A General Framework for International Conflict under Incomplete Information

Mark M. Dekker*[1, 4], Dawid Walentek*[2], Jonas M. B. Haslbeck[3], and Joris Broere[4]

[1]Department of Information and Computing Sciences, Utrecht University
[2]Amsterdam Institute for Social Science Research, University of Amsterdam
[3]Psychological Methods Group, University of Amsterdam
[4]Centre for Complex Systems Studies, Utrecht University

* Shared first authorship

Abstract

The crisis bargaining literature identifies the incentives and constraints that states face in an international conflict. It shows that the size of and the relation between economic cost resulting from engaging in a coercion and domestic audience cost resulting from issuing empty threats determines states' prospects to succeed at the threat stage and engage in coercion. This article further develops the crisis bargaining framework by providing a specification of uncertainty about the economic and domestic audience cost. Specifically, we introduce uncertainty by modelling economic and domestic audience costs as probability distributions, instead of fixed values as in earlier studies. We develop our argument with the use of a game theory model, expanding the current crisis bargaining framework. In contrast to earlier work, we find that an increase in uncertainty decreases the effectiveness of threats of coercion and decreases the probability of engagement in coercion.

1   Introduction

Conflict between states comes at a great cost (Pinker, 2011; Goldstein, 2011; Pond, 2017; Neuenkirch and Neumeier, 2016), motivating extensive research into the conditions for the onset of war or economic sanctions (Chandler, 2013; Gartzke, 2007; Bueno de Mesquita et al., 1999; Pinker, 2011; Maoz and Russett, 1993; Dixon, 1994; Hufbauer et al., 2007). To this end, scholars have developed a large number of theories to understand how conflict in international relations develops. In this article, we engage with the crisis bargaining framework — one of the most popular approaches to the onset of war or sanctions (Whang et al., 2013; Schultz, 1999, 2001; Drezner, 1999). The crisis bargaining framework focuses on the question of under which conditions, actions (e.g., issuing a threat of an economic sanction or engaging in war) occur and succeed (i.e., leading to the desired outcome). A common fundamental assumption is that the costs and benefits of any conflict situation drive the success rate and choice of strategy. Benefits are in the form of policy concessions made by a state targeted with a threat or coercion, in favour of the other party. Commonly studied conflict costs are the domestic audience cost created by backing down from a threat, and the economic cost resulting from engaging in coercion (Schultz, 1999).

Perhaps the most influential model of crisis bargaining in international conflict is the work of Schultz (1999), which incorporates (i) the relation between democracy and the domestic audience cost for backing-down on a threat, (ii) the relation between trade and the economic cost of engaging in a conflict and (iii) the role of uncertainty.[1] Schultz's crisis bargaining model (in the sequel, we will refer to this as the 'Schultz model') allows scholars to study the conditions for failure of threats and prospects of engagement in coercion. Specifically, it shows how the relation between economic cost and domestic audience cost influences the actors' decisions. However, a notable limitation in the Schulz model is that it does not account for variation in information about costs of benefits — i.e., the knowledge about one's own and the opposite party's costs and benefits in a conflict situation. Modelling this knowledge, or the uncertainty related to it, in a crisis bargaining model allows one

---

[1]The work of Schultz (2001, 1999) is, to an extent, based on past insight into interstate conflict, for example the work of Fearon (1997, 1994).

to model knowledge asymmetry or reason about the more realistic situation of imperfect knowledge about the opponent's costs and benefits. Consequently, information does not play a separate role, next to costs and benefits, in determining the outcome of a conflict. For scholars, the absence of variation in information in the model limits its applicability to empirical data.

In this article, we address this issue by proposing a crisis bargaining model, in which the probability of success of a threat and engagement in a coercion is not only influenced by domestic audience and economic cost, but also by the uncertainty in a state's belief about the costs and benefits of the opposing party. We make these costs and benefits and accompanying probabilities explicit, which allows for application of the model to various conflict scenarios. Our model contributes to existing literature in two ways. First, we offer a game theory crisis bargaining model and derive success and engagement probabilities formally, which allows for easy application to various conflict scenarios, whereas the Schultz model does not specify any type of probabilities. We explore the model outcomes across various values of the parameters, leading to a number of propositions on the effects of economic costs, audience costs and payoff uncertainty on the conflict's outcome. Second, we distinguish the constraining role of information and domestic audience cost as two separate effects in determining a situation's outcome; whereas previous literature did not differentiate between domestic audience cost and information, assuming it to be a single mechanism. We make the code of all our analyses available in a well-documented archive, thereby inviting other researchers to adjust our baseline conditions and generate own results with our crisis bargaining model.

Our paper is structured as follows. First, in Section 2, we offer an overview of the literature on crisis bargaining and list the limitations of the current approaches. Then, we develop a novel game theory model for crisis bargaining under complete and incomplete information in Section 3. Next, in Section 4, we analyse the prediction of our model in respect to effectiveness of threats and onset of coercion; we also derive a number of propositions. Finally, we conclude the article in Section 5.

2 Literature review

2.1 Overview of the literature on crisis bargaining

A common way of modelling conflict, either economic or military, is by treating it as a sequential game (e.g., Drezner, 2003; Schultz, 1999; Lacy and Niou, 2004). In such a game, the two players, the sender (reactor on a certain status quo) and the target (initiator of a new status quo), decide sequentially whether to (1) issue a threat of coercion (sender's turn), (2) resist the threat of coercion if a threat is issued by the sender (target's turn), and (3) follow up on the threat with engagement if the target resists (sender's turn). The sequential nature of these models leads to the concept of a 'game tree', consisting of several layers (turns), with various ending branches (outcomes). The goal of these models is to identify under what conditions threats succeed and coercion engagement occurs. We list a number of common assumptions and characteristics among these game theory models of crisis bargaining in the following, based on the work of Fearon (1994); Drezner (2003); Schultz (1999); Signorino (1999); Whang et al. (2013) and Lacy and Niou (2004):

- The sender and target are assumed to play a game of incomplete information, meaning that the payoffs at each instance of the game are private information — actors do know their own payoffs, but do not know the exact payoffs of the opponent.

- The sender and target are assumed to know a 'mean' payoffs of the opponent, where 'mean' relates to the costs and benefits averaged over all potential conflict that the sender could face. In other words, the target and the sender do not know what are the exact costs for their conflict and operate under a form of 'common knowledge' available to all actors.

- The domestic audience cost and economic cost for engaging in coercion are assumed to have a uniform probability distribution (between 0 and 1); thus, each payoff that is possible is equally likely. Nevertheless, the exact probabilities are not specified in the literature.

- Scholars assume a (negative) domestic audience cost as a penalty for issuing an empty threat and (non-positive) economic costs for engaging in coercion.

- The sender is assumed to only follow-up on a threat if she expects that the cost of engaging in a coercive action is smaller than the domestic audience cost. Related to this, scholars assume that the 'mean' domestic audience costs increase with the level of democracy of the sender state.

- Finally, in the crisis bargaining literature, audience costs are (partly) dependent on the 'resolve' characteristic of senders. Two types of senders are distinguished: those with high resolve and those with low resolve ('Nature' draws the sender's type), and this is assumed to be private information for each player. High resolve senders have higher domestic audience cost for backing down on a threat relative to economic cost for engaging in a coercion. And the opposite holds for low resolve players. Consequently, high resolve players follow up, engaging in coercion after a threat if the target stands firm. Low resolve players issue only empty threats and do not follow with an imposition.[2]

2.2 Limitations of current models

The crisis bargaining models proposed by scholars of international conflict (in particular, the work of Fearon (1994); Drezner (2003); Schultz (1999); Signorino (1999); Whang et al. (2013) and Lacy and Niou (2004)) suffer from a number of overlapping mis-specifications. First, the outcome of engaging in conflict (or likeliness of winning) is often treated as known (Schultz, 1999; Whang et al., 2013; Lacy and Niou, 2004), while literature does not explain why this should commonly be the case. If anything is known to the other party, it is more likely another quantity: the expected cost of coercion. For example, researchers have access to data on the expected economic costs of sanctions (Morgan et al., 2014).

Second, researchers indicate that players face incomplete information about each other payoffs, yet in the models, only the costs of economic sanctions is treated as unknown — the opponent's audience costs for backing down (sender) or concession to a threat (target) are considered known, while there is no clear indication why this should be the case.

Third, researchers argue that threatening with coercion (e.g., in the form of threatening an economic sanction) can change the target's belief about the sender's payoffs, a mechanism commonly referred to as 'signalling' (Fearon, 1994; Drezner, 2003; Schultz, 1999; Whang et al., 2013). Yet, at the same time they argue that the payoffs follow a uniform distribution, determined *a priori* (i.e., before the game starts) — making it contradictory to attain changes in the payoffs after the start of the game, and in particular, after the sender's threat. Hence, there is no room in current formal models for a change in target's estimate of the sender's costs for engaging in coercion. Yet, researchers do expect to observe this relation in data on conflict.

Fourth, the target's decision to comply with a threat is based on the estimate of a probability of a threat being followed up on, which in turn is based on an estimate of the opponent's payoffs. As mentioned before, a signalling effect of threats can shift the estimate of the target about the sender's payoffs, but how certain the target is about this estimate is not specified. In other words, scholars focus on signalling, largely ignoring the uncertainty around and the value of payoffs, which may be the sole driver of threats success and engagement in coercion. The importance of specification of uncertainty as a distribution around payoffs is acknowledged (Schultz, 1999), however, models have not yet been extended accordingly.

In this article, we attempt to address the listed shortcomings in the literature and derive a general framework for international conflict. We do this by offering an improved model of crisis bargaining, with a complete specification of the role of costs (domestic audience and economic) and uncertainty. First, we formalise uncertainty on all the costs faced by the sender and the target. Second, our model focuses on the more fundamental quantity of cost of engaging in coercion, and allows one to calculate success probabilities based on these costs. Finally, we distinguish uncertainty as a separate mechanism, next to economic and domestic audience cost, in determining the outcome of the game.

3 A Model for complete and incomplete information

3.1 Crisis bargaining model

In this section, we present the general structure of a crisis bargaining game between a sender and a target. The sender is a state that issues a threat of war or economic sanctions, while the target is a state towards which this threat is directed. For simplicity, we model the sender and the target as unitary actors and assume that the sender engages only in unilateral conflict. The game consists of five phases, in three of which one of the two players has to choose an action. Figure 1 displays the game's phases t_0 to t_4. The game is sequential — in each phase only one of the two players chooses

---

[2]A strategy referred in the literature as bluffing, see Whang et al. (2013)

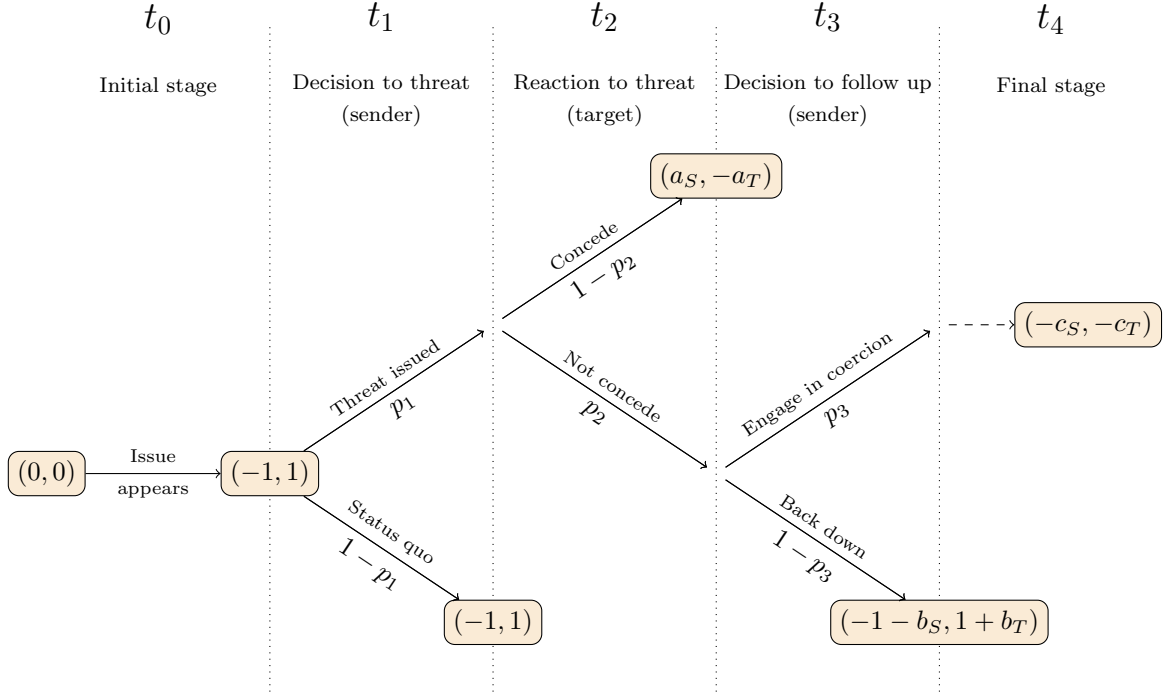an action, based on the action of their opponent in the previous phase. Below we discuss the various phases.

Figure 1: The game consists of five phases (t_0 to t_4). The phases are sequential and the phase's active player is indicated in brackets. The boxed tuples indicate the payoffs in the order of (*Sender,Target*). The payoff parameters are the domestic audience cost a_S and a_T due to the target's concession to a threat (audience benefit for the sender), the domestic audience cost b_S and b_T for the sender's backing down after the threat (i.e., audience 'benefit' for the target) and two cost terms c_S and c_T reflecting the economic costs resulting from engaging in coercion. The probabilities stated underneath every branch are discussed in Section 3.4.

Phase t_0 - Initial stage

The start of phase t_0 represents the baseline phase before a contested policy event occurred. No player takes any action (i.e., there is no 'game' element in this phase). A policy issue appears, with a positive effect on the target and a negative effect on the sender.[3] The resulting payoffs are marked by $(-1, 1)$: a negative (-1) payoff for the sender, and a positive (1) payoff for the target. The values -1 and 1 are taken for simplicity.

Phase t_1 - Decision to threat (sender's turn)

In phase t_1, the sender needs to decide how to address the policy issue. If she decides to settle with the status quo, the game ends with the same payoffs as phase t_1 started with: $(-1, 1)$. The alternative is to issue a (economic or military) threat of coercion, which continues the game to phase t_2.

Phase t_2 - Reaction to threat (target's turn)

The target state responds to a threat in phase t_2, either conceding to it or not. If the target concedes, the payoffs do not return to the initial $(0, 0)$ in t_0, because the target faces a domestic audience cost of $-a_T$ for abandoning its policy and the sender receives a domestic audience benefit of $+a_S$ for a foreign policy success. If the target state decides *not* to concede, the game enters phase t_3.

---

[3]For example, an uranium enrichment programme conducted by an authoritarian state.

Phase t_3 - Decision to follow up (sender's turn)

Now, the sender state has to decide whether or not to follow up on the threat and engage in coercion. If she backs down, the game ends, returning the initial issue to the status quo $(-1, 1)$ and adding a domestic audience cost $-b_S$ for the sender, penalising issuing an empty threat, and a domestic audience benefit $+b_T$ for the target for standing firm in face of foreign pressure.

Phase t_4 - Final stage

If the sender state decides to engage in coercion, the game continues and ends in phase t_4. We parametrise the coercion cost with (c_S, c_T), which are a combination of the eventual conflict outcome (status quo or not), domestic audience costs and economic costs. We assume that the costs of engagement in a coercion are large, so that $c_S \gg a_S, b_S$ and $c_T \gg a_T, b_T$.

3.2   Dynamics under complete information

We start by analysing possible dynamics of the crisis bargaining model, under the assumption that each player has complete information about the payoffs (i.e., are known) of both oneself and the opponent. This setup reflects previous work on crisis bargaining (Schultz, 1999).

Knowing that all parties are badly affected by engaging in coercion (i.e., economic costs $c_S, c_T \gg 1$), we conclude that threats will only be issued if the sender knows that the target concedes in t_2. The reasoning is as follows, starting from the final stage t_4, working backwards. If the game ends up in this phase, a sequence of decisions (after the initiation of the game via a threat) must have occurred: a threat was issued, the target did not concede to the threat, and the sender followed up on the threat, engaging the coercion. In such a case, as all payoffs are known, the target must have known in t_2 that the sender would not back down in t_3. Because the target is better off conceding in t_2 (as $-a_T \gg -c_T$), the target would never let the game progress to t_3 (and t_4), so she will always concede in t_2. Analogously, the sender is better off with the status quo than either the outcomes of backing down (t_3) (as $-1 > -1 - b_S$) or engaging in coercion (t_4) as (as $-1 \gg -c_T$), so she will only issue a threat in t_1 if she is sure that the target will concede in t_2.

Hence, the complete information game never reaches t_3 or t_4 and either ends with the status quo, or a successful threat (i.e., target concedes to threat). That is, states never engage in coercion and targets never resists a threat of coercion. Of course, this is not what is observed in reality — states do go to war and threats are resisted[4]. Still, the complete information game offers us a reference point "as with all comparative-static exercises, though, the important results deal with the relative probabilities of an outcome across two games, not the absolute probability in either game" (Schultz, 1999). To further expand our theoretical framework, in the following section we relax the assumption of complete information.

3.3   Introducing uncertainty about opponent's payoff

In contrast to the previous section, this section is devoted to analyse the crisis bargaining model *without* assuming that both parties know all payoffs. We set up such a model by introducing a measure of uncertainty for each player about the payoffs (most importantly, about the payoffs of the opposing player). In addition, we vary the uncertainty which allows us to generate more refined results about the effect of uncertainty on the behaviour of states in the crisis bargaining situation. We thereby add a third component, next to domestic audience and economic cost, that influences the dynamics of conflicts. Note that we could set the uncertainty to zero, which would simplify the incomplete information model to a complete information model, effectively making it a special case.

We introduce a notation $C_x^{(y)}$, where C is one of the cost or payoff variables. The lower script x refers to the state (sender or target) to which the cost or payoff applies, e.g., b_S is the audience cost *for the sender*. The upper script y refers to from which perspective (sender or target) we are viewing the cost, e.g. $b_S^{(T)}$ is the audience cost for the sender, *from the perspective of the target*. Note that some variables are fully known by the party itself — $b_S^{(S)}$ is the sender's own audience cost as estimated by itself, hence we assume that she knows this, i.e., the uncertainty of this variable is zero. However, the uncertainty of other variables (like $b_S^{(T)}$) are dependent on how much the states know about each other. We introduce uncertainty by modelling each of these variables as

---

[4]This can be concluded from, e.g., the TIES data set, which indicates a substantial number of cases of failed threats of economic sanctions (Morgan et al., 2014).

Gaussian distributions, with (if not stated otherwise) a mean and standard deviation of 1. The unimodal structure of Gaussian distributions is chosen deliberately to reflect the assumption that payoff estimates are generally centred around a certain best guess (in contrast to multi-model distributed variables, that seem implausible).

We also assume that the costs of engaging in coercion $(-c_S, -c_T)$ are not fully known to the players — both their own payoffs and the opponents payoff. This allows us to account for the uncertainty related to a military intervention or economic coercion; factors like success, duration or intensity of the conflict are not fully known by both parties beforehand.

3.4 Dynamics under incomplete information

We now analyse the dynamics of the crisis bargaining model depicted in Figure 1 in the incomplete information framework, where each player faces uncertainty about the payoffs of the opposing state (and own payoff in case of engaging in coercion). In order to compare the pay-offs of each player at each decision node, we need to compute all the payoffs starting from phase t_4 backwards. While the complete information framework only has deterministic outcomes based on the distribution of payoffs, the incomplete information game is determined by a set of probabilities — p_1, p_2 and p_3 (see Figure 1). In contrast to existing literature, where these probabilities (if explicitly defined) are commonly given as a constant, we express these probabilities in terms of the payoff variables and their uncertainties, as we will illustrate in the remainder of this section. We use notation $N(\bar{x}, \sigma_x)$ for the normal distribution of variable x, with mean $\bar{x}$ and standard deviation σ_x. True values are denoted as just 'x', like in Figure 1, and in derivations we put $N(\bar{x}, 0) = x$ (although in theory, $\bar{x}$ can still be wrong). Analogous to payoffs, we write the probability $p_i^{(y)}$ as the probability seen from the perspective of state y.

Phase t_3

Similarly to the analysis of the complete information framework in Section 3.2 we begin by analysing the game tree from the end, starting with p_3, defined as the probability of the sender engaging in coercion (given that phase t_3 is reached), having pay-off $-c_S$. At the same time, the probability that the sender backs down is $1 - p_3$, achieving payoff $-1 - b_S$. As only the sender has an active choice in this phase, we only need to compare the payoffs of the sender to calculate both $p_3^{(S)}$ and $p_3^{(T)}$. Thus, the probability of imposing a sanction p_3 boils down to:

$$
\begin{aligned}
p_3^{(S)} &= P(-c_S^{(S)} > -1 - b_S^{(S)}) \\
&= P(-N(\bar{c_S}^{(S)}, \sigma_{c_S}^{(S)}) > -1 - b_S)) \\
&= P(N(0, \sigma_{c_S}^{(S)}) > -1 - b_S + \bar{c_S}^{(S)})
\end{aligned}
$$

and similarly for the target (for whom both $c_S^{(T)}$ and $b_S^{(T)}$ are uncertain):

$$
p_3^{(T)} = P(N(0, \sqrt{(\sigma_{b_S}^{(T)})^2 + (\sigma_{c_S}^{(T)})^2}) > -1 - \bar{b_S}^{(T)} + \bar{c_S}^{(T)})
$$

Phase t_2

Now that we have an expression for p_3, we work backwards towards phase t_2. We define p_2 (see Figure 1) as the probability that the target does not concede to a threat, given that a threat had been issued (analogously, $1 - p_2$ is the probability that the target concedes). To estimate p_2, we need to calculate the expected payoff of the target for not conceding (i.e., entering t_3), as estimated by the sender and by the target itself:

$$
E_T^{(T)}(t_3) = p_3^{(T)} \cdot (-c_T^{(T)}) + (1 - p_3^{(T)}) \cdot (1 + b_T^{(T)})
$$

using the expression for $p_3^{(T)}$ from above. Furthermore, $b_T^{(T)} = b_T$ is known and $c_T^{(T)} = N(\bar{c_T}^{(T)}, \sigma_{c_T}^{(T)})$ is normal distributed, making $E_T^{(T)}$ a normal distributed variable. Similarly:

$$
E_T^{(S)}(t_3) = p_3^{(S)} \cdot (-c_T^{(S)}) + (1 - p_3^{(S)}) \cdot (1 + b_T^{(S)})
$$

where $b_T^{(S)} = N(\bar{b}_T^{(S)}, \sigma_{b_T}^{(S)})$ and $c_T^{(S)} = N(\bar{c}_T^{(S)}, \sigma_{c_T}^{(S)})$ are normal distributed. Using the expressions for the expected payoff of the target in t_3, we can calculate p_2 as seen from both parties:

$$
\begin{aligned}
p_2^{(T)} &= P(E_T^{(T)}(t_3) > -a_T^{(T)}) \\
&= P(N(0, \sigma_{E_T^{(T)}(t_3)}) > -a_T^{(T)} - \bar{E}_T^{(T)}(t_3)) \\
&= P(N(0, \sigma_{c_T^{(T)}}) > -a_T^{(T)} - \bar{E}_T^{(T)}(t_3))
\end{aligned}
$$

and:

$$
\begin{aligned}
p_2^{(S)} &= P(E_T^{(S)}(t_3) > -a_T^{(S)}) \\
&= P(N(0, \sqrt{(\sigma_{c_T}^{(S)})^2 + (\sigma_{b_T}^{(S)})^2 + (\sigma_{a_T}^{S})^2}) > -\bar{a}_T^{(S)} - \bar{E}_T^{(S)}(t_3))
\end{aligned}
$$

Phase t_1

Analogously, we calculate the expected payoff of the sender for entering t_2 (by issuing a threat), viewed from both perspectives:

$$
E_S^{(S)}(t_2) = p_2^{(S)} \cdot (E_S^{(S)}(t_3)) + (1 - p_2^{(S)}) \cdot (a_S^{(S)})
$$

and:

$$
E_S^{(T)}(t_2) = p_2^{(T)} \cdot (E_S^{(T)}(t_3)) + (1 - p_2^{(T)}) \cdot (a_S^{(T)})
$$

where $E_S^{(S)}(t_3)$ and $E_S^{(T)}(t_3)$ are normally distributed. This leads to the following expression of p_1:

$$
\begin{aligned}
p_1^{(S)} &= P(E_S^{(S)}(t_2) > -1) \\
&= P(N(0, \sigma_{E_S^{(S)}(t_3)}) > -1 - \bar{E}_S^{(S)}(t_2))
\end{aligned}
$$

and:

$$
\begin{aligned}
p_1^{(T)} &= P(E_S^{(T)}(t_2) > -1) \\
&= P(N(0, \sigma_{E_S^{(T)}(t_3)}) > -1 - \bar{E}_S^{(T)}(t_2))
\end{aligned}
$$

3.5 Summary

We have derived expressions for the probabilities p_1, p_2 and p_3 under different values of the parameter set, consisting of the averages and standard deviations of the domestic audience costs (a_S, a_T, b_S and b_T), the payoffs in the case of status quo $(-1, 1)$, and economic costs resulting from a potential coercion (c_S and c_T). When asking the question what will happen in each of the game tree phases, one is mainly interested in the probabilities as seen from the active player's perspectives; that is, in phase t_1, the sender's decision to issue a threat (determining $p_1^{(S)}$), in phase t_2, the target's decision to concede or not ($p_2^{(T)}$) and in phase t_3, the sender's final decision ($p_3^{(S)}$).

In Table. 1, we give a summary of the most important probabilities in scenarios with varying knowledge about the opposing player's payoffs. We calculate these values using a reference parameter set with 'incomplete' information (all means and standard deviations equal to 1), a 'complete' information column (reference with standard deviations 0) and a 'no-information' case (reference with a very large standard deviation equal to 1,000) — the latter is added as a comparison, but its results are trivial (i.e., all actions are fully random). We can see that under complete information, the sender always threats, the target concedes in half of the cases, and the sender always follows up on the threat if it comes to that decision. In the reference case (i.e., non-trivial uncertainty values), the sender does not always issue a threat, the target concedes more often, and the sender does not always follows up on the threat. These differences between the complete and incomplete information cases display the role of information on its own, reflecting the importance of including uncertainty in a crisis bargaining model. Not only allows this framework to vary uncertainty, the model can also be tailored for different values of the payoffs. In the next section, we explore several parameter settings that reflect empirical data.

Result	Symbol	Complete $\sigma \to 0$	Incomplete $\sigma = 1$	No information $\sigma \to \infty$
Action probabilities:				
Sender threats*	$p_1^{(S)}$	1	0.754	0.5
Target does not concede*	$p_2^{(T)}$	0.5	0.764	0.5
Sender engages in coercion*	$p_3^{(S)}$	1	0.841	0.5
End results:				
Status quo	$1 - p_1^{(S)}$	0	0.246	0.5
Concession target	$p_1^{(S)}(1 - p_2^{(T)})$	0.5	0.178	0.25
Backing down sender	$p_1^{(S)} p_2^{(T)}(1 - p_3^{(S)})$	0	0.091	0.125
Coercion onset	$p_1^{(S)} p_2^{(T)} p_3^{(S)}$	0.5	0.485	0.125

Table 1: Summary of the probabilities in the reference parameter settings. In the last two columns we use settings with either very low or very high standard deviations. The asterisk indicates that the probability is conditional on that the particular phase has been reached.

4 Analysing model predictions

In this section, we investigate the success probabilities of certain actions in the game tree for a range of parameter settings and derive empirically testable propositions. As noted above, if not otherwise stated, all means are equal to the true values, and the standard deviations of all distributions (i.e., opposing player's payoffs and own payoff in t_4) are equal to 1. Analogous to Table 1, we express the model results in probabilities. In particular, we map out the probability of a threat's success $(1 - p_2^{(T)})$ and the probability of engaging in coercion $(p_3^{(S)})$ (conditional on having reached these stages) for different payoffs settings.

4.1 Effectiveness of threats

Figure 2a relates this success rate to the true value (horizontal axis) and uncertainty (vertical axis) of the target's economic cost of coercion. Panel (b) relates the threat's success probability to the true value (horizontal axis) and standard deviation (vertical axis) of the sender's audience cost of backing down after the target have resisted a threat. The panels show that an increase in the economic cost to the target [panel (a)] or an increase in the audience cost to the sender [panel (b)] causes the threat's success rate to rise, which is in line with the predictions in the crisis bargaining literature (Schultz, 1999).

Furthermore, regarding uncertainty (the vertical axes), panel (a) shows that for relatively low true economic cost to the target, more uncertainty in this cost leads to an increase in the probability of threat's success, while for relatively high economic cost to the target more uncertainty leads to a lower probability of threat's success. In panel (b), where we vary the (true value and uncertainty of the) sender's audience cost, an increase in the uncertainty results in threats being less likely to succeed. This is a novel prediction in the crisis bargaining literature, in which the effects of information and domestic audience cost are separated and distinguished.
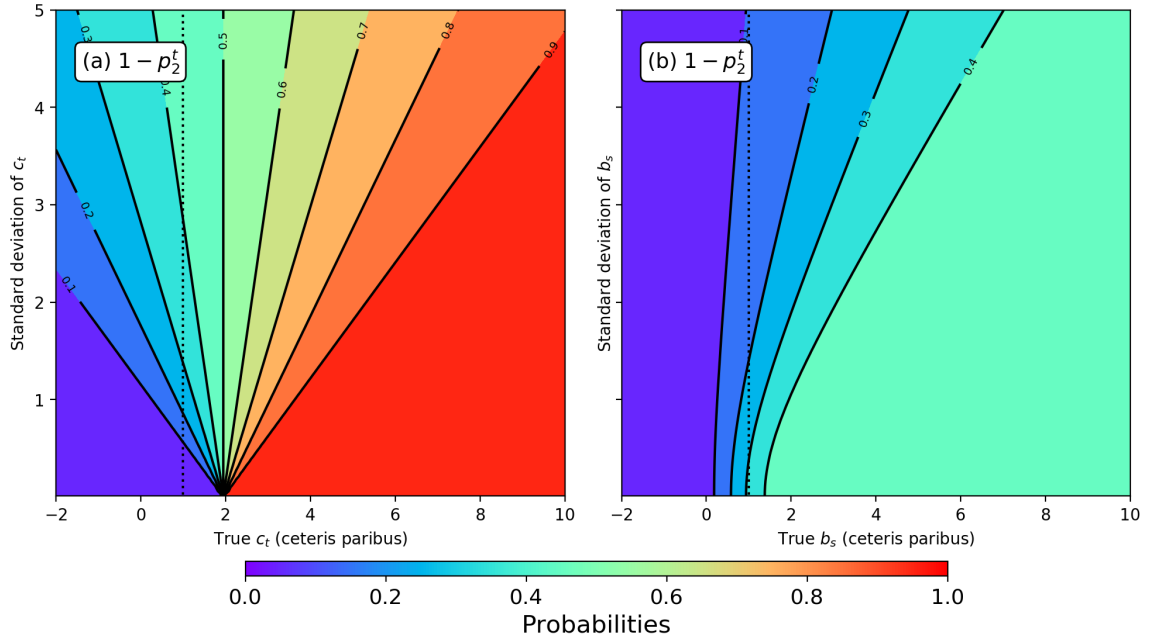
Figure 2: Probability of threat success $(1 - p_2^t)$ in colours, given that a threat has been issued. Panel (a) shows the effect of varying the true economic coercion costs to the target (horizontal) and standard deviation of these costs (vertical). Panel (b) shows the effect of varying the true sender's audience cost of an empty threat (horizontal) and the standard deviation of this value (vertical).

Based on the above analysis, we can derive the following three propositions:

Proposition 1. In a conflict between states, increase in the economic cost to a target increases the effectiveness of a threat.

Proposition 2. In a conflict between states, increase in the domestic audience cost to a sender increases the effectiveness of a threat.

Proposition 3a. In a conflict between states, increase in the uncertainty about the economic cost to a target decreases (increases) the effectiveness of threats, for relatively high (low) true value of the economic cost.

Proposition 3b. In a conflict between states, increase in the uncertainty about the domestic audience cost to a sender decreases the effectiveness of threats.

The first two propositions that we offer relate to hypotheses that have already been established in the literature on crisis bargaining. Our first proposition reflects the coercive hypothesis — an expectation that with an increase of the cost of economic sanctions or war, relative to the size of the sender's and target's economy, coercion is more likely to end, and succeed, at the threat stage (Whang et al., 2013; Katagiri and Min, 2019). Our second proposition reflects the public commitment hypothesis — an expectation that states with a high domestic audience cost (e.g. democracies) are more likely to succeed at a threat stage (Fearon, 1994; Schultz, 2001, 1999; Katagiri and Min, 2019).

The third proposition is a novel contribution to the crisis bargaining literature and concerns the impact of information. It shows that uncertainty is a separate determinant next to domestic audience and economic cost in driving success of threats of coercion. This results is also in contrast with part of the crisis bargaining theory, where information and domestic audience cost were presented as a single determinant (Schultz, 1999, 2001; Whang et al., 2013). Here, we disentangle these mechanisms and show that they can work in parallel.

4.2 Engagement in coercion

Next, we consider the behaviour of states in respect to engagement in coercion, based on our crisis bargaining model. In particular, the probability that the sender decides to engage in coercion, when the game had already reached t_3. Next to the effectiveness of threats, this is another focal point in the theory development on international conflict. Figure 3 shows the probability of a sender

engaging in coercion ($p_3^{(S)}$, given that a threat has been resisted) in colours. Analogous to Figure 2, panel (a) displays this probability for true values (horizontal axis) and standard deviation (vertical axis) of the sender's economic costs (in contrast to the *target's* costs in Figure 2). Panel (b) shows this probability for varying true values (horizontal axis) and standard deviation (vertical axis) of the sender's audience cost of backing down.

After an inspection of Figure 3, we conclude that a higher economic cost to the sender (as we would intuitively predict) reduces the probability of sender's engagement in a coercion. What is more, based on panel (b) in Figure 3, we observe that a higher audience cost to the sender increase the probability of coercion onset, what also follows an intuitive prediction. Interestingly, the role of uncertainty is again highlighted in these panels — high probabilities of engaging in coercion (for low economic costs or high audience costs) are decreased with higher uncertainty, and vice versa.
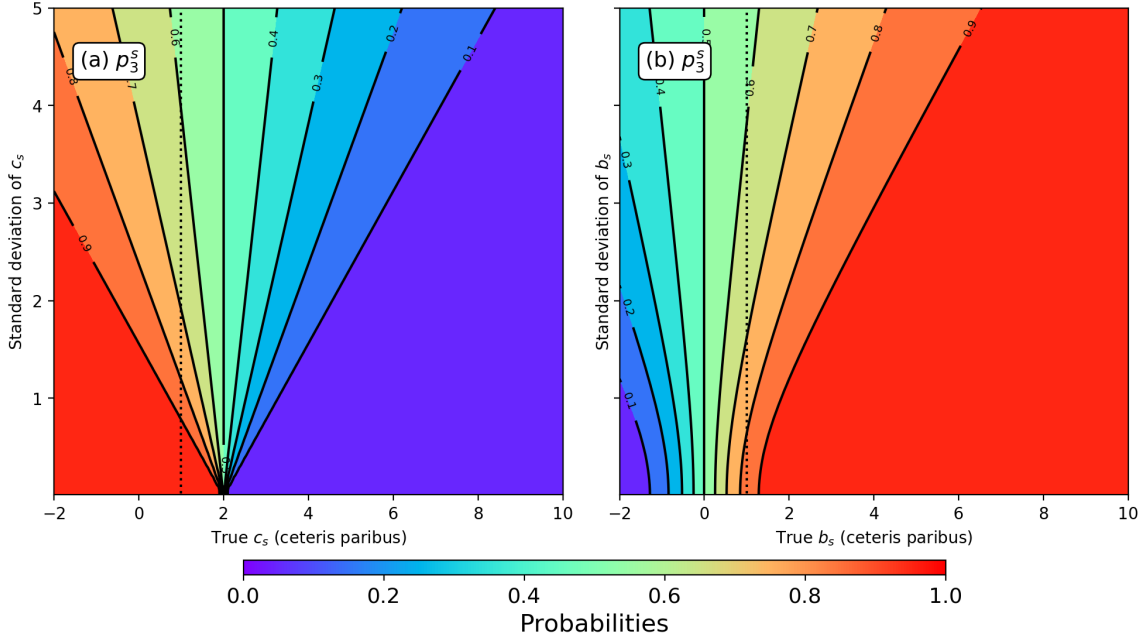

Figure 3: Probability of engaging in coercion (p_3^s) in colours, given that a target did not concede to a threat. Axes as in Fig. 2.

Based on our model's prediction, we have generated three propositions on coercion onset:

Proposition 4. In a conflict between states, increase in the economic cost to a sender decreases the probability of a coercion onset.

Proposition 5. In a conflict between states, increase in the audience cost to a sender increases the probability of a coercion onset.

Proposition 6a. In a conflict between states, increase in uncertainty decreases the probability of coercion onset, for relatively high true value of both economic and audience cost to the sender.

Proposition 6b. In a conflict between states, increase in uncertainty increases the probability of coercion onset, for relatively low true value of both economic and audience cost to the sender.

In the next two sections, we explore our crisis bargaining model by applying it to two conflict scenarios: (I) a democratic sender threatens an autocratic target and (II) both sender and target experience a high audience cost. Furthermore, we offer a programme that allows scholars to themselves change our assumptions and generate scenarios and respective predictions.[5]

4.3 Scenario I - A democratic sender threatens an autocratic target

In the first scenario, the target's domestic audience cost is lower than the cost of engagement in coercion ($a_T < c_T$) and the sender's domestic audience cost is higher than the cost of an engagement in coercion ($a_S > c_S$). This scenario mimics a typical conflict situation between a democratic sender and an autocratic target; scholars indicate that authoritarian regimes can successfully shield

---

[5]This programme is available via a Github repository.

themselves from popular discontent (Allen, 2008) (i.e., low domestic audience costs for authoritarian regimes), that voters in democratic regimes punish political leaders for issuing empty threats (Kertzer and Brutger, 2016) (i.e., high domestic audience costs for democracies), and that democracies, on average, boast more complex and resilient economies than authoritarian regimes (Farrell and Newman, 2019) (i.e., economic costs of coercion are relatively lower for democracies). This theoretical set-up is particularly relevant to instances of economic sanctions and non-proliferation, for example US sanctions against Iran or North Korea.

To analyse this scenario in the framework of our crisis bargaining model, we set $a_T = b_T = 0.5, c_T = 1.5, a_S = b_S = 1.5$ and $c_S = 0.5$. Table 2 shows the resulting probabilities for three cases of the standard deviations as in Table 1. In a complete information case (second-last column), the game always ends in the target conceding, as it knows that the sender will impose. The incomplete information case column (middle column) shows that the probability of the target not conceding becomes greater than 0. Yet, the probability of the sender backing down is still marginal, pointing towards the dominating effect of domestic audience costs at the final decision stage. The audience costs-driven determination of the sender to engage in coercion in this scenario setting may reflect the determination of the Trump administration in sanctioning of Iran over nuclear proliferation concerns.

Table 2: Summary of probabilities in the parameter settings of Scenario I. In the last two columns we use settings with either very low or very high standard deviations. Standard deviations are only non-zero (in any of the cases), when considering pay-offs of the opposing player or own conflict costs.

Result	Symbol	Incomplete $\sigma = 1$	Complete $\sigma \to 0$	No information $\sigma \to \infty$
Action probabilities:				
Sender threats*	$p_1^{(S)}$	0.971	1	0.5
Target does not concede*	$p_2^{(T)}$	0.222	0	0.5
Sender engages in coercion*	$p_3^{(S)}$	0.977	1	0.5
End results:				
Status quo	$1 - p_1^{(S)}$	0.029	0	0.5
Concession target	$p_1^{(S)}(1 - p_2^{(T)})$	0.755	1	0.25
Backing down sender	$p_1^{(S)} p_2^{(T)} (1 - p_3^{(S)})$	0.005	0	0.125
Coercion onset	$p_1^{(S)} p_2^{(T)} p_3^{(S)}$	0.211	0	0.125

* given that the particular phase has been reached.

4.4 Scenario II - High audience costs for sender and target

A second scenario considers the case of both the sender and the target state experiencing higher domestic audience costs than economic coercion costs. This theoretical set-up reflects cases in which democratic senders and target states that are newly formed (authoritarian) regimes which have not yet consolidated power and do not want to appear weak to domestic competitors (Spaniel and Smith, 2015).[6] This scenario may reflect the US-Venezuela conflict, where the regime of Maduro is in a power-transition period since year 2013, after the death of Hugo Chavez, and faces a number of domestic competitors.

To analyse this scenario in our model, we set $a_T = b_T = a_S = b_S = 1.5$ and $c_T = c_S = 0.5$ (see Fig. 1). In the complete information case (second-last column), the parties eventually always engage in coercion, and therefore, threats are never successful. However, in an incomplete information game (middle column), we observe that the effectiveness of threats increases (11% chance of target's concession). Nevertheless, we observe that both the target is not very likely to concede and the sender is very likely to engage in a conflict. This may serve as a reference point to the on-going US-Venezuela conflict, where the Maduro regime rejected humanitarian aid to Venezuelans in order to not appear weak and there appears little sight of a solution. This example points to two potential proxies of uncertainty. One could be duration of a regime; actors cannot rely on past experience (own and of other states) when dealing with a relatively young regime. This may add to the uncertainty about the costs and benefits that the regime faces. Second, diplomatic ties may reflect the level of uncertainty. States with frequent diplomatic exchanges may be more likely to

---

[6]It may also be applied to a conflict between two democracies.

precisely estimate the costs and benefits of the opponent in a crisis bargaining situation.

Table 3: Summary of probabilities in the parameter settings of Scenario II. In the last two columns we use settings with either very low or very high standard deviations.

Result	Symbol	Incomplete $\sigma = 1$	Complete $\sigma \to 0$	No information $\sigma \to \infty$
Action probabilities:				
Sender threats*	$p_1^{(S)}$	0.842	1	0.5
Target does not concede*	$p_2^{(T)}$	0.892	1	0.5
Sender engages in coercion*	$p_3^{(S)}$	0.977	1	0.5
End results:				
Status quo	$1 - p_1^{(S)}$	0.158	0	0.5
Concession target	$p_1^{(S)}(1 - p_2^{(T)})$	0.091	0	0.25
Backing down sender	$p_1^{(S)}p_2^{(T)}(1 - p_3^{(S)})$	0.017	0	0.125
Coercion onset	$p_1^{(S)}p_2^{(T)}p_3^{(S)}$	0.734	1	0.125

* given that the particular phase has been reached.

5   Conclusions

In this article, we put forward an extended game theory model of crisis bargaining for international conflict which goes beyond the current literature in the following ways: (a) we introduce uncertainty on both domestic audience and economic cost, (b) we formalise what uncertainty means and quantify its distribution, we (c) show that uncertainty and costs are separate mechanisms and finally, (d) we offer a distinction between uncertainty and signalling. This extended framework is consistent with the main predictions of the existing models but at the same time implies novel and empirically testable propositions relating to the response of a target state to a threat and the response of a sender to a target that did not concede to a threat.

Consistent with earlier work, our model predicts that both an increase in the economic cost to a target and an increase in the domestic audience cost to a sender increases the effectiveness of a threat. In addition, our model predicts that an increase in the uncertainty about the economic cost to a target decreases the effectiveness of threats, if the actual value of the economic cost is relatively high. We also observe that an increase in the uncertainty about the domestic audience cost to a sender decreases the effectiveness of threats. When uncertainty increases, the target is more likely to decide to resist a threat of coercion that would have been successful under complete information, or with less uncertainty. With respect to engagement in coercion, our model predicts that an increase in the economic cost to a sender decreases the probability of a coercion onset, while an increase in the audience cost (to a sender) increases the probability of engagement in coercion. Finally, we observe that uncertainty has a equalising effect on coercion onset and decreases (increases) the probability of coercion, for relatively high (low) actual values of both economic and audience cost to the sender.

More and more empirical data that offers information on conflicts that did not escalate beyond a threat is available. For example, the TIES data set (Morgan et al., 2014) includes cases of economic sanctions that terminated at the threat stage (failures and successes). Our general framework offers guidance for research into the conditions that systematically affect effectiveness of threats and the prospects of coercion onset conditional on a failed threat. A large body of scholarly work — for example findings of Whang et al. (2013); Whang and Kim (2015); Lacy and Niou (2004) on effectiveness of threats and of Peksen (2019); Hafner-Burton and Montgomery (2008); Schultz (1999); Spaniel and Smith (2015) on onset of coercion — may be reconsidered in the light of this article. A set of hypotheses based on the Schultz (1999) crisis bargaining framework would not allow empirical research to take full advantage of, for example, the TIES data set (Morgan et al., 2014) and observations of conflict in international relations that terminated at the threat stage. In addition, our model is readily accessible to researchers and can easily be adjusted for desired values and generate visualisation of predictions.

The model can be explored further by including signalling to further account for the role of information. Currently, we only study the effect of change in variance around a mean value — our measure of uncertainty. However, a shift in the mean value of an estimate would, due to a

change from status quo to a threat, could potentially reflect what scholars refer to in the literature as signalling.

In summary, this article offers a synthesis and an advancement of scholarship on crisis bargaining. We offer a single framework for international conflict under complete and incomplete information, in the form of a game theory model that allows for variation in uncertainty. Crucially, our model makes the novel prediction that the probabilities of a threat's success and coercion onset depend both on the economic and domestic audience costs, as well as on the uncertainties about these costs. Finally, the generality of our model allows for broad application in empirical research and offers scope for further investigation of many subjects related to information, such as signalling or bluffing.

References

Allen, S. H. (2008). The Domestic Political Costs of Economic Sanctions. *Journal of Conflict Resolution 25*(6), 916–944.

Bueno de Mesquita, B., J. D. Morrow, R. M. Siverson, and A. Smith (1999). An Institutional Explanation of the Democratic Peace. *American Political Science Review 93*(4), 791–807.

Chandler, D. (2013). Promoting democratic norms? Social constructivism and the 'subjective' limits to liberalism. *Democratization 20*(2), 215–239.

Dixon, W. J. (1994). Democracy and the Peaceful Settlement of International Conflict. *American Political Science Review 88*(1), 14–32.

Drezner, D. (1999). *The Sanctions Paradox: Economic Statecraft and International Relations*. Cambridge: Cambridge University Press.

Drezner, D. (2003). The Hidden Hand of Economic Coercion. *International Organization 57*(03), 643–659.

Farrell, H. and A. L. Newman (2019). Weaponized Interdependence. *International Security 44*(1), 42–79.

Fearon, J. (1994). Domestic Political Audiences and the Escalation of International Disputes. *American Political Science Review 88*(03), 577–592.

Fearon, J. (1997). Signaling foreign policy interests: Tying hands versus sinking costs. *Journal of Conflict Resolution 41*(1), 68–90.

Gartzke, E. (2007). The Capitalist Peace. *American Journal of Political Science 51*(1), 166–191.

Goldstein, J. S. (2011). *Winning the War on War: The Decline of Armed Confict Worldwide*. New York: Dutton.

Hafner-Burton, E. M. and A. H. Montgomery (2008). Power or Plenty. How Do International Trade Institutions Affect Economic Sanctions? *Journal of Conflict Resolution 52*(2), 213–242.

Hufbauer, G. C., J. J. Schott, K. A. Elliott, and B. Oegg (2007). *Economic Sanctions Reconsidered* (3 ed.). Washington, D.C.: Peterson Institute for International Economics.

Katagiri, A. and E. Min (2019). The Credibility of Public and Private Signals: A Document-Based Approach. *American Political Science Review 113*(1), 156–172.

Kertzer, J. D. and R. Brutger (2016). Decomposing Audience Costs: Bringing the Audience Back into Audience Cost Theory. *American Journal of Political Science 60*(1), 234–249.

Lacy, D. and E. M. Niou (2004). A Theory of Economic Sanctions and Issue Linkage: The Roles of Preferences, Information, and Threats. *Journal of Politics 66*(1), 25–42.

Maoz, Z. and B. Russett (1993). Normative and Structural Causes of Democratic Peace , 1946-1986. *American Political Science Review 87*(3), 624–638.

Morgan, T. C., N. Bapat, and Y. Kobayashi (2014). Threat and imposition of economic sanctions 1945–2005: Updating the TIES dataset. *Conflict Management and Peace Science 31*(5), 541–558.

Neuenkirch, M. and F. Neumeier (2016). The impact of US sanctions on poverty. *Journal of Development Economics 121*, 110–119.

Peksen, D. (2019). When Do Imposed Economic Sanctions Work? A Critical Review of the Sanctions Effectiveness Literature. *Defence and Peace Economics 30*(6), 635–647.

Pinker, S. (2011). *The Better Angels of Our Nature*. New York: Viking.

Pond, A. (2017). Economic Sanctions and Demand for Protection. *Journal of Conflict Resolution 61*(5), 1073–1094.

Schultz, K. (1999). Do Domestic Institutions Constrain or Inform? *International Organization 53*(2), 233–266.

Schultz, K. (2001). *Democracy and Coercive Diplomacy*. Cambridge: Cambridge University Press.

Signorino, C. S. (1999). Strategic Interaction and the Statistical Analysis of International Conflict. *American Political Science Review 93*(02), 279–297.

Spaniel, W. and B. C. Smith (2015). Sanctions, Uncertainty, and Leader Tenure. *International Studies Quarterly 59*(4), 735–749.

Whang, T. and H. J. Kim (2015). International Signaling and Economic Sanctions. *International Interactions 41*(3), 427–452.

Whang, T., E. V. Mclean, and D. W. Kuberski (2013). Coercion, Information, and the Success of Sanction Threats. *American Journal of Political Science 57*(1), 65–81.