilibrium become resolved because of audience costs (i.e., $-a < w_c \leq 0$).
bargaining leverage by getting the other side to offer more of the disputed good than it otherwise would. In addition, audience costs may allow for credible signaling, if the different types at least partially differ in (i.e., “separate”) the public demands that they make.

In this setting, we find that audience costs can help stabilize the crisis by reducing the probability of war and causing agreement to be reached with types with whom war would otherwise occur. Hence, audience costs can help stabilize crisis bargaining even when their role is limited in a way that they cannot create a commitment to fight where no such commitment previously existed. However, in this setting the mechanism has nothing to do with credible signaling, but is instead an extension of the bargaining-leverage (greater share of the good) effect to the incomplete-information setting. In fact, we construct pooling, and hence completely uninformative, equilibria in which audience costs have these stabilizing effects.

We then consider the more standard setting where only the higher-resolve type of the informed actor begins the interaction with a credible commitment to fight; the lower-resolve type begins the interaction strictly preferring the status-quo to war. In this setting, audience costs have the potential to be so significant as to create a commitment to fight, because the low-resolve type can potentially generate large enough audience costs as to commit itself to war if the status-quo is not revised in its favor. This setting is similar to previous analyses of audience costs, with the exception that we use a model in which the disputed good is treated as fully divisible and the informed actor endogenously chooses the magnitude of its public demand. We find that the same type of stabilizing equilibria exist as in the earlier setting, in which it is the extension of the bargaining-leverage effect rather than credible signaling that reduces the probability of war. However, unlike in the earlier setting, stabilizing equilibria
also exist in which credible signaling is the reason why the probability of war goes down. These equilibria are similar to the informative equilibria that emerge in previous analyses of audience costs, in that information-transmission is the mechanism through which audience costs exert their effect. But in contrast to other analyses, we find that there exist such equilibria in which the low-resolve type never chooses such large public demands as to create a commitment to fight unless the status-quo is revised in its favor. That is, we find that audience costs can allow for credible signaling even when such large audience costs are not being generated as to create a commitment to fight where no such commitment previously existed.

Thus, our two main results are that (i) audience costs can have stabilizing effects on incomplete-information crisis bargaining through a non-informative, bargaining-leverage mechanism, and (ii) audience costs can have stabilizing effects, whether through the bargaining-leverage or information-revealing mechanism, even when they are not so significant as to create a commitment to fight where no such commitment previously existed. The basic intuition is that when the disputed good is fully divisible, moderate audience costs can generate moderate bargaining leverage that can influence crisis bargaining, unlike in indivisible-good settings in which audience costs only matter if they are large enough to reverse one’s preference between war and living without the disputed good. The results contradict the usual assumption in the literature that only large audience costs can have any effect in international crises, and suggest that the recent empirical literature that focuses mainly on whether large audience costs can be generated is only looking at one of the relevant empirical issues surrounding audience costs. Our theoretical analysis suggests that even limited audience costs can affect international crises, and future empirical work should examine this. In addition,
our results suggest that audience costs can exert stabilizing effects through a non-informative, bargaining-leverage mechanism, and in the Conclusion we discuss two prominent historical cases that are consistent with this mechanism.

The rest of the paper is organized as follows. In the next section, we describe the model. We then present the results when each type of the informed actor begins the interaction with a credible commitment to fight. Next, we present the results when only the high-resolve type begins the interaction with a credible commitment to fight. We conclude by discussing some additional implications of the analysis, and illustrate the stabilizing bargaining-leverage mechanism in the context of the 1962 Cuban Missile Crisis and the 1898 Fashoda Crisis.

2 Model

Consider the following crisis bargaining setting.\footnote{The model is somewhat similar to that analyzed by Tarar and Leventoglu (2009), with the following differences. Tarar and Leventoglu (2009) use an alternating-offers bargaining protocol and only consider the complete-information setting. We use an ultimatum-offer (take-it-or-leave-it) bargaining protocol and examine the complete as well as incomplete-information settings. Later, we consider the robustness of the results to alternative bargaining protocols.} Two countries, labeled $D$ (pronoun “he”) and $S$ (pronoun “she”), are involved in a dispute over a divisible good (e.g., territory) whose value to both sides is normalized to 1. If war occurs, it is treated as a game-ending costly lottery, with $D$ winning with probability $p \in (0, 1)$ and $S$ winning with probability $1 - p$, and the costs of war are $c_D, c_S > 0$. We assume risk-neutrality. Then, country $D$’s expected utility from war is $EU_D(\text{war}) = p - c_D$, and country $S$’s expected utility from war is $EU_S(\text{war}) = 1 - p - c_S$. Thus, as seen in Figure 1, $[p - c_D, p + c_S]$ is the set of agreements that both sides prefer to war (Fearon 1995).
There is some status quo division of the disputed good, \((q, 1 - q)\), where \(q \in [0, 1]\) is \(D\)'s share and \(1 - q\) is \(S\)'s share. Suppose that one side, say \(D\), is dissatisfied with the status quo, i.e., strictly prefers war to the status quo \((q < p - c_D)\). Then, the status quo has to be revised in \(D\)'s favor if war is to be avoided. To model the new agreement, if any, that is reached, we use the following ultimatum-offer bargaining protocol, as in Fearon (1995). \(S\) makes a take-it-or-leave-it offer \((y, 1 - y)\) to \(D\), where \(y \in [0, 1]\) is \(D\)'s proposed share and \(1 - y\) is \(S\)'s proposed share. \(D\)'s options are to (i) accept the offer, in which case each side's utility is its proposed share, (ii) reject by choosing the status quo, in which case each side's utility is its status quo share, or (iii) reject by going to war, in which case each side's utility is its expected payoff from war. When \(D\) is dissatisfied (i.e., \(q < p - c_D\)), this model has a unique subgame-perfect equilibrium (SPE) in which \(S\) offers exactly \(y^* = p - c_D\), which \(D\) accepts. The status quo is peacefully revised in \(D\)'s favor, with \(S\) getting her most preferred outcome within the preferred-to-war bargaining range.

Now suppose that before the ultimatum crisis bargaining game (described above) begins, the dissatisfied leader \(D\) can make a public statement about the minimal amount of the disputed good that he expects to obtain in the negotiations. We will henceforth call this statement leader \(D\)'s “public demand,” and denote it by \(\tau_D\). We will assume that this can be anything from 0 to 1, i.e., \(\tau_D \in [0, 1]\). The demand \(\tau_D = 0\) amounts to not making any public demand at all, whereas \(\tau_D = 1\) amounts to publicly demanding the entire disputed good. Any intermediate public demand is also allowed. Following leader \(D\)'s choice of \(\tau_D\), leader \(S\) makes her offer.

The payoffs are as follows. If \(D\) chooses some \(\tau_D\) and then accepts an agreement \((y, 1 - y)\)

\(^5\)This is without loss of generality, as Powell (1999) shows that at most one side can be dissatisfied at a time.
such that $y \geq \tau_D$, then leader $D$’s personal payoff is simply $y$ (the share of the disputed good that he receives). However, if $D$ accepts some $y < \tau_D$, then leader $D$’s personal payoff is $y - a_D(\tau_D - y)$ rather than $y$, where $a_D(\tau_D - y)$ is the audience cost that leader $D$ pays for accepting less than what he publicly demanded. We call $a_D > 0$ (which is an exogenous parameter of the model) leader $D$’s “audience cost coefficient,” which indicates how costly it is to violate a public demand by a given amount: the bigger $a_D$ is, the more costly it is. Moreover, for any given value of $a_D$, the bigger the deficit between leader $D$’s public demand and what he actually obtains (i.e., the bigger the difference $\tau_D - y$ is), the bigger the audience cost that he pays (e.g., perhaps his reelection probability decreases as the gap increases).\footnote{This is the same cost structure used by Leventoğlu and Tarar (2005) and Muthoo (1999) to examine how public demands affect non-crisis bargaining (where there is no outside option of war).}

We assume that the audience cost applies to any agreement reached (if it is such that $y < \tau_D$) and to the status quo (if $D$ chooses some $\tau_D > q$ and then rejects $S$’s offer by opting for the status quo), but not to the war payoff. That is, if the leader publicly demands a share of the disputed good that is greater than his war payoff but war occurs, the leader does not pay an audience cost in that case, since he is not backing down from a public demand by accepting an agreement that is less than what he demanded.

## 3 Complete-Information Results

We are primarily interested in the incomplete-information results, but because they are useful for understanding these results, we first briefly present the complete-information results.

Recall that in the ultimatum crisis bargaining model without public demands, when $D$ is
dissatisfied (i.e., \( q < p - c_D \)), there is a unique SPE in which \( S \) proposes \( y^* = p - c_D \) for \( D \), which \( D \) accepts. In the model with public demands, Figure 2 shows the offer \( y \) that \( S \) makes (dashed line), as well as leader \( D \)’s personal payoff (\( y \) minus any audience cost paid; solid line) in the SPE of the ultimatum crisis bargaining subgame, for any given public demand \( \tau_D \in [0, 1] \). When \( D \) publicly demands no more than his utility from war (\( \tau_D \leq p - c_D \)), the demand has no effect on the bargaining subgame, because \( S \)’s usual offer, \( y^* = p - c_D \), continues to be acceptable.

However, when \( \tau_D > p - c_D \), then leader \( D \) pays an audience cost in \( S \)’s usual proposal \( y^* = p - c_D \), and hence leader \( D \)’s personal payoff for accepting this proposal is less than his utility from war. Therefore, \( S \) has to offer more than \( p - c_D \) if she wants to avoid war (which she does, until \( D \)’s public demand becomes too high). In particular, \( S \) offers \( y^{**} = \frac{p - c_D + a_D \tau_D}{1 + a_D} \), which is a bigger offer that, with audience costs (i.e., it is the case that \( y^{**} < \tau_D \)), leaves leader \( D \)’s personal payoff for accepting it to be just \( p - c_D \).

Note that the effect of the public demand (in this range) is essentially to increase \( D \)’s “reservation value” from \( p - c_D \) to \( y^{**} \), thereby reducing the size of the preferred-to-war bargaining range in \( D \)’s favor from \([p - c_D, p + c_S]\) to \([y^{**}, p + c_S]\). Note how this is significantly different from the effect of audience costs in models that treat the disputed good as indivisible. In those models, audience costs can only affect complete-information crisis bargaining if they are large enough to reverse one’s preference between war and living without the disputed good (e.g., Fearon 1997, 78-9), and these analyses find that under incomplete information, audience costs affect crisis bargaining through partially-informative equilibria in which at least some types are generating such large audience costs. In our analysis, \( D \) is dissatisfied and hence begins the interaction with a credible commitment to fight, and audi-
enence costs would therefore have no effect if we treated the disputed good as indivisible. But by treating it as fully divisible, we find that moderate audience costs can generate moderate bargaining leverage, and it will turn out that under incomplete information, this will lead to audience costs affecting crisis bargaining even when they are not so large as to create commitments to fight.

To summarize, as seen by the dashed line in Figure 2, D’s public demand provides bargaining leverage by causing S to offer more than what she otherwise would; the bigger the public demand that D makes, the bigger is S’s offer. However, note that under complete information, S calibrates her offer so that leader D’s personal payoff for accepting the offer (with audience costs) remains at \( p - c_D \) (the solid line in Figure 2); this is the smallest offer that S needs to make to avoid war, and she has no incentive to offer any more. Thus, “bargaining leverage” refers to a greater share of the disputed good for the leader’s country, but not a greater personal payoff for the leader. The leader is not personally better off by using public demands, but citizens who care about achieving a greater share of the disputed good for their country are (which by itself provides a reason for them to impose audience costs on their leader for backing down from public demands; i.e., a microfoundation for imposing audience costs).

Finally, if D’s demand exceeds a threshold denoted by \( \tau_{D_{\text{war}}} = \frac{(p + c_S)(1 + a_D) - p + c_D}{a_D} \) (the “war threshold,” and which is about 0.95 in Figure 2), then D is publicly demanding so much that the preferred-to-war bargaining range is eliminated in the sense that the minimal offer that D now needs to avoid war (\( y^{**} \)) leaves S with less than her payoff for war, and hence S chooses to make an unacceptable offer and allows war to occur.
4 Incomplete-Information Results

We now want to examine the effect of public demands on incomplete-information crisis bargaining, first in a setting where we limit audience costs in a way that they cannot create a commitment to fight where no such commitment previously existed, and then where we allow audience costs to potentially have this commitment-creating effect. To do this, we simply add incomplete information to the previous analysis. We consider a case of one-sided uncertainty, where \( S \) is uncertain about \( D \)'s cost of war \( c_D \) (this implies that \( S \) is uncertain about \( D \)'s payoff for war, or “resolve”).\(^7\) We assume that \( D \)'s cost of war can take on two possible values, i.e., there are two possible types of \( D \). “Nature” chooses \( D \)'s cost to be \( c_{D_l} \) with probability \( s \in (0,1) \) and \( c_{D_h} \) with probability \( 1 - s \), with \( c_{D_l} < c_{D_h} \). Thus, \( c_{D_l} \) is the more resolved type, because his expected utility from war is higher (see Figure 3).

However, in this section we assume that even the low-resolve type begins the interaction being dissatisfied with the status-quo, i.e., we assume that \( q < p - c_{D_h} \). Since each type begins the interaction dissatisfied with the status-quo, neither type can use a public demand to create a commitment to fight where no such commitment previously existed; the commitment is already there. Substantively, this captures a scenario in which \( S \) assigns probability 1 to \( D \) being dissatisfied, but is uncertain about how large a revision of the status-quo she needs to offer to avoid war.

In the model without the option of public demands, the well-known “risk-return tradeoff” result is that if the probability \( (s) \) that \( S \) assigns to \( D \) being the high-resolve type exceeds

\(^7\)Note that Fearon’s (1994), Kurizaki’s (2007), and Schultz’s (1999, 2001a, 2001b) models of audience costs also consider uncertainty about the costs of war, whereas Fearon (1997) considers uncertainty about the valuation of the disputed good.
the threshold \( s^* = \frac{c_{Dh} - c_{Dl}}{c_{Dh} + c_S} \in (0, 1) \), then \( S \) makes the big offer \( y^* = p - c_{Dl} \) that satisfies both types, and war is avoided with certainty. On the other hand, if the probability is less than this threshold, then \( S \) makes the small offer \( y^* = p - c_{Dh} \) that only the low-resolve type accepts, and war occurs with positive probability (namely, with probability \( s \), the probability that \( D \) ends up being the high-resolve type). In this latter scenario, where there is a positive probability of war, we suppose that, upon realizing his type, \( D \) endogenously chooses a public demand \( \tau_D \in [0, 1] \). \( S \) observes this demand, and then makes her offer. In this setting, the public demand may generate bargaining leverage, but may also allow for information-transmission if the two types at least partially separate the public demands that they make.

In the incomplete-information setting, we are interested in the possibility of equilibria in which public demands allow agreement to be reached where otherwise war would occur. The audience cost literature points out that audience costs can have mixed effects on the overall stability of crisis bargaining; on the one hand, their signaling role may allow some agreements to be reached where otherwise war would occur, but on the other hand, they may lock some actors into war, thereby possibly causing a net increase in the probability of war (e.g., Fearon 1994, 586; Fearon 1997, 71; Schultz 1999, 246). In the risk-return setting of our crisis bargaining model (i.e., when \( s < s^* \)), agreement is reached with probability 1 with the low-resolve type, and with probability 0 with the high-resolve type. Therefore, at a minimum we are interested in the possibility of equilibria in which public demands allow agreement to be reached with positive probability with the high-resolve type (when \( s < s^* \)). It will turn out that in all such equilibria, agreement is still being reached with probability 1 with the low-resolve type, and hence the probability of war (\textit{ex ante} as well as \textit{ex post}) is
lower than in the model without public demands. Therefore, it does not matter whether we define our criterion for public demands having a “stabilizing” effect on crisis bargaining as causing agreement to be reached with positive probability with the high-resolve type, or as the more demanding criterion of reducing the probability of war, as these two criteria are effectively equivalent. We therefore use the more modest criterion, and define the following:

**DEFINITION**: A “stabilizing PBE” is a perfect Bayesian equilibrium (PBE) of the public demand model in which agreement is reached with positive probability with the high-resolve type. (More formally, it is a PBE in which there is at least one \( \tau_D \) that the high-resolve type chooses with positive probability such that upon observing that \( \tau_D \), with positive probability \( S \) makes an offer that the high-resolve type would accept.)

In general, the PBE of a signaling model can be characterized as either separating (the different types of the informed actor choose completely different signals), pooling (they choose the exact same signal), or semi-separating (they choose some common signals, and some different ones). In a pooling equilibrium, no information is transmitted to the uninformed actor, perfect information-transmission occurs in a separating equilibrium, and partial information-transmission is achieved in a semi-separating equilibrium.

We have the following result, which follows directly from propositions and corollaries established in the appendix. We consider the possibility of PBE in pure as well as mixed strategies.

**RESULT 1**: Stabilizing PBE exist if and only if \( S \)'s prior belief satisfies the condition 
\[
s \geq \frac{c_{D_h} - c_{D_l}}{c_{D_h} + c_{S} + a_{D}(c_{D_l} + c_{S})} \in (0, s^*)\).
\]
These equilibria can be pooling or semi-separating, but not fully-separating.

That fully-separating stabilizing PBE do not exist is not surprising, as they generally
do not exist in other formal analyses of audience costs as well. Rather, these analyses find that audience costs exert their effects through partial information-transmission via semi-separating equilibria. That pooling stabilizing PBE exist is surprising, and hence we now describe the logic of these equilibria, and then turn to the semi-separating equilibria, which do not differ from the pooling equilibria in any fundamental way. We then discuss the condition on the prior belief under which stabilizing PBE exist.

How can pooling stabilizing PBE exist? Their existence is due to what we call the “threshold-lowering effect” of a moderate public demand. To see how this effect occurs, first recall that in the model without public demands, $S$ chooses to make the big offer $(p - c_{D_h})$ that both types accept when her prior belief (denoted by $s$) that $D$ is the high-resolution type exceeds a certain threshold (denoted by $s^*$); otherwise she chooses to make the small offer $(p - c_{D_h})$ that only the low-resolution type accepts. How is this threshold $s^*$ determined? If $S$ makes the big offer of $p - c_{D_h}$, it is accepted for sure, and hence $S$’s payoff is $1 - (p - c_{D_h})$. If she makes the small offer of $p - c_{D_h}$ instead, only the low-resolution type accepts it; the high-resolution type rejects it and goes to war, and hence $S$’s expected payoff for making the small offer is $s(1 - p - c_{S}) + (1 - s)(1 - (p - c_{D_h}))$. Therefore, the threshold $s^*$ is determined by solving the inequality $1 - (p - c_{D_h}) > s(1 - p - c_{S}) + (1 - s)(1 - (p - c_{D_h}))$ for $s$. In particular, note that as the difference between the big and small offers increases (e.g., as $c_{D_h} - c_{D_l}$ increases), $S$ is more tempted to make the small offer, resulting in the threshold $s^* = \frac{c_{D_h} - c_{D_l}}{c_{D_h} + c_{S}}$ becoming larger. On the other hand, as the difference between the large and small offers becomes smaller, $S$ is more willing to make the large offer, resulting in the threshold $s^*$ becoming smaller.

Now suppose that in the public demand model, both types of $D$ pool and choose some
\[ \tau_D \leq p - c_{D_h} \] (Figure 3 is useful in following this argument). At that small demand level, neither type of \( D \) has publicly demanded more than his war-payoff, and hence type \( c_{D_h} \) still needs just \( p - c_{D_h} \) to avoid war, and type \( c_{D_l} \) still needs just \( p - c_{D_l} \) to avoid war. Hence, at that demand level, the big and small offers that \( S \) considers making are the same as in the model without public demands, and therefore \( S \)'s belief-threshold for making the large offer is the same \( s^* \) that it is in the model without public demands. However, now suppose that both types of \( D \) choose some \( p - c_{D_h} < \tau_D \leq p - c_{D_l} \). At that larger demand level, type \( c_{D_l} \) has still not demanded more than his war-payoff \((p - c_{D_l})\), and hence still just needs \( p - c_{D_l} \) to avoid war. Thus, the big offer that \( S \) might make (if she wants to avoid war with certainty) remains the same as before. However, type \( c_{D_h} \) is now publicly demanding more than his war-payoff \((p - c_{D_h})\), and hence he now needs \( y^{**}(c_{D_h}) \) to avoid war, where \( y^{**}(c_{D_h}) \) is the offer that (with audience costs) makes leader \( D \)'s personal payoff for accepting it to be \( p - c_{D_h} \). This offer is of course larger than \( p - c_{D_h} \) (but smaller than \( p - c_{D_l} \)). Hence, the small offer (if she wants to just satisfy the low-resolve type) is now larger than before. Because the large offer is the same as before but the small offer has increased, \( S \)'s belief-threshold (that we denote by \( s_{critical} \)) for making the large offer is below the earlier threshold of \( s^* \). This is seen in Figure 4, which plots \( s_{critical} \) as a function of \( \tau_D \). As \( \tau_D \) increases within the range \( p - c_{D_h} < \tau_D \leq p - c_{D_l} \), \( s_{critical} \) strictly decreases, because the small offer \( y^{**}(c_{D_h}) \) is strictly increasing (whereas the large offer \( p - c_{D_l} \) remains the same). Once \( \tau_D \) exceeds \( p - c_{D_l} \), then the public demand also exceeds type \( c_{D_l} \)'s war-payoff, and hence he now needs \( y^{**}(c_{D_l}) \) to avoid war, where \( y^{**}(c_{D_l}) \) is the offer that (with audience costs) makes leader \( D \)'s personal payoff for accepting it to be \( p - c_{D_l} \). This offer is of course larger than \( p - c_{D_l} \), and hence the large offer is becoming larger than it was before. The small offer is still increasing as well.
and the net effect is that $s_{\text{critical}}$ starts strictly increasing. By the time $\tau_D$ has increased to $p + c_S$ (which is 0.6 in Figure 4), $s_{\text{critical}}$ has increased to $s^*$ again. As $\tau_D$ increases even more, $s_{\text{critical}}$ increases above $s^*$. Finally, as $\tau_D$ approaches the war-threshold of the high-resolve type (that we denote by $\tau_{D_{\text{war}}}(c_{D_l})$), which is 0.75 in Figure 4, $s_{\text{critical}}$ approaches 1.8

Thus, in the public demand model, the threshold for making the big offer, $s_{\text{critical}}$, is strictly below the threshold in the model without public demands, $s^*$, for a range of moderate public demand levels, namely $p - c_{D_h} < \tau_D < p + c_S$. We call this the “threshold-lowering effect” of moderate public demands. Note that this effect is just an extension of the bargaining-leverage role of public demands to the incomplete-information setting: a moderate public demand causes $S$ to have to compromise more with the low-resolve type than the high-resolve type (in relative terms, i.e., relative to the offers in the model without public demands; note that the big offer is still larger than the small offer), and this causes the belief-threshold to decrease.

This threshold-lowering effect allows for the existence of pooling, and hence completely uninformative, equilibria in which agreement is reached with probability 1 with the high-resolve type even when $s < s^*$. For example, for the parameter values in Figure 4, $s^* = 0.4$. Suppose that $S$’s prior belief is $s = 0.3 < s^*$. Then, in the model without public demands, we are in the risk-return setting where $S$ makes the small offer and war occurs if $D$ ends up being the high-resolve type. Suppose that in the model with public demands, both types of $D$ choose some $\tau_D$ such that, at that demand level, the critical belief-threshold is less than the prior 0.3 (i.e., $s_{\text{critical}}(\tau_D) < 0.3$; as seen in Figure 4, there is a continuum of such $\tau_D$, and any will do). Then, upon observing this $\tau_D$, $S$’s posterior belief remains at her prior belief

---

8For $\tau_D > \tau_{D_{\text{war}}}(c_{D_l})$, $s_{\text{critical}}$ is undefined, because $S$ never chooses to make the big offer.
since no updating of beliefs has occurred, but since the prior/posterior now exceeds the critical threshold (i.e., $s > s_{\text{critical}}(\tau_D)$), she now makes the big offer (that results in leader $D$’s personal payoff for accepting it to be $p - c_{D_l}$), and hence war is avoided with certainty.\(^9\)

In this pooling equilibrium, information-transmission obviously has nothing to do with why a public demand allows agreement to be reached with the high-resolve type in a situation where otherwise war would occur: this is entirely due to the threshold-lowering effect of a moderate public demand. Note that we can construct semi-separating equilibria as well, in which agreement is reached with probability 1 with the high-resolve type. For example, in the pooling equilibrium described above, specify that the high-resolve type makes an alternative demand level (some $\tau'_D$) with some very small probability $\epsilon > 0$. $S$’s posterior belief will be slightly below the prior belief $s$ upon the common (pooling) demand level ($\tau_D$) being observed, but will still (for $\epsilon$ small enough) exceed the critical threshold at that demand level (recall that we stipulated that the pooling demand level satisfies $s_{\text{critical}}(\tau_D) < s$), and hence $S$ will still make the big offer upon observing the common (pooling) demand level. Thus, the low-resolve type will not have an incentive to deviate to $\tau'_D$ (and again, the off-the-equilibrium-path beliefs can be anything for this to be a PBE).

Note that in this semi-separating equilibrium, perfect information-transmission is achieved when $S$ observes the demand level that only the high-resolve type makes with positive probability ($\tau_D$), and hence it may appear that the stabilizing effect of public demands is due

\(^9\)In addition, note that the off-the-equilibrium-path beliefs (if $S$ observes some non-equilibrium $\tau'_D$) for this to be a PBE can be anything, because $S$ is already making the big offer (with certainty) upon observing $\tau_D$, which is the best possible offer she can be making upon observing any demand level (i.e., she never makes an offer that leaves leader $D$’s personal payoff for accepting it to be greater than $p - c_{D_l}$), and hence neither type of $D$ has an incentive to deviate to some non-equilibrium demand level.
to information-transmission. However, the only reason that this type of semi-separating stabilizing PBE can exist is because $S$ is making the big offer upon observing the common (pooling) demand level $\tau_D$, and $S$ is doing so because of the threshold-lowering effect. If the threshold was not being lowered at $\tau_D$, then $S$ would make the small offer upon observing $\tau_D$ (recall that we are considering the risk-return setting of $s < s^*$, so that $S$ is inclined to make the small offer), and hence the low-resolve type would strictly benefit by deviating to $\tau'_D$, and the equilibrium would break down. Thus, even in this semi-separating equilibrium, and not just in the pooling equilibrium characterized earlier, agreement is being reached with the high-resolve type because of the threshold-lowering effect, not because of information-revelation. The threshold-lowering effect is doing all of the hard work, and information-revelation is incidental.

In fact, we have the following result.

**RESULT 2**: In every stabilizing PBE, at least one threshold-lowering common (pooling) demand level is being made. (More formally, there exists at least one $\tau_D$ that is being chosen with positive probability by both types such that $s_{\text{critical}}(\tau_D) < s^*$.)

This result establishes that the threshold-lowering effect is a necessary condition for agreement to be reached with positive probability with the high-resolve type. On the other hand, information-revelation is not a necessary condition, as proven simply by the existence of the pooling equilibrium described above.

The following result makes a similar point.

**RESULT 3**: In every stabilizing PBE, the probability with which $S$ makes the big offer upon observing any demand level that the low-resolve type makes with positive probability is at least as large as the probability with which $S$ makes the big offer upon observing any
demand level that the high-resolve type makes with positive probability.

These two results make it hard to argue that the stabilizing effect of public demands is in any way due to information-transmission, when both types of $D$ begin the interaction with credible commitments to fight.

Finally, recall from Result 1 that stabilizing PBE exist if and only if the prior satisfies $s \geq \frac{c_{D_l} - c_{D'_l}}{c_{D_h} + c_S + a_D(c_{D'_l} + c_S)} \in (0, s^*)$. This can be graphically seen from Figure 4. In Figure 4, the critical belief-threshold ($s_{critical}$) reaches a minimum when $\tau_D = p - c_{D_l}$, and this minimum is $s_{critical,min} = \frac{c_{D_h} - c_{D'_l}}{c_{D_h} + c_S + a_D(c_{D'_l} + c_S)}$. If the prior belief $s$ is below this minimum, then even if $D$ chooses the demand level that minimizes the threshold, $S$ still chooses to make the small offer upon observing that demand level, and hence stabilizing PBE cannot exist. Public demands can help stabilize crisis bargaining, but not when $S$ begins the interaction very confident that she faces the low-resolve type.

However, note that in the model without public demands, $S$ makes the large offer only if the prior belief satisfies $s \geq s^* = \frac{c_{D_h} - c_{D'_l}}{c_{D_h} + c_S}$, which is a strictly higher threshold than $\frac{c_{D_h} - c_{D'_l}}{c_{D_h} + c_S + a_D(c_{D'_l} + c_S)}$. Thus, allowing for public demands enlarges the range of priors for which peaceful equilibria exist, which provides further support for the notion that public demands can help stabilize crisis bargaining. Also note that $\frac{c_{D_h} - c_{D'_l}}{c_{D_h} + c_S + a_D(c_{D'_l} + c_S)}$ strictly decreases as $a_D$ increases, and that $\frac{c_{D_h} - c_{D'_l}}{c_{D_h} + c_S + a_D(c_{D'_l} + c_S)}$ approaches 0 as $a_D$ approaches infinity. Hence, increasing the audience cost coefficient increases the range of priors for which peaceful equilibria exist, suggesting that greater audience-cost-susceptibility increases the “range” for peace. But for any given value of $a_D$, no matter how high, there exists a range of priors for which no stabilizing PBE exists; if $S$ begins the interaction very confident that she faces the low-resolve type, then public demands cannot lower the belief-threshold enough to cause her to
make an offer suitable for the high-resolve type. Public demands can help stabilize crisis bargaining, but not for all values of the prior.

5 Robustness to Alternative Bargaining Protocols

Our analysis has been based on a model in which $S$ makes a take-it-or-leave-it (TILI) offer, and hence has all of the bargaining leverage in the sense that she just needs to offer $D$ his payoff from war and gets to keep the remaining surplus for herself. Before discussing the robustness of the results to alternative bargaining protocols, we point out that good reasons can be given for why in crisis bargaining between a satisfied and a dissatisfied leader, the satisfied leader should generally have most of the bargaining leverage.\textsuperscript{10} Because it will be the dissatisfied leader that ultimately starts a war, the satisfied leader just needs to offer enough (or slightly more, to be on the safe side) to make the dissatisfied leader indifferent between accepting the offer and going to war. The dissatisfied leader cannot credibly claim that “I will go to war unless you offer me significantly more than my war payoff” or “I will go to war unless you accept this proposal of mine that offers me significantly more than my war payoff” (and if he is even slightly risk-averse, he strictly prefers an offer of his war-payoff over going to war). In addition, if the disputed good is territory, we can imagine the two sides having their troops manning the status-quo division of the territory, and the satisfied leader just needs to withdraw her troops to the point where the dissatisfied leader is indifferent between accepting the new division and going to war. The dissatisfied leader cannot credibly demand a greater withdrawal. By the same logic, under incomplete information the satisfied leader never has an incentive to offer more than (or withdraw her troops further back from) the war-payoff of the highest-resolve type of dissatisfied leader that she could be facing.

\textsuperscript{10}Our thanks to Robert Powell for suggesting these arguments.
Nevertheless, we now want to consider whether our results are robust to some alternative bargaining protocols. The results are discussed in more detail in the appendix, and are summarized here. First consider a bargaining protocol in which $D$ gets to make a TILI offer. Now $D$ has all of the bargaining leverage, and hence a public demand cannot provide additional bargaining leverage (under complete or incomplete information). In addition, the probability of war is 0 even in the model without public demands, and public demands hence have no role in reducing the probability of war.

Now suppose that “nature” chooses $S$ to make a TILI offer with probability $\pi \in (0, 1)$, and $D$ gets to make a TILI offer with probability $1 - \pi$. This captures the full range of intermediate bargaining leverage in expectation. For $\pi$ close to 1, $S$ has almost all of the bargaining leverage in expectation, and $D$ has almost all of the bargaining leverage in expectation when $\pi$ is close to 0. Any other amount of expected bargaining leverage is captured by setting $\pi$ at the appropriate value. For any value of $\pi \in (0, 1)$, all of the earlier results continue to hold, and the only change is that now neither type of $D$ publicly demands more than $p + c_S$ in equilibrium (which is a possibility earlier, as in the semi-separating PBE described above if $\tau'_D > p + c_S$). If $D$ is chosen to make the TILI offer, he proposes dividing the good at $(p + c_S, 1 - p - c_S)$, and hence pays an audience cost in his own proposal if $\tau_D > p + c_S$. This causes the high-resolve type to never have an incentive to choose some $\tau_D > p + c_S$, which causes the low-resolve type to never have an incentive to do so as well. Thus, any $\tau_D > p + c_S$ is now ruled out in equilibrium, but because the threshold-lowering range is $p - c_{Dh} < \tau_D < p + c_S$, the threshold-lowering range remains fully intact, as well as all of the stabilizing PBE based on it in which no $\tau_D > p + c_S$ is chosen with positive probability (and this restriction is not consequential in any way).
6 Audience Costs Creating a Commitment to Fight

We now want to examine what happens when we allow audience costs to be so significant as to be able to create a commitment to fight where no such commitment previously existed, as in other analyses of audience costs. We first briefly discuss the results under complete information, and then under incomplete information.

6.1 Complete-Information Results

Going back to the ultimatum-offer model with complete information, suppose that $D$ and $S$ both begin the interaction satisfied with the status-quo, i.e., $p - c_D < q < p + c_S$. In the model without public demands, $S$ would simply propose the status-quo $y^* = q$, which $D$ accepts. Because both players are satisfied with the status-quo, the status-quo is not revised in equilibrium.

Now allow $D$ to make a public demand $\tau_D \in [0, 1]$ before $S$ makes her offer. If $D$ chooses some $\tau_D \leq q$, then $S$ continues to propose $y^* = q$. Now suppose $\tau_D > q$. Now leader $D$’s personal payoff for the status-quo is $q - a_D(\tau_D - q)$ rather than $q$. If $\tau_D > q$ is small enough that $q - a_D(\tau_D - q) \geq p - c_D$, which simplifies to $\tau_D \leq \frac{a(1 + a_D) - p + c_D}{a_D}$, then $D$ still prefers the status-quo (even with the audience costs paid) to war, and hence is still satisfied with the status-quo. Therefore, for $q < \tau_D \leq \frac{a(1 + a_D) - p + c_D}{a_D}$, $S$ still offers $y^* = q$, which $D$ accepts.

However, when $\tau_D > \frac{a(1 + a_D) - p + c_D}{a_D}$, then $D$ now prefers war to the status-quo, i.e., $D$ is now dissatisfied. Therefore, we call this right-hand-side the “dissatisfaction threshold” $\tau_{D_{diss}}$. If $D$ chooses some $\tau_D > \tau_{D_{diss}}$, then he is creating a commitment to fight unless the status-quo is sufficiently revised in his favor in a situation where he previously lacked that commitment, i.e., he endogenously makes himself dissatisfied. This is the divisible-
good equivalent of using audience costs to create a commitment to stand firm at the final
decision-node in other models of audience costs (e.g., Fearon 1997, 78).

If $D$ chooses some $\tau_D > \tau_{D_{\text{diss}}}$, then $S$ offers $y^{**} = \frac{p-c_D+a_D\tau_D}{1+a_D}$, which is the offer that, with audience costs (i.e., it is the case that $y^{**} < \tau_D$), leaves leader $D$’s personal payoff for accepting it to be just $p - c_D$. This is the minimal offer that $S$ needs to offer $D$ when $D$ makes himself dissatisfied. It is of course the case that $y^{**} > q$; the status-quo is revised in $D$’s favor when $D$ makes himself dissatisfied, but leader $D$’s personal payoff is just his war-payoff ($p - c_D$). These results are graphically shown in Figure 5, which plots country $D$’s share of the pie (i.e., $S$’s offer) as well as leader $D$’s personal payoff as a function of $\tau_D \in [0, 1]$.

Finally, when $D$’s public demand exceeds a threshold denoted by $\tau_{D_{\text{war}}} = \frac{(p+c_S)(1+a_D) - p+c_D}{a_D}$ (the “war threshold,” and which is about 0.92 in Figure 5), then $D$ is publicly demanding so much that the preferred-to-war bargaining range is eliminated in the sense that the minimal offer ($y^{**}$) that $D$ needs to avoid war now leaves $S$ with less than her payoff for war, and hence $S$ chooses to make an unacceptable offer and allows war to occur.

Note from Figure 5 that under complete information, $D$ never chooses to make himself dissatisfied, in fact he never chooses any $\tau_D > q$. If he chooses some $q < \tau_D \leq \tau_{D_{\text{diss}}}$, then he continues to receive the status-quo but needlessly pays an audience cost. If he chooses some $\tau_{D_{\text{diss}}} < \tau_D \leq \tau_{D_{\text{war}}}$, then he gets a bigger offer than the status-quo, but his personal payoff is $p - c_D$, which is worse than the original status-quo. In contrast to indivisible-good models in which it is rational to use audience costs to create a commitment to fight if the other side lacks that commitment (e.g., Fearon 1997, 78-9), when the disputed good is divisible and the other side can choose how much to offer, it is not rational to do so. However, it will turn
out that under incomplete information, it can be rational for a satisfied type to choose some \( \tau_D > q \), because this can cause \( S \) to make a big offer (suitable for the high-resolve type), either because of the threshold-lowering effect or because of information-transmission.

### 6.2 Incomplete-Information Results

Now suppose again that there are two types of \( D \), \( c_{D_l} \) and \( c_{D_h} \), and only the high-resolve type (\( c_{D_l} \)) is dissatisfied, i.e., \( p - c_{D_h} < q < p - c_{D_l} \). Then the low-resolve type can potentially use audience costs to create a commitment to fight where previously he lacked that commitment. Substantively, this captures a scenario in which \( S \) is uncertain about whether \( D \) is really dissatisfied, and hence whether she really needs to offer a revision of the status-quo to avoid war.\(^{11}\) In the model without public demands, the standard risk-return tradeoff result is that \( S \) chooses to make the big offer of \( p - c_{D_l} \), which both types accept, if \( s > s^{**} = \frac{(p-c_{D_l})-q}{(p+c_S)-q} \in (0,1) \), and chooses to make the small offer of \( q \), which only type \( c_{D_h} \) accepts, if \( s < s^{**} \). In the scenario where \( s < s^{**} \), we are again interested in the possibility of PBE of the public demand model in which agreement is reached with positive probability with the high-resolve type.

We have the following result.

**RESULT 4:** The threshold-lowering effect exists (i.e., there exist \( \tau_D \) for which \( s_{critical}(\tau_D) < s^{**} \)) if and only if \( a_D > \frac{q-(p-c_{D_h})}{(p-c_{D_l})-q} \ (> 0) \).

When the threshold-lowering effect exists, namely when the audience-cost coefficient exceeds a certain threshold, then there exist stabilizing PBE of exactly the same nature as in the case where both types are dissatisfied to begin with. That is, there exist pooling as well as semi-separating stabilizing PBE as long as the prior satisfies \( s \geq \frac{c_{D_h}-c_{D_l}}{c_{D_h}+c_S+\theta_{D}(c_{D_l}+c_S)} \in \)

\(^{11}\)Powell (1999) refers to this as \( D \) being “potentially dissatisfied” *ex ante.*
(0, s**), i.e., as long as S does not begin the interaction extremely confident that she faces the low-resolve type. As before, although there exist semi-separating, and hence partially-informative, stabilizing PBE, agreement is being reached with positive probability with the high-resolve type because of the threshold-lowering effect rather than information-transmission (which is just incidental).

When it is not the case that $a_D > \frac{q-(p-c_{D_h})}{(p-c_{D_l})-q}$, do stabilizing PBE still exist, i.e., can public demands still allow agreement to be reached with positive probability with the high-resolve type? We have the following result.

RESULT 5: For any $a_D > 0$, there exist semi-separating stabilizing PBE in which information-transmission is genuinely the reason for why agreement is reached with positive probability with the high-resolve type. These PBE have the following form. Type $c_{D_h}$ chooses some small demand $\tau_{D_h} \leq q$ with positive probability less than 1, and some larger demand $\tau_{D_p} > q$ with the remaining probability. Type $c_{D_l}$ chooses $\tau_{D_p}$ with probability 1 (hence, the subscript $p$ stands for “pooling”). Upon observing $\tau_{D_p}$, S mixes (randomizes) between the large offer (suitable for the high-resolve type at that demand level) and the small offer (suitable for the low-resolve type at that demand level) with such probability that $c_{D_h}$ is indifferent between $\tau_{D_h}$ and $\tau_{D_p}$.

These equilibria are similar to the partially-informative PBE that exist in other audience cost models, in that upon observing the large demand $\tau_{D_p}$, S’s belief that she faces the high-resolve type increases, and this is the reason that she makes the large offer with positive probability in a situation where her prior inclination was to make the small offer with certainty. That is, these are stabilizing PBE in which information-transmission is genuinely the reason for why agreement is reached with positive probability with the high-resolve type,
in contrast to the stabilizing PBE based on the threshold-lowering effect.

To summarize, when only the high-resolve type begins the interaction dissatisfied, then stabilizing PBE based on the threshold-lowering effect exist if and only if (i) \( a_D > \frac{q-(p-c_{Dh})}{(p-c_{Dl})} > 0 \) and (ii) \( s \geq \frac{c_{Dh}-c_{Dl}}{c_{Dh}+c_{Dl}+a_D(c_{Dh}+c_{S})} \in (0, s^*) \). Stabilizing PBE based on information-transmission (described in Result 5) exist for all values of \( a_D > 0 \) and all values of the prior \( s \). It should be noted that when conditions (i) and (ii) are satisfied, then stabilizing PBE based on the threshold-lowering effect can be constructed in which agreement is reached with the high-resolve type with probability 1, and hence the probability of war is 0. In the stabilizing PBE based on information-transmission, on the other hand, agreement is reached with the high-resolve type with maximum probability \( \frac{q-(p-c_{Dh})}{c_{Dh}-c_{Dl}} \in (0, 1) \), and hence there is still a positive probability of war.\(^{12}\)

We now consider whether the low-resolve type is using audience costs to create a commitment to fight (i.e., endogenously making himself dissatisfied) in each type of stabilizing PBE. Recall that in the earlier analysis where both types of \( D \) begin the interaction dissatisfied, and hence the parameters of the model are such that no type has the opportunity to use audience

\(^{12}\)As discussed in the appendix, when \( a_D \geq \frac{c_{S}+c_{Dl}}{1-p-c_{S}} \), then fully-separating stabilizing PBE also exist in which information-revelation is the reason that agreement is reached with positive probability with the high-resolve type. In these equilibria, \( c_{Dh} \) chooses some \( \tau_{Dh} \leq q \) with probability 1, and \( c_{Dl} \) chooses \( \tau_{Dl} = \tau_{Dwar}(c_{Dl}) \) with probability 1. That is, the high-resolve type all but eliminates the preferred-to-war bargaining range, at which point \( S \) is indifferent between making an acceptable and unacceptable offer. In the PBE, she makes the large offer with positive probability (which is why it is a stabilizing PBE) and the small offer with high enough probability to deter the low-resolve type from deviating to \( \tau_{Dl} \). Although these are fully-separating PBE, agreement is reached with the high-resolve type with at most the same probability as in the semi-separating stabilizing PBE of Result 5, and hence they are not fundamentally different from these PBE.

27
costs to create a commitment to fight, we found that public demands can help stabilize crisis bargaining, but only through the threshold-lowering effect, and not through information-transmission. That is, those were stabilizing PBE based on the threshold-lowering effect in which no type is using audience costs to create a commitment to fight where no such commitment previously existed. In the current analysis, where the low-resolve type begins the interaction satisfied, and hence can use audience costs to create a commitment to fight, we found that public demands can help stabilize crisis bargaining through the threshold-lowering effect (under some conditions) as well as through information-transmission. It turns out that in this scenario, every threshold-lowering public demand \( \tau_D \) satisfies \( \tau_D > \tau_{D_{\text{diss}}} (c_{D_h}) \), and hence these are stabilizing PBE based on the threshold-lowering effect in which the low-resolve type is using audience costs to create a commitment to fight where no such commitment previously existed. Stabilizing PBE based on information-revelation (described in Result 5) can be constructed in which the high common (pooling) demand level satisfies \( q < \tau_{D_p} \leq \tau_{D_{\text{diss}}} (c_{D_h}) \), and hence the low-resolve type is not creating a commitment to fight, as well as in which it satisfies \( \tau_{D_p} > \tau_{D_{\text{diss}}} (c_{D_h}) \), and hence the low-resolve type is creating a commitment to fight. In contrast to earlier work on audience costs, we find that audience costs can have stabilizing effects on crisis bargaining through a non-informative, threshold-lowering (bargaining leverage) effect, and that audience costs can have stabilizing effects on crisis bargaining, whether through the threshold-lowering or information-revealing mechanism, even without such large audience costs being generated as to create a commitment to fight where no such commitment previously existed.\(^{13}\)

Finally, we consider why stabilizing PBE based on information-revelation exist when

\(^{13}\)Also note that in the fully-separating stabilizing PBE described in the previous footnote, the low-resolve type is not creating a commitment to fight.
the low-resolve type begins the interaction satisfied, but not when he begins dissatisfied. Stabilizing PBE based on information-revelation depend on there being some risk that deters the low-resolve type from fully mimicking the behavior of the high-resolve type. When the low-resolve type begins the interaction dissatisfied, the small offer that $S$ is inclined to make (in the risk-return scenario where $s < s^*$) is just the low-resolve type’s payoff for war, which the low-resolve type can still achieve by just going to war if mimicking the high-resolve type’s public demand does not achieve the desired end of getting $S$ to make the large offer. That is, mimicking is risk free, and this prevents the existence of stabilizing PBE based on information-revelation. When the low-resolve type begins the interaction satisfied, on the other hand, mimicking now carries a risk. If mimicking does not get $S$ to make the large offer, the low-resolve type of leader $D$ gets a personal payoff of either $q - a_D(\tau_{D_p} - q)$ (if the mimicking demand level satisfies $q < \tau_{D_p} \leq \tau_{Diass}(c_{D_h})$) or $p - c_{D_h}$ (if it satisfies $\tau_{D_p} > \tau_{Diass}(c_{D_h})$), both of which are worse than the original status-quo. Hence, stabilizing PBE based on information-revelation exist in which the high-resolve type chooses some $\tau_{D_p} > q$, which the low-resolve type finds risky to mimic and hence does not fully mimic, allowing for information-transmission that gets $S$ to make the big offer with positive probability.

7 Conclusion

The question of whether and how leaders can credibly signal their private information in international crises, thereby possibly overcoming informational asymmetries that can lead to costly and inefficient war, is an important topic in the study of international conflict. In a seminal article, Fearon (1994) argued that generating audience costs through public
threats/demands is a primary mechanism by which leaders can credibly signal private information, and this theory has proven very influential, with a large theoretical and empirical literature now dealing with the topic of audience costs. Existing theoretical work on audience costs finds that they affect crisis bargaining through information-revealing equilibria in which at least some types of the informed actor are generating such large audience costs as to create a commitment to fight where no such commitment previously existed. This has been taken to imply that audience costs exert their effect through information-transmission and that only large audience costs can affect crisis bargaining. In part because of these theoretical results, the question of whether leaders can generate large audience costs is central to recent empirical research that uses survey experiments to examine audience cost theory.

In this paper, we consider a divisible-good, endogenous-offer crisis bargaining model with endogenous-magnitude public demands, and find in addition to the above (standard) type of equilibria, (a) informative equilibria in which such large audience costs are not being generated, (b) non-informative bargaining-leverage (based on the threshold-lowering effect) equilibria in which such large audience costs are not being generated, and (c) non-informative bargaining-leverage equilibria in which such large audience costs are being generated. To summarize our main findings, we find that (i) audience costs can affect incomplete-information crisis bargaining through a non-informative, bargaining-leverage effect, and that (ii) audience costs can affect such crisis bargaining even when they are not so significant as to create a commitment to fight where no such commitment previously existed.

How can we empirically examine these theoretical results? The results are what are often called “existence results” in the game-theoretic literature, i.e., we establish the existence of new types of equilibria, some with new causal mechanisms, in which audience costs affect
crisis bargaining. Thus, if there are empirical cases that are consistent with any of the new
types of equilibria discussed in the previous paragraph, that would be empirical evidence
supportive of our results.

Although a full empirical analysis is beyond the scope of this paper, a number of promi-
nent historical episodes seem to be consistent with equilibria of type (c), i.e., non-informative
bargaining-leverage equilibria in which large audience costs are being generated. For exam-
ple, according to U.S. Attorney General Robert Kennedy’s memoir of the 1962 Cuban Mis-
side Crisis, during the course of internal administration deliberations, U.S. President John
Kennedy speculated about the likely Soviet reaction to a U.S. attack on Cuba, including
airstrikes and/or a U.S. invasion. When Air Force Chief of Staff General Curtis LeMay
responded that there would be no Soviet response, the president was skeptical, stating that
“They, no more than we, can let these things go by without doing something. They can’t,
after all their statements, permit us to take out their missiles, kill a lot of Russians, and
then do nothing. If they don’t take action in Cuba, they certainly will in Berlin” (Kennedy
1999, 28-9; emphasis added). This statement does not convey the sense that information-
transmission occurred, as would be an inference along the lines of “The fact that Soviet
leaders made strong public pronouncements indicates that they are probably more resolved
than we initially believed them to be.” Rather, it is consistent with our bargaining-leverage
argument, in that the Soviet statements created a situation where they had to respond to
a U.S. attack on Cuba or face significant domestic and/or international audience costs, re-
gardless of the initial level of resolve. This caused the U.S. to be more conciliatory in its
policy choice; Kennedy opted for a blockade rather than airstrikes.\footnote{Robert Kennedy
goes on to write of the president: “What guided all his deliberations was an effort not to
disgrace Khrushchev, not to humiliate the Soviet Union... This was

31
Similarly, in the 1898 Fashoda Crisis between Britain and France, British leaders generated large potential audience costs by publicly releasing the dispatches between the two countries and belligerently stating in public that they would not compromise at all on the issue, and the only choice facing France was to entirely withdraw its troops from Fashoda or face war. The French ambassador in London, Baron de Courcel, stated to his superiors: “It seems that, with this haughty language, the English government will cut itself off from all retreat, and that it will be impossible for it to back down from demands made in such a manner…Lord Salisbury has entrenched himself in English public opinion thereby preventing his government from negotiating as long as French forces occupy Fashoda” (Schultz 2001a, 41-2). Again, this statement does not convey the sense that French leaders updated their beliefs about the pre-existing resolve of British leaders, but rather that they felt that the English government had created a situation where it could not now back down regardless of what its initial level of resolve was, consistent with our bargaining-leverage argument. Ultimately, France chose to withdraw its forces from Fashoda rather than risk war. These cases suggest the plausibility of our result about the non-informative, bargaining-leverage mechanism by which audience costs can help stabilize crisis bargaining, and indicate that why he was so reluctant to stop and search a Russian ship; this was why he was so opposed to attacking the missile sites. The Russians, he felt, would have to react militarily to such actions on our part. Thus the initial decision to impose a quarantine rather than to attack; our decision to permit the Bucharest to pass; our decision to board a non-Russian vessel first; all these and many more were taken with a view to putting pressure on the Soviet Union but not causing a public humiliation…During our crisis talks, he kept stressing the fact that we would indeed have war if we placed the Soviet Union in a position she believed would adversely affect her national security or cause such public humiliation that she lost the respect of her own people and countries around the globe…‘I am not going to push the Russians an inch beyond what is necessary’ ” (Kennedy 1999, 95-8).
more empirical work is needed to better understand the mechanisms by which leaders can reduce the likelihood of costly war, whether through information-revelation or bargaining leverage.

8 Appendix

Due to space constraints, the appendix with the full formal results and proofs is provided as a supplementary online file.
References


Figure 2: $D$’s Share of the Pie and Personal Payoff as a Function of His Public Commitment/Demand $\tau_D \in [0,1]$, when $D$ is Dissatisfied
Figure 4: $S$’s Critical Belief Threshold ($s_{critical}$) as a Function of $D$’s Public Commitment/Demand Level ($\tau_D$)
Figure 5: D’s Share of the Pie and Personal Payoff as a Function of His Public Commitment/Demand $\tau_D \in [0, 1]$, when $D$ Begins the Interaction Satisfied.
preferred-to-war bargaining range

Figure 1: The Bargaining Range

Figure 3: Uncertainty About $D$’s Cost of War