Should Audiences Cost?

Optimal Domestic Constraints in International Crises[*]

Scott Ashworth
Harris School of Public Policy Studies
University of Chicago

Kristopher W. Ramsay
Department of Politics
Princeton University

July 14, 2017

Abstract

We study voting strategies in a model of crisis bargaining. The model allows for the possibility that voters punish a leader who agree to settlements or punishes leaders who initiate war. Citizens make voting decisions with the incentives they generate for leaders in mind. In the optimal strategy, citizens always punish leaders who initiate crises and then back down. Whether they punish leaders for backing down rather than going to war, on the other hand, depends on the status quo and on the costs of war. We identify the conditions under which citizens use re-selection incentives to offset the different preferences over fighting (which involves punishing fighting), and the conditions under which they use punishments to exaggerate that preference divergence (which involves rewarding fighting). We relate these results to the theoretical and empirical literatures on audience costs.

Word count: 9776

International crises often place political leaders in domestic jeopardy. Sanctions for failure range from loss of office to death. Unsurprisingly then, it has become a commonplace of international relations scholarship that leaders act with one eye on retaining domestic support (Snyder, 1991; Fearon, 1994a; Smith, 1998; Schultz, 2001a; Bueno de Mesquita et al., 1992; Goemans, 2000; Debs and Goemans, 2010). Two prominent literatures suggest strategies that citizens could adopt to optimally exploit leaders' interest in retention.

One literature emphasizes *audience costs.* Fearon (1994a) observes that a domestic audience can encourage resolve, enhancing a leader's bargaining posture. In the canonical version, a leader pays a domestic cost after backing down from fighting in a crisis they themselves initiated. Anticipating this cost, only leaders with high values of fighting initiate crises, credibly influencing a target's beliefs about their own war payoff.

The other literature emphasizes *political bias* (Bueno de Mesquita and Lalman, 1992; Bueno de Mesquita et al., 1992; Jackson and Morelli, 2007). These authors argue that leaders will typically want to fight in circumstances the citizens would not. This could be, for example, because leaders themselves do not bear the physical costs of fighting, or because they expect to appropriate a disproportionate share of the rewards of victory. Domestic consequences for belligerent foreign policy decisions give leaders lower crisis and war payoffs, leading to more peace and fewer challenges.

In this paper, we characterize the optimal retention strategy in a model that captures concerns from both strands of the literature.[1] These results clarify when citizens should use retention incentives to offset the political bias, and when they should use audience costs to exaggerate the political bias.[2]

We build on the standard crisis model (Bueno de Mesquita and Lalman, 1992; Fearon, 1994b; Schultz, 1999). Two countries, Home and Foreign, start with a status quo division of some resource. Home's leader has private information about the war payoffs, and can decide whether or not to challenge the status quo. If challenged, Foreign offers a new division of the resource that Home can accept or reject. Rejection leads to a war, modeled as a costly lottery with the resource as

---

[1]Downs and Rocke (1994) also look at the agency problem associated with the decision to go to war and are interested in optimal re-selection rules from the principal's [citizens'] perspective. Their model, however, does not allow for an audience cost interpretation, and they do not fully characterize the optimum for their model.

[2]Jackson and Morelli (2007), study a model in which citizens always have an incentive to create more biased leaders. Unlike ours, their model does not have asymmetric information.

1

the prize. The Home leader cares both about the crisis outcome and about being retained as leader. Citizens decide whether or not to retain the leader, creating a political agency model of foreign policy decisions in the canonical form of Barro (1973), Ferejohn (1986), and Austen-Smith and Banks (1989).

We analyze the model under two different assumptions about what citizens can observe. We first study the model with private settlements, where citizens can observe crisis initiation and the fact of settlement, but not the settlement's details. Substantively, consider negotiations over classified matters, or issues that are too technical for the non-expert citizens to fully grasp. The recent Iran nuclear deal, for example, is difficult for lay people to evaluate. We then study the model with public settlements, where citizens can observe the full details of the settlement. Substantively, a settlement reseting territorial boundaries would be an example.

Our setup gives citizens an important kind of flexibility in how they punish the leader. The citizens might punish the leader for backing down, *relative to the status quo.* That is, the retention probability can be lower for a leader who settles a crisis than for a leader who stays out of the crisis in the first place. This is like the audience cost in Tomz's (2007) survey experiments. Alternatively, the citizens might punish the leader for backing down, *relative to escalating.* That is, the retention probability can be lower for a leader who settles a crisis than for a leader who rejects the settlement in favor of fighting. This is like the audience costs in Fearon's (1994a) war of attrition model. While IR scholars often think about these two notions of punishment interchangeably, they play importantly different roles in our analysis.

Punishing settlement relative to the status quo is always part of an optimal retention strategy. To see why, notice that the stronger Foreign believes Home must be to initiate a crisis, the more Foreign fears fighting. This fear in turn makes Foreign more generous in bargaining. Home would thus like to convince Foreign that it will only initiate a crisis when it is strong, since this keeps offers high. But, strength is private information, so if Home convinces Foreign to make a generous offer, there will be an ex-post incentive for the leader to start a crisis even when weak. The Home citizens can eliminate this ex-post problem by punishing settlement relative to the status quo.

Whether to punish settlement relative to escalation requires more nuance. Take this in two steps, starting with secret settlements. Political bias means the leader's private cost/benefit ratio for war differs from the citizens'. Thus, absent retention

incentives, that leader rejects too many offers. Citizens can correct this misalignment if they reward settlement relative to escalation. But leaping to the conclusion that settlement must be rewarded, relative to escalation, is a mistake. Foreign's offer might depend on the Home leader's payoff difference between settlement and escalation. The optimal retention strategy must trade off the effects on each of these margins. With high aggregate war costs, the optimal retention strategy leads to peace for sure. In this high cost case, punishing settlement enhances Home's bargaining power without risking war. But with low aggregate war costs, the optimal retention strategy allows both settlement and war with positive probability. In this low cost case, the logic of correcting the leaders' political bias is correct, and the leader is indeed rewarded for settlement.

Now consider public settlements. Here, citizens observe the details of any settlement and can condition retention on these details. This added flexibility greatly increases how much Home can extract in bargaining with Foreign. Foreign's option to fight puts a natural upper bound on how much Home can extract. In Section 4 we show that Home can push Foreign all the way to that constraint: The optimal retention strategy induces Foreign to make an offer so generous that it is indifferent between what is left over and fighting with probability one.

These results help sort out the foundations of the literature on audience costs.[3] Fearon, in his seminal paper, gave an informal optimality-based defense of the assumption:

> The [audience cost] results here suggest that ..., if the principal [read citizen] could design a 'wage contract' for the foreign policy agent, the principal would want to commit to punishing the agent for escalating a crisis and then backing down. (Fearon, 1994a, p. 581)

While plausible, this defense has not convinced all scholars. For example, Schultz asked:

> [w]hy voters would punish their leaders for getting caught in a bluff, if bluffing is sometimes an optimal strategy. After all, anyone who has ever

---

[3]This work includes theory (Fearon, 1994a; Smith, 1998; Schultz, 1999), empirics (Schultz, 2001b; Tomz, 2007), and applications to everything from causes of war (Bueno de Mesquita and Siverson, 1995) to dispute settlement in GATT (Busch, 2000) to the role of regional organizations in the consolidation on new democracy (Pevehouse, 2002).

3

> played poker understands that bluffing is not always undesirable behav-
> ior.... Clearly, additional work remains to be done on why and under
> what conditions rational voters would impose audience costs. (Schultz,
> 1999, p. 237, ft. 11)

Our model of optimal incentives shows both conditions under which Fearon's intu-
ition is right, and the limits of that intuition.[4]

In the concluding section, we reflect on what these limits imply for the empirical
literature on audience costs.

1 The Model

There are two countries, Home and Foreign. Foreign is a unitary actor, but Home
is made up of a leader, who makes international decisions, and a citizen who decides
whether or not to retain the leader.

The countries share a perfectly divisible unit of resource, with Home's status quo
share y and Foreign's status quo share $1 - y$. At the outset, Home has the option
to keep the status quo or to demand more, initiating a crisis. If it initiates, Foreign
gets to propose either war or a new allocation, $(x, 1 - x)$, with $0 \leq x \leq 1$. If Foreign
proposes an allocation and Home accepts it, then that allocation is implemented.
Otherwise (Foreign proposes war or Home rejects a proposed allocation), there is a
war. This war costs Home's citizen c_H and Foreign c_F, and the winner takes all of
the territory. We assume that $c_H \leq 1 - y$ and $c_F \leq 1 - y$.

In a war, the winner is determined by the relative strengths of the countries.
No player knows this quantity for sure. Without loss of generality, the Home leader
gets a signal $p \in [0, 1]$, which we can interpret as his posterior probability of victory
in war. The prior distribution of this posterior probability is uniform on $[0, 1]$.[5]

---

[4]Smith (1998) and Slantchev (2006) are also interested in equilibrium models of audience costs,
but they consider how rational voters might respond to various crisis strategies of leaders to attempt
to screen out decision-makers based on their leadership quality. Those models, therefore, ask if
actions that are based on optimal leader selection might be observationally equivalent to audience
costs. A similar story comes out of Hess and Orphanides (2001), where leaders also have quality
types. We, on the other hand, focus on isolating the agency problem and have only one type of
leader quality. This allows us to address the direct question regarding how we might optimally
incentivize a leader to use her private information about the outcome from war in the best way,
from the citizen's perspective, even if the leader and the citizen have different preferences when it
comes to war.

[5]This information structure is consistent with a model with two payoff-relevant states: $\theta \in$

After the crisis, the Home citizen decides to retain or dismiss the Home leader. We consider two different specifications of what the citizen observes before making this decision. In the case of **secret settlements**, the citizen observes whether the Home leader kept the status quo and, if not, whether a settlement was reached or a war was fought. In the case of **public settlements**, by contrast, the Home citizen observes the precise allocation that a settlement calls for.

All players evaluate outcomes based on the final allocation of the territory and whether or not there is a war; in addition, the Home leader prefers retaining office to losing it. To specify payoffs formally, we use the following notation: π is Home's final share of the territory, w is an indicator function taking the value 1 if there is a war and zero otherwise, and ρ is an indicator function taking the value 1 if the leader is retained and 0 otherwise. Foreign ranks outcomes according to the expectation of $(1 - \pi) - wc_F$; the Home citizen ranks outcomes according to the expectation of $\pi - wc_H$; and the Home leader ranks outcomes according to the expectation of $\pi - w\gamma c_H + \rho$, where γ is a parameter less than or equal to 1

For a fixed reward strategy followed by the citizen, there is an extensive form game of incomplete information played between the leader of Home and Foreign. We first fix an arbitrary retention strategy and characterize the perfect Bayesian equilibria of this game. Then, following the lead of Fearon (1994a), we characterize the reward strategy whose associated equilibrium maximizes the citizen's ex-ante welfare.

We focus on equilibria of a particularly simple form. Say that a strategy for the Home leader is **monotone** if:

(i) Home initiates with signal p' and $p > p'$ together imply that Home initiates with signal p, and

(ii) Home rejects offer x' with signal p' and $p > p'$ together imply that Home rejects offer x' with signal p.

We will study perfect Bayesian equilibrium in which the Home leader's strategy is monotone. Lemma 1 (below) implies that any PBE satisfies the second condition, but there might be equilibria that involve non-monotone entry decisions. However,

$\{H, F\}$. In state H, Home wins a war if it happens, and in state F, Foreign wins a war if it happens. When the signals have conditional densities $f(p \mid \theta = H) = 2p$ and $f(p \mid \theta = F) = 2(1 - p)$ a simple application of Bayes's rule gives the posterior probability that the state is H is $\Pr(\theta = H \mid p) = p$. The prior distribution of this posterior probability is uniform on $[0, 1]$.

these profiles would not be equilibria if the model were modified so that entry always carried some very small risk of war, independent of the settlement proposed by Foreign. Thus we focus on monotone equilibria for the remainder of the analysis.

Formally, a **monotone assessment** is a profile $\big((\underline{p}, \overline{p}(\cdot)), (\mu, x)\big)$, where:

- the Home leader initiates if $p \geq \underline{p}$,

- the Home leader accepts offer x' if $p \geq \overline{p}(x')$,

- Foreign believes that a Home leader who initiates has p distributed uniformly on $[\mu, 1]$,

- Foreign offers x.

A monotone assessment is a **monotone equilibrium** if:

- the strategy profile $(\underline{p}, \overline{p}(\cdot), x)$ is sequentially rational given beliefs μ, and

- $\mu = \underline{p}$.

Let a tuple of retention probabilities $\mathbf{r} = (r_Q, r_S, r_W)$ represent the probability that the citizen retains the Home leader under the status quo, settlement, and war, respectively. (In the case of public settlements, r_S is a function from $[0, 1]$ to $[0, 1]$.) Write $\mathcal{E}(\mathbf{r})$ for the set of strategy profiles associated with monotone equilibria of the game induced by $\mathbf{r}$.

We will use the following terminology to describe the qualitative features of retention strategies. A retention strategy **punishes backing down relative to fighting** if $r_W > r_S$. A retention strategy **punishes backing down relative to the status quo** if $r_Q > r_S$.

1.1 Interpreting the Assumptions

Before turning to the analysis, four of our assumptions require further comment.

First, the parameter γ in the Home leader's payoff measures the degree of conflict of interest between the leader and the citizen. Such conflicts make the leader more eager than the citizen to initiate a war. This is natural when the leader has full access to the spoils of war but does not do the actual fighting (Bueno de Mesquita, Siverson and Woller, 1992; Goemans, 2000; Chiozza and Goemans, 2004). Differences in γ may also reflect institutional differences—compulsory, universal military service, for

example, should induce high values of γ (at least for leaders with combat-aged children), while an all volunteer military might insulate leaders more from the costs of war. Either way, divergent preferences between citizens and leaders is key to our model.

Second, we will think of $c_H + c_F$ as the total cost of war. This might seem to ignore the Home leader's cost of war, but we do not want to take the three player structure that literally. Instead, we interpret each country as a unit mass of citizens, with c_H and c_F the per capita costs of fighting. The Home leader is bears a lower cost than the average citizen, captured by γ. Because she is only one person in a very large population, per capita costs are unaffected.

Third, the assumption of private settlements captures any setting in which citizens cannot condition on the full details of settlements. This is certainly the case when those details are literally unobservable, as is the case for classified concessions of the sort involved in the Cuban missile crisis. But the assumption can also capture settings in which citizens can observe the settlement, but do not know enough to evaluate it. For example, citizens will have a hard time evaluating the consequences of detailed nuclear negotiations.

Fourth, if a war is fought, we do not allow the citizen to condition retention on which side wins. Although it would be more natural to assume that the citizen can also observe the outcome of a war, we argue in the supplementary Appendix that there is no real loss of substantive insight from ignoring this possibility.

The key intuition for why we can ignore the winning and losing of wars is that a citizen in our model actually wants to reward the leader for losing the war. Intuitively, this gives a relatively weak type a kind of insurance policy against adverse war outcomes. Insurance makes her more willing to fight, which, in turn, increases the offer Foreign is willing to make.

In our view, this result pushes too hard on the fact that the war outcome does not depend on any action of the Leader. If it did, rewarding losers would create a moral hazard problem that would work against the citizen's interest. As such, a reasonable model that allowed for different retention probabilities for winners and losers would involve a constraint to ensure that winners are retained with at least as high a probability as losers. With such a constraint, the optimum reward scheme will result in the same reward for winners and losers. To keep the notation simple, we have simply imposed that equality in the main text.

2 Secret Settlements: Crisis Equilibrium

It's easiest to start with the case of secret settlements. We will handle this is two steps. This section characterizes equilibrium for an arbitrary retention rule. The following section characterizes the retention rule that leads to the best equilibrium for the Home citizen.

So let $\mathbf{r} = (r_Q, r_S, r_W) \in [0,1]^3$ be arbitrary. How will the crisis unfold?

Start at the end of the game. Home will accept the offer $(x, 1-x)$ exactly when $x + r_S \geq p + r_W - \gamma c_H$. Solve this to establish:

Lemma 1. *Home accepts x if and only if*

$$p \leq \overline{p}^*(x) \equiv \min\langle 1, x + (r_S - r_W) + \gamma c_H \rangle.$$

Next we consider Foreign's optimal offer. In a monotone equilibrium, Foreign believes that, conditional on initiating, Home's signal is uniform on $[\mu, 1]$, for some μ. So if Foreign offers $(x, 1-x)$, it's payoff is

$$
\begin{aligned}
U(x) = {}& \Pr(p \leq \overline{p}^*(x) \mid p \geq \mu)(1 - x) \\
& + \Pr(p > \overline{p}^*(x) \mid p \geq \mu)\left(\mathbb{E}(1 - p \mid p \geq \overline{p}^*(x)) - c_F\right).
\end{aligned}
\tag{1}
$$

The first term is the settlement times the probability of acceptance, while the second term is the expected payoff conditional on war times the complementary probability.

The function U breaks naturally into three components:

(i) If $x \geq 1 - (r_S - r_W) - \gamma c_H$, the offer is accepted for sure, and $U(x) = 1 - x$.

(ii) If $x \leq \mu - (r_S - r_W) - \gamma c_H$, the offer is rejected for sure, and $U(x) = 1 - \mathbb{E}(p \mid p \geq \mu) - c_F$, a constant.

(iii) If $\mu - (r_S - r_W) - \gamma c_H < x < 1 - (r_S - r_W) - \gamma c_H$, both acceptance and war have positive probability and U is equal to the quadratic

$$Q(x) = (1 - x)\int_\mu^{\overline{p}(x)} \frac{dt}{1 - \mu} + \int_{\overline{p}(x)}^1 (1 - t - c_F)\frac{dt}{1 - \mu}.$$

The basic idea of the optimal offer is then clear. It would be foolish to offer more than $1 - (r_S - r_W) - \gamma c_H$. Doing so would amount to giving additional territory to

an opponent who was going to accept anyway. Given that upper bound, the offer should maximize Q. We then have:

Lemma 2. *Let*

$$x^* = \min\langle \mu + c_F, 1 - (r_S - r_W) - \gamma c_H \rangle.$$

(i) *If $x^* > \mu - (r_S - r_W) - \gamma c_H$, then x^* is the unique optimal offer.*

(ii) *If $x^* \leq \mu - (r_S - r_W) - \gamma c_H$, than any $x \leq \mu - (r_S - r_W) - \gamma c_H$ is an optimal offer. In particular, x^* is optimal.*

(Proofs omitted from the main text are the Appendix.)

Given an equilibrium offer (x^*) and private information (p), we can calculate Home's continuation value of starting a crisis as

$$J(p, x^*) = \begin{cases} p + r_W - \gamma c_H & \text{if } p > x^* + (r_S - r_W) + \gamma c_H \\ x^* + r_S & \text{otherwise} \end{cases}.$$

This function is graphed in Figure 1. The dashed lines represent Home's utilities to settlement and war as a function of their private information about success in war. Home will enter if $J(p, x^*) > y + r_Q$, will not enter if $J(p, x^*) < y + r_Q$, and is indifferent if $J(p, x^*) = y + r^S$. As an immediate implication we have:

Lemma 3. *In any equilibrium in which there is positive probability that Home keeps the status quo and positive probability that Home enters and accepts the offer,*

$$y + r_Q = x^* + r_S.$$

To help show how these components fit together, we compute an equilibrium for the simplest case.

Example 1. Suppose $r_Q = r_S = r_W = \gamma = 1$, and $c_H = c_F < y$. H will accept the offer x whenever $p \leq \overline{p}^*(x) = x + c_H$. Guess that the equilibrium has some (but not all) types initiate and some (but not all) of those accept the appeasement offer. That means we must find a x^*, μ, and $\underline{p}^*$ such that $x^* = \mu + c_F$, $y = x^*$, and $\mu = \underline{p}^*$. (The last equality is consistence of Foreign's beliefs with the Home
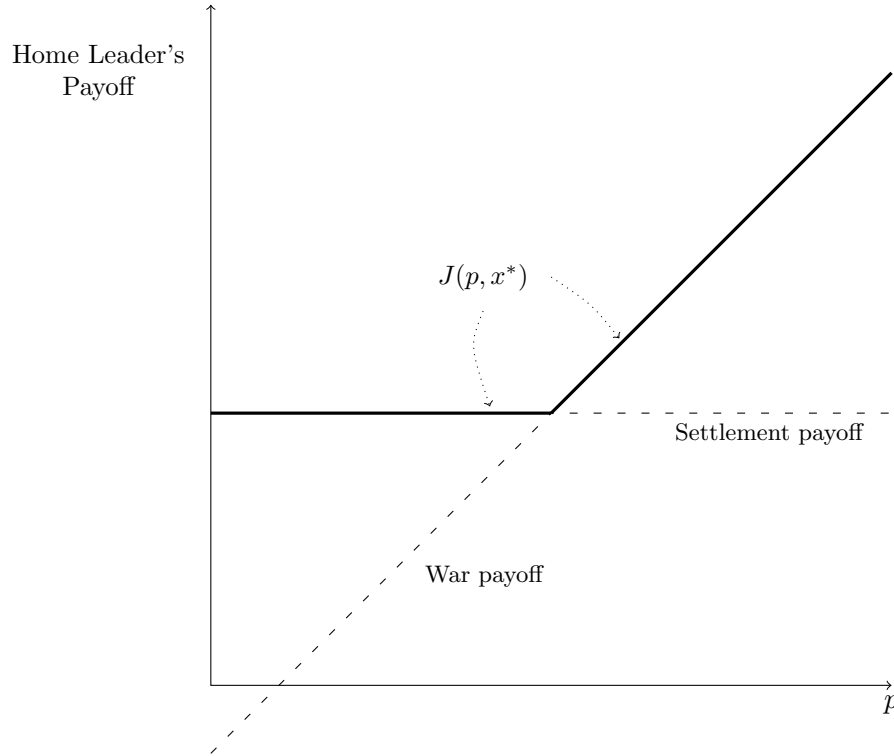
Figure 1: $J(p, x^*)$, in bold, is the upper envelope of the Home leader's payoffs to settlement, $x^* + r_S$, and to war, $p - \gamma c_H + r_W$.

leader's strategy.) That's easy: $x^* = y$, $\underline{p}^* = y - c_F$, and $\mu = y - c_F$. By sequential rationality, we now have an equilibrium, unless $\underline{p} > \overline{p}^*(x^*)$, invalidating our guess. But $\overline{p}^*(x^*) = y + c_H > y - c_F$, so we do in fact have an equilibrium.

Constructive arguments like that used in the example can be used to establish that, for any $\mathbf{r}$, there exists a pure-strategy monotone equilibrium. Rather than characterize this entire equilibrium correspondence, we ask: which is optimal?

3   Secret Settlements: Optimal Retention

The Home citizen's optimal retention strategy is the one that induces an equilibrium maximizing the citizen's ex-ante payoff. Write $\underline{p}$ for the least type that initiates, x for the offer, and $\overline{p}$ for the least type that fights. Then the citizen's expected payoff

10

is

$$U(\underline{p}, \overline{p}, x) = \underline{p}y + (\overline{p} - \underline{p})x + \int_{\overline{p}}^{1} (t - c_H)\, dt. \tag{2}$$

The first term is the probability that Home keeps the status quo times Home's status quo share, the second term is the probability that Home's leader initiates a crisis and then accepts the appeasement offer times the value of that offer, and the third term is the expected payoff on the event that there is war. We do not rule out the possibility that $\underline{p} = \overline{p}$ or that one or both of the cutpoints is 0 or 1, so this payoff function covers all of the cases discussed above.

3.1 Incentives for fixed x

As a benchmark, it's helpful to calculate the optimal strategy when x is fixed. In this case, the only problem the citizen needs to solve is the discrepancy between her own cost of fighting and the leader's cost of fighting. Comparing this solution to the full solution discussed below will highlight the role strategic bargaining between Home and Foreign in shaping incentives.

Start at the accept/reject decision. The citizen gets $p - c_H$ from rejection and x from acceptance. Thus the appropriate critical type for acceptance is the type that makes the citizen indifferent: $p^* = x^* + c_H$. We can make the leader implement this rule by setting $r_S - r_W = (1 - \gamma)c_H$. This is quite intuitive—the re-selection differential makes the ruler exactly internalize the extra cost borne by the citizen in the case of war.

Now roll back to the entry decision. Write the continuation value for the game conditional on choosing to enter as $\hat{J}(p, x^*) = \max\langle x^*, p - c_H \rangle$. The citizen would not enter iff $y \geq \hat{J}(p, x^*)$. Therefore, to implement a play of the game such that the leader of Home does exactly what the citizen would do at each instance, it suffices to take $r_Q - r_S = 0$ and $r_S - r_W = (1 - \gamma)c_H$. Such a strategy looks like this: choose any number $\kappa \in [(1 - \gamma)c_H, 1]$, and set $r_Q = r_S = \kappa$ and $r_W = \kappa - (1 - \gamma)c_H$.

In this benchmark, $r_S > r_W$—the citizen rewards settlement relative to war. This is a direct response to the leader's political bias. It can be thought of as a scheme that results in an apparently "dovish electorate" observed by Snyder and Borghard (2011).

Things get more interesting when the offer is not fixed. Then, the Home leader's incentives can also be used to affect the other state's behavior. The optimal strategy

in our model responds to exactly that incentive: the citizen manipulates x^* through her choice of $\mathbf{r}$. Recall that

$$x^* = \min\langle \underline{p} + c_F, 1 - (r_S - r_W) - \gamma c_H \rangle.$$

The two arguments of the max represent two different ways the Home citizen can manipulate x^*. If some offers are rejected, then the citizen can manipulate x^* only by changing the critical type who enters. If all offers are accepted, the Home citizen can manipulate the offer by changing $r_S - r_W$.

The rest of this section shows how the citizen's ex-ante optimal incentive strategy balances these two considerations.

3.2 The full optimum

To characterize optimal retention strategy, we maximize the citizen's payoff (given by (2)), subject to the constraints that:

(i) the retention rule is feasible: $\mathbf{r} \in [0,1]^3$,

(ii) the strategies are feasible: $\underline{p} \in [0,1]$, $x \in [0,1]$, and $\bar{p}(x) \in [0,1]$ for all $x \in [0,1]$, and

(iii) the strategies of H and F are part of an assessment that is a equilibrium: there is a μ such that $\big((\underline{p}, \bar{p}(\cdot)), (x, \mu)\big) \in \mathcal{E}(\mathbf{r})$.

Write $\mathcal{F}$ for the set of $(\mathbf{r}, \underline{p}, \bar{p}(\cdot), x)$ satisfying the first two constraints. Consider the Program:

$$\max_{(\mathbf{r}, \underline{p}, \bar{p}(\cdot), x, \mu) \in \mathcal{F}} \underline{p}y + (\bar{p}(x) - \underline{p})x + \int_{\overline{p}}^{1} (t - c_H)\, dt \tag{3}$$
$$\text{st } \big((\underline{p}, \bar{p}(\cdot)), (x, \mu)\big) \in \mathcal{E}(\mathbf{r}).$$

A retention strategy $\mathbf{r}$ is **optimal** if there is an equilibrium $\big((\underline{p}, \bar{p}(\cdot)), (x, \mu)\big)$ such that $\big(\mathbf{r}, \underline{p}, \bar{p}(\cdot), x, \mu\big)$ solves Program 3.

Proposition 1.

(i) Suppose $2c_H + c_F < 1 - y$. Then a retention strategy is optimal if and only if

it is of the form

$$r_Q = \kappa \quad r_S = \kappa - c_H - c_F \quad r_W = \kappa - (2 - \gamma)c_H - c_F$$

for some $\kappa \in [(2 - \gamma)c_H + c_F, 1]$.

In this case, the induced equilibrium has a positive probability of both initiation followed by backing down and initiation followed by war: $0 < \underline{p} < \overline{p}(x) < 1$.

(ii) *Suppose $2c_H + c_F > 1 - y$. Then a retention strategy is optimal if and only if it is of the form*

$$r_Q = \kappa \quad r_S = \kappa - \tfrac{1}{2}(1 - y + c_F) \quad r_W = \kappa - (1 - y - \gamma c_H)$$

for some $\kappa \in \left[\max\left\langle \tfrac{1}{2}(1 - y + c_F), 1 - y - \gamma c_H \right\rangle, 1\right]$.

In this case, the induced equilibrium has positive probability of initiation followed by backing down occurs with positive probability, but probability zero of initiation followed by war: $0 < \underline{p} < \overline{p}(x) = 1$.

The proof that these are the optimal schemes is somewhat involved, so we defer the details to Appendix B. We focus here on the interpretation. To help elucidate the implications of the optimal strategy, we present the salient implications of the characterization as a series of facts.

First we highlight the behavior is induced by the optimal strategy.

Fact 1. *The optimal strategy always induces positive probability of initiating a crisis and of backing down.*

Every optimal strategy generates crises with positive probability. Importantly, the probability that the crisis results in a war varies with the underlying parameters. This fact speaks directly to the concern found in the audience cost literature noted by Gowa (1999) and Schultz (2001*b*). Citizens want their leaders to "bluff" strategically by initiating crises that end peacefully, but may or may not reward settling more than fighting.

Fact 2. *The optimal strategy induces war with positive probability if and only if the total cost of war are low enough $(2c_H + c_F < (1 - y))$.*

That war is possible only when total costs are low is quite intuitive. To see why, recall that war is avoided with probability 1 only when the appeasement offer is so high that all types of Home accept. Since the optimal offer is

$$x^* = \min\langle \underline{p} + c_F, 1 - (r_S - r_W) - \gamma c_H \rangle,$$

we see that increasing F's cost raises the offer (ignoring the cap), while increasing H's cost lowers the cap. Both changes tend to make the cap binding.

The possibility of war in the optimal incentive strategy casts doubt on the classical liberal argument that, if the citizens of a country were in control, the country would be peaceful because it is the citizens who pay the "price of war in blood and money" (Russet, 1993, p.30).[6] Political control by the citizens does not eliminate risky behavior. Citizen re-election schemes respond to the risk-reward trade off common in unitary actor models of crisis and war.

Interestingly, the decision about whether or not to use a strategy with positive probability of war is independent of the preference divergence between the leader and the citizen. The form of the optimal strategy, on the other hand, does depend on the preference divergence. This is because the citizen uses the reward for settlement to get the degree of political bias that is optimal for manipulating the offer. How much of that optimal level of political bias needs to come from the reward strategy obviously depends on how much political bias is inherent in the preference divergence.

The form of the optimal retention strategy depends on what kind of behavior the citizen wants to induce.

From Proposition 1, we have $r_S - r_W < 0$ if and only if $1 - y < 2\gamma c_H + c_F$. Thus:

Fact 3. *The optimal strategy punishes backing down relative to fighting if and only if the total costs of war are high enough.*

Fact 3 is illustrated in Figure 2.

In the audience cost case, the citizen wants to support certain settlement at the highest possible level. To do this, she needs to punish backing down just enough to offset the leader's cost of fighting. Here the incentives work in a manner reminiscent of the intuitions of Fearon (1994a) and Schultz (2001b): they make backing down

---

[6]Also see Kant (1903); Snyder (1991).

costly, stiffen the leader's stance in bargaining, and lead to a bigger share of the pie. In fact, the optimal audience costs exactly offsets the leader's cost of fighting, and Home gets everything when it initiates a crisis.

In the internalized costs case, on the other hand, the leader will choose to fight with positive probability. And the set of offers he rejects does not affect the offer that is actually made. Thus the citizen wants to offset the existing political bias by rewarding backing down relative to fighting. After the crisis has begun, the citizen wants the leader to internalize fully the cost of war. For $\gamma < 0$, this means making war less attractive than a settlement. In the space between there is no war, but the citizen does not punish settlement relative to war.
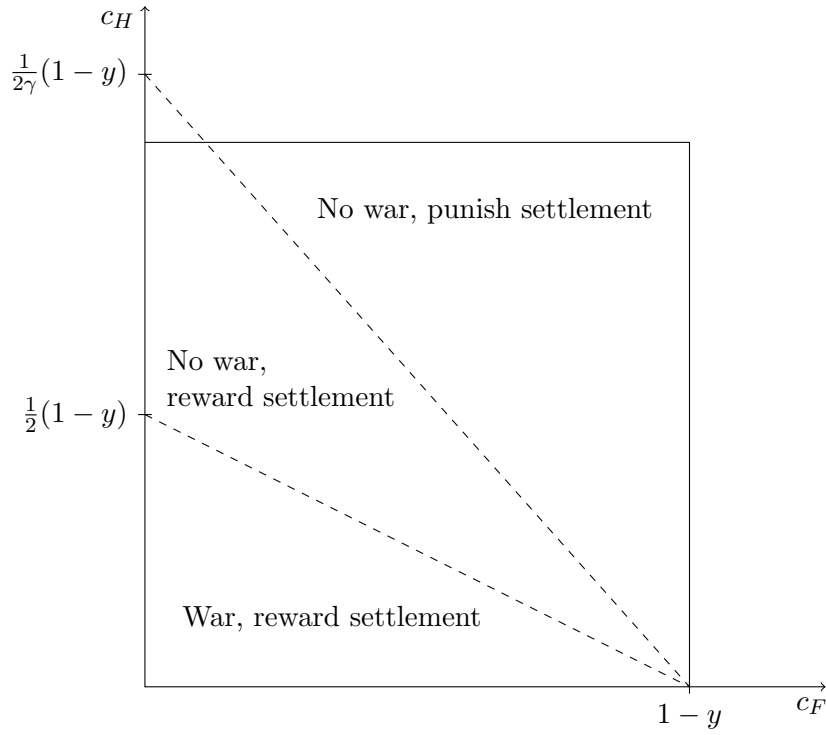

Figure 2: Characteristics of the equilibrium with the optimal retention strategy, as a function of the costs of fighting. The box delimits the set of costs consistent with the assumption that $c_H \leq 1 - y$ and $c_F \leq 1 - y$.

Whether or not backing down is rewarded relative to fighting depends on the underlying parameters. By contrast, r_Q and r_S are unambiguously ordered:

Fact 4. *The optimal strategy always punishes backing down relative to the status quo.*

Whatever offer the citizen is trying to extract form F, initiation must be limited. In the internalized cost regime, this is obvious: the offer is an increasing function of the least type who enters. The argument for the audience cost regime is a bit more subtle, but just as intuitive: if F is making a very generous offer, then weak types will want to initiate a crisis to take advantage of it. But if F expects that reaction, it will no longer be willing to make the generous offer. Thus offers can be generous in equilibrium only if accepting those generous offers is costly to leaders. This robust feature of optimal schemes is the one for which Tomz (2007) finds support in his survey experiments on audience costs.

Finally, the model's comparative statics shed some light on when Fearon-Schultz audience costs, i.e. those where backing down leads to a lower probability of retention than fighting, are a good incentive strategy for citizens to use.

Fact 5. *As Home's status quo share increases, the audience cost regime becomes more likely, in the sense that more combinations of c_H and c_F make the audience cost regime optimal.*

We should expect to see evidence of Fearon-Schultz audience costs where backing down (settling) is punished electorally compared to fighting, in either the post-Cold War period for the U.S. or in conflicts where there is large asymmetry in favor of the democratic initiator under the status quo. This might explain the apparent lack of audience costs in the Suez crisis, as the status quo was very unfavorable to the challengers, or in the Cuban Missile Crisis, with the Soviet weapons already on the island, in Snyder and Borghard (2011).

4   Public Settlements

The previous Section characterized the retention strategy a citizen would optimally commit to when the details of settlements are secret. But it is also of interest to characterize optimal incentives when such details are public.

When settlements are public, a retention strategy no longer specifies a single number r_S for the probability of retention in the event of a settlement. Instead, it

specifies a function $r_S : [0,1] \to [0,1]$, where $r_S(x)$ is the probability of retention when the settlement is x.

We will restrict attention to **cutoff reward schemes**—functions of the form

$$r_S(x) = \begin{cases} \overline{r} & \text{if } x \geq \overline{x} \\ \underline{r} & \text{if } x < \overline{x} \end{cases}$$

for some $\overline{x} \in [0,1]$ and $\underline{r}, \overline{r} \in [0,1]$ with $\underline{r} \leq \overline{r}$.

The apparent restrictiveness of this family of retention strategies will not turn out to limit what the Home citizen can achieve. Indeed, we will show that the Home citizen can use a strategy from this family to drive Foreign all the way to indifference between settling and fighting eventhough Foreign has all the bargaining power.

4.1 What Payoffs are Possible?

In any equilibrium, an interval of types $[\underline{p}, 1]$ will initiate a crisis, while the complementary interval $[0, \underline{p})$ will keep the status quo. The search for the best possible equilibrium payoff for the Home citizen can be carried out in two steps, one that asks what can be attained given $\underline{p}$, and a second that asks which $\underline{p}$ is best, given the first answer.

Suppose that Foreign believes that H initiates a crisis if and only if $p \geq \mu$. If F follows the strategy Fight, it gets payoff

$$\mathcal{W}_F(\mu) = 1 - \mathbb{E}(p \mid p \geq \mu) - c_F$$
$$= \frac{1 - \mu}{2} - c_F,$$

where the second equality uses the uniform distribution of p.

Since this strategy is feasible, F's payoff in any equilibrium must be at least $\mathcal{W}_F(\underline{p})$. And since the maximal total surplus is 1, the Home citizen's payoff from a continuation equilibrium after all types $p \geq \underline{p}$ have initiated is at most:

$$\mathcal{I}(\underline{p}) = \min\langle 1, 1 - \mathcal{W}_F(\underline{p})\rangle$$
$$= \min\left\langle 1, \frac{1 + \underline{p}}{2} + c_F \right\rangle$$

Call a pair (r_S, r_W) **fully extractive at** $\underline{p}$ if, given (r_S, r_W), equilibrium play in

17

the continuation game gives H payoff $\mathcal{I}(\underline{p})$. Because war is costly, Foreign's best response to a fully extractive pair must be to offer $x^{\dagger} = \mathcal{I}(\underline{p})$.

Clearly, the best payoff that the Home citizen can hope for in an equilibrium with initiation cutoff $\underline{p}$ is $\underline{p}y + (1 - \underline{p})\mathcal{I}(\underline{p})$. And the best possible payoff any equilibrium could give is

$$\max_{\underline{p} \in [0,1]} \underline{p}y + (1 - \underline{p})\mathcal{I}(\underline{p}). \tag{4}$$

Call a retention strategy **maximally extractive** if it induces an equilibrium in which the Home citizen's payoff attains the maximum in Program 4.

4.2 How to Extract the Surplus

As in the case of secret settlements, a key step is to find out which types accept a given offer.

Type p is willing to accept offer x if and only if $x + r_S(x) \geq p - \gamma c_H + r_W$. For a cutoff retention scheme, there are two cases. If $x < \overline{x}$, then the offer is accepted by types $p \leq x + \gamma c_H - r_W + \underline{r}$. If $x \geq \overline{x}$, then the offer is accepted by types $p \leq x + \gamma c_H - r_W + \overline{r}$.

Again, write the acceptance strategy as a critical type, $\overline{p}(x)$ such that types $p \leq \overline{p}(x)$ accept and types $p > \overline{p}(x)$ reject. This will be easier to write with the function $\iota_{\underline{p}}$ given by

$$\iota_{\underline{p}}(x) = \begin{cases} 1 & \text{if } 1 \leq x \\ x & \text{if } \underline{p} < x < 1 \\ \underline{p} & \text{if } x \leq \underline{p} \end{cases}.$$

Now we can write the equilibrium acceptance strategy:

Lemma 4. *Fix a cutoff reward scheme (r_S, r_W). In any equilibrium in which the Home leader initiates a crisis if and only if $p \geq \underline{p}$, the Home leader's acceptance strategy is given by:*

$$\overline{p}^{*}(x) = \begin{cases} \iota_{\underline{p}}(x + \gamma c_H - r_W + \underline{r}) & \text{if } x < \overline{x} \\ \iota_{\underline{p}}(x + \gamma c_H - r_W + \overline{r}) & \text{if } x \geq \overline{x} \end{cases}.$$

Figure 3 shows that, if $\overline{r} > \underline{r}$, there is a discrete jump in the critical type at $\overline{x}$. Moreover, by taking $\overline{r} - \underline{r} \geq 1 - \underline{p}$, that jump can be made to cover the entire range of possible types. This makes it easy to get full extraction.
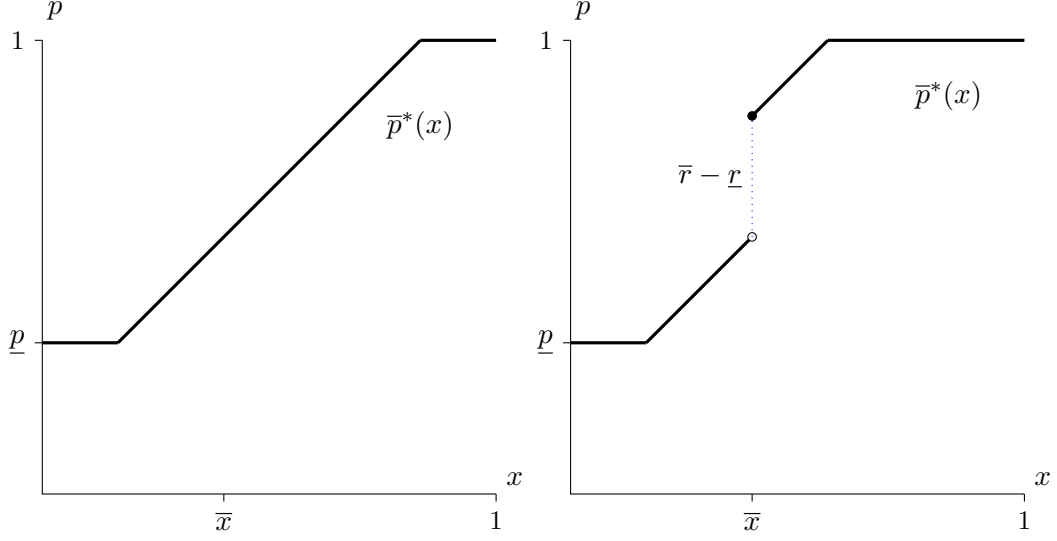
Figure 3: Each panel shows the Home leader's acceptance function, $\bar{p}$, for a cutoff retention strategy with $\bar{x}$. The panel on the left has $\underline{r} = \bar{r} = 0$. The panel on the right has $\bar{r} > \underline{r} = 0$.

Example 2. Let $x^\dagger = \min\langle 1, 1 - \mathcal{W}_F(\underline{p})\rangle$, and suppose $\underline{p} \leq x^\dagger + \gamma c_H \leq 1$. Set $\bar{r} = 1$, $\underline{r} = \underline{p}$, and $r_W = x^\dagger + \gamma c_H$. The equilibrium acceptance strategy simplifies to:

$$\bar{p}^*(x) = \begin{cases} \underline{p} & \text{if } x < x^\dagger \\ 1 & \text{if } x \geq x^\dagger \end{cases}.$$

That is, any relevant type is willing to accept an offer of all of the conditional surplus, and none is willing to accept anything less. Clearly, Foreign will offer $x^\dagger$, and it will be accepted.

While the argument is not always as simple as the one just given, it turns out that the Home citizen can always fully extract at $\underline{p}$.

Proposition 2. *Fix $\underline{p}$. There is a pair (r_S, r_W) that is fully extractive at $\underline{p}$.*

The proof of Proposition 2 is constructive—for any values of the parameters and $\underline{p}$, we present a concrete pair (r_S, r_W) that is fully extractive. In those constructions, we have $\underline{r} = 0 < r_W$. Substantively, this says that the Home leader would be punished for settlement, relative to war, were he to settle for less than $x^\dagger$. Impor-

tantly, this punishment is off the path, and would not be observable by a researcher studying outcomes of this game.

While the proof of Proposition 2 uses a pair with off-the-path punishments, it leaves open the question of whether such punishments are necessary for full extraction. The next result shows that they are, at least in our class of cutoff reward schemes. (Notice, Example 2 has such off-path punishments but does not use $\underline{r} = 0$.)

Proposition 3. *Fix $\underline{p}$ and let the cutoff reward scheme (r_S, r_W) be a fully extractive pair. Then $r_W > \underline{r}$.*

4.3 The Full Optimum

Proposition 2 shows that Home can full extract the surplus that is available conditional on (equilibrium) crisis initiation. But a full discussion of the optimal retention strategy requires attention also be paid to which types initiate a crisis. The following example shows that keeping $\underline{p} > 0$ can strictly increase the Home citizen's ex-ante expected payoff, even when Home's status quo share is lower than the fully extractive offer that would be made conditional on $\underline{p} = 0$.

Example 3. Suppose $c_F = \frac{1}{3}$ and $y = \frac{1}{2}$. If all types of Home initiate a crisis, then the Home citizen's ex-ante expected payoff is $1 - (\mathbb{E}(1-p) - c_F) = \frac{5}{6}$. If Home initiates a crisis if and only if $p > \frac{1}{6}$, then the fully-extractive offer is

$$ x^\dagger = \frac{1 + \underline{p}}{2} + c_F = \frac{11}{12}. $$

Thus the ex-ante expected payoff to the Home citizen is

$$ \underline{p}y + (1 - \underline{p})x^\dagger = \frac{61}{72} > \frac{5}{6}. $$

Example 3 illustrates the importance of being able to extend a fully extractive retention scheme to a complete retention strategy that creates the appropriate incentives for crisis initiation. Not all fully extractive retention schemes can be so extended.

Example 4 (Example 2, continued). Suppose $0 < \underline{p} \le x^\dagger + \gamma c_H \le 1$, and $y < x^\dagger$. Consider the fully extractive scheme given by $\bar{r} = 1$, $\underline{r} = \underline{p}$, and $r_W = x^\dagger + \gamma c_H$. To

extend this to a full retention strategy that gives types $p < \underline{p}$ an incentive to keep the status quo, we must have $y + r_S = x^\dagger + 1$, or

$$r_S = (x^\dagger - y) + 1.$$

But this is not feasible, since r_S cannot exceed 1.

Example 4 shows that it is too much to hope for that every full-extractive scheme can be extended to a complete retention strategy. Fix some parameters y and c_F, and let $\underline{p}$ be the maximizer from Program 4. Proposition 2 shows that there are retention schemes that are fully extractive in the continuation equilibrium following crisis initiation by types $p \geq \underline{p}$. The next result shows that at least one of these fully-extractive schemes can in fact be extended to a complete retention strategy that is maximally extractive.

Proposition 4. *Fix c_F and y. There is a maximally extractive retention strategy. That strategy induces all types to initiate a crisis if and only if $y \leq c_F \leq \frac{1}{2}$.*

Figure 4 gives more detail on how the the optimal retention strategy varies as a function of c_F and y. The triangle is the set of (c_F, y) pairs that are consistent with the assumption that $c_F \leq 1 - y$. Each of the four regions is labeled according to whether all types initiate a crisis or not, and whether or not the offer is equal to 1. Observe that an offer equal to 1 implies that initiation is limited, but the converse does not hold.

5 Conclusion

We have characterized the optimal incentive strategy for a citizen to give to a ruler who might engage in a simple form of crisis bargaining, allowing us to endogenize the domestic political constraints that have played such a large role in recent IR theory. These political constraints are sensitive to two different agency problems. The first, which is well-explored in the literature, is the divergence between the leader's private payoff to war and the citizens' payoffs to war. The second, which to our knowledge is new, is a commitment problem faced by the leader. This commitment problem comes from the fact that committing to keep the status quo unless the private information received by the leaders is quite favorable for Home's prospects in war
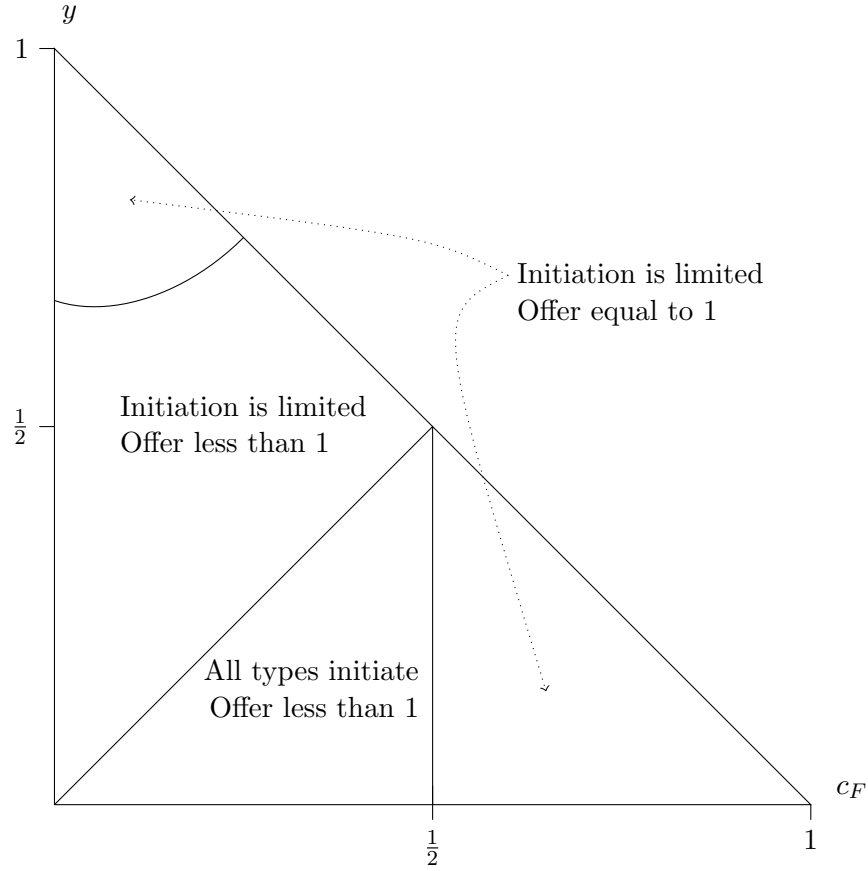
21

Figure 4: The form of maximally extractive retention strategies as a function of c_F and y.

leads to high appeasement offers in the case of initiation, but those same high offers make it attractive to initiate crisis when their signal is not so favorable.

The citizen's optimal response to the commitment problem is quite robust: the leader should alway faces a lower probability of retention after backing down than if she keeps the status quo. This allows a citizen to manage the leader's commitment problem. But the incentive we thought we knew well, the decision to enhance or offset the leader's political bias, is more sensitive to the environment. When total costs of war are low, the citizen offsets the leader's political bias by rewarding backing down more than going to war. This leads the leader to fully internalize the costs of fighting, an action she nonetheless takes with positive probability in the

22

equilibrium. When total costs of fighting are large, on the other hand, the citizen designs a strategy that leads to peace with probability one. In this case, the citizen enhances the bargaining power of the leader by punishing her for backing down rather than fighting.

While the model is much too stylized to capture the full richness of empirical discussions of audience costs, it does help clear up an ambiguity in the literature's treatment of the issue. In Fearon's (1994a) canonical theoretical discussion, as well as in the cases discussed by Schultz (2001b), audience costs refer to punishment for backing down once a crisis has started. Another take on audience costs, one tested in Tomz's (2007) survey experiments, is that they are costs associated with initiating and then backing down.[7] In terms of our model, Fearon and Schutz are discussing the contrast between retention probabilities conditional on backing down and on war, while Tomz is talking about the contrast between retention probabilities conditional on backing down and on keeping the status quo from the beginning. The model highlights that these are conceptually different parts of the optimal incentive strategy, and that whether or not each is used responds to different aspects of the agency problem.

The model also allows us to think about the relationship between incentivizing leaders, through mechanisms like audience costs, and the citizen's foreign policy preferences. Snyder and Borghard (2011) argue that voters' reaction to a leader's foreign policy actions is less about preferences for inconsistency and more about policy preferences. In our approach the re-election scheme is not based on a direct preference for consistency. Rather, it is based on managing the principal agent problem faced by voters. But we can still ask how their concern about hawkish or dovish citizen preferences relate to our approach.

What would it mean, in the context of our approach, for a voter to be hawkish or dovish? One answer would say that a voter is hawkish if they preferred to start a crisis, given the conditions, their initial beliefs about the probability of victory, and the expectations about how their leader and the foreign rival would act. They are dovish if they prefer the status quo ex ante. The voters' relevant benchmark is what foreign policy choice they would choose given their information. In terms of the formal model, a voter is hawkish if $\frac{1}{2}(1 + c_F) - c_H(1 - c_F) > y$. A voter prefers

---

[7]Fearon (1994a) comments on p.585 that audience costs may "affect entry decisions" and discusses how his model must be modified to consider this dimension of audience costs.

their leader to initiate a crisis given their information if the Foreign's costs are high, Home's costs are low, and the status quo is bad for Home.

Comparing this condition to those in Proposition 1, it is easy to see that all combinations of retention regimes and policy preferences are possible. Sometimes hawkish voters should reward peace and backing down. Sometimes dovish voters should punish settlement and reward war. The voters' policy preferences over initiating a crisis and how they should optimally use their political support to incentivize their leader in foreign policy are two separate and largely unrelated issues.

The optimal retention strategy for public settlements problematizes some of the empirical literature on punishment and audience costs. With public settlements, Home's leader is thrown out for sure if she accepts any offer leaving Foreign some surplus. This corresponds to the classical idea of punishment for initiating a crisis and then backing down. The result shows there can be a microfoundation for the standard audience cost argument without implying that punishments are ever observed. This is because these costs occur only off the equilibrium path.

There is clearly much work to be done before we have a full picture of leader and citizens' incentives for the agency problem during a crisis. One obvious question is how sensitive are the results to the precise form of uncertainty. Our model has asymmetric information about the relative strength of the two states. Few formal IR models follow such an approach; most feature asymmetric information about resolve (operationalized as uncertainty about costs of fighting). Since Fey and Ramsay (2009) show that the difference can be consequential, it is interesting to think about optimal incentives in a model with known strength but uncertain resolve.

It's also easy to think of more substantial extensions. The pure moral hazard approach we have adopted allows the citizen to exercise substantial commitment power. As Fearon (1999) forcefully argues, that commitment power will be substantially curtailed if there is heterogeneity in the pool of potential leaders. (Smith (1998) and Slantchev (2006) have constructed adverse selection models of audience costs in which commitments are not possible.) An interesting extension would ask what kinds of political institutions would best help manage these commitment problems and allow the citizen to get close to the optimal strategy we characterize here. We also hope to be able to analyze a more symmetric model, one in which both leaders are given incentives by their citizens. These extensions create a rich menu of options for further research.

A Proofs for Section 2

Proof of Lemma 2. Evaluating the integral shows that q is quadratic with the coefficient on x^2 is $-\frac{1}{2}$, so q is concave. Since the other two components of U are linear, this shows that U is strictly concave on $(\underline{p} - (r_S - r_W) - \gamma c_H, \infty)$ and constant otherwise, so it is globally quasiconcave. Furthermore, U is continuous and is differentiable except possibly at $x = \underline{p} - (r_S - r_W) - \gamma c_H$.

The next step is to look more closely at Q as a function on all of $\mathbb{R}$. Multiply through by $(1 - \underline{p})$ and differentiate (remembering that $\overline{p}'(x) = 1$) to get

$$(1 - \underline{p})Q'(x) = -(\overline{p}(x) - \underline{p}) + (1 - x) - (1 - \overline{p}(x) - c_F).$$

Equate this to 0 and solve to see that Q is maximized at $x = \underline{p} + c_F$.

This means that U is nondecreasing up to

$$x^* = \min\langle \underline{p} + c_F, 1 - (r_S - r_W) - \gamma c_H \rangle,$$

and is strictly decreasing thereafter. Thus x^* is always an optimal offer. If $x^* > \underline{p} - (r_S - r_W) - \gamma c_H$, then x^* is the unique optimizer. If $x^* \leq \underline{p} - (r_S - r_W) - \gamma c_H$, on the other hand, then any $x \leq \underline{p} - (r_S - r_W) - \gamma c_H$ is optimal. $\qquad\square$

B Proofs for Section 3

B.1 Proof of Proposition 1

We will consider two optimization problems. First, the **equilibrium problem (EP)** is to maximize the citizen's ex-ante payoff subject to the constraints that

(i) the retention rule is feasible: $\mathbf{r} \in [0, 1]^3$,

(ii) the strategies are feasible: $\underline{p} \in [0, 1]$, $x \in [0, 1]$, and $\overline{p}(x) \in [0, 1]$ for all $x \in [0, 1]$, and

(iii) the strategies of H and F are part of an assessment that is a PBE: $(\underline{p}, \overline{p}(\cdot), x) \in \mathcal{E}(\mathbf{r})$.

Write $\mathcal{F}$ for the set of $(\mathbf{r}, \underline{p}, \overline{p}(\cdot), x)$ satisfying the first two sets of constraints.

Then, by Lemmas 1–3, we can formally define EP as[8]:

$$\max_{(r_Q, r_S, r_W \underline{p}, \overline{p}(\cdot), x) \in \mathcal{F}} U(\underline{p}, \overline{p}(\cdot), x)$$

$$\text{st } \underline{p} \begin{cases} = 1 & \text{if } y + r_Q < x + r_S \\ \in [0, 1] & \text{if } y + r_Q = x + r_S \\ = 0 & \text{if } y + r_Q > x + r_S \end{cases}$$

$$\overline{p}(x) = \min\langle 1, x + (r_S - r_W) + \gamma c_H \rangle \quad \forall x \in [0, 1]$$

$$x = \min\langle \underline{p} + c_F, 1 - (r_S - r_W) - \gamma c_H \rangle$$

Let $\mathcal{V}^{EP}$ be the value of EP.

Rather than attack the EP directly, we will instead consider a simpler problem, and show that solutions to the simpler problem allow us to construct solutions to EP. Let $\Delta = r_S - r_W$, and let

$$\widetilde{U}(\Delta, x) = (x - c_F)y + ((x + \Delta + \gamma c_H) - (x - c_F)) x + \int_{(x + \Delta + \gamma c_H)}^{1} (t - c_H) \, dt,$$

and define the **relaxed problem (RP)** as:

$$\max_{(\Delta, x) \in [-1, 1] \times [0, 1]} \widetilde{U}(\Delta, x)$$

$$\text{st } x \leq 1 - \Delta - \gamma c_H$$

Let $\mathcal{V}^{RP}$ be the value of the relaxed problem.

Lemma 5. *The value of RP is an upper bound on the value of EP: $\mathcal{V}^{RP} \geq \mathcal{V}^{EP}$.*

Proof.

---

[8]It might seem that our constraints are too restrictive—Lemma 2 allows offers greater than x for some retention probabilities. But notice that that can only happen when the offer is sure to be rejected, so setting the offer arbitrarily to x in such cases does not miss any possible payoffs to the citizen.

26

(i) Start by considering the intermediate program:

$$\max_{(\Delta, \underline{p}, \hat{p}, x) \in [-1,1] \times [0,1]^3} U(\underline{p}, \hat{p}, x)$$

$$\text{st } \hat{p} = x + \Delta + \gamma c_H$$

$$x \leq \underline{p} + c_F$$

$$x \leq 1 - \Delta - \gamma c_H.$$

Inspection of the constraints in EP and this intermediate program shows that the intermediate program has weakly greater value. (Notice that the constraint $x \leq 1 - \Delta - \gamma c_H$ ensures that $\hat{p}$ is in fact less than or equal to 1.)

(ii) At any solution to the intermediate program, the constraint $x \leq \underline{p} + c_F$ must bind.

Proof. The solution neglecting both inequality constraints is to set $x = 1$, $\underline{p} = 0$, and $\Delta = -\gamma c_H$. But that violates the first inequality constraint.

Now assume that the second inequality constraint binds. In this case, the offer is $x = 1 - \Delta - \gamma c_H$. Given that, the equality constraint implies $\hat{p} = 1$, and the payoff reduces to

$$\underline{p} y + (1 - \underline{p}) x.$$

If the first inequality constraint does not bind, then the solution must be for $x = 1$ and $\underline{p} = 0$. But then $\underline{p} + c_F = c_F < 1 = x$, contradicting the claim that the first constraint is slack. $\qquad \square$

(iii) Substitute $\hat{p} = x + \Delta + \gamma c_H$ and $x = \underline{p} + c_F$ into U to get $\widetilde{U}$.

(iv) Thus RP has the same value as the intermediate program. $\qquad \square$

Lemma 6.

(i) *Fix y, c_H, and c_F. RP has a unique solution.*

(ii) *Suppose $2c_H + c_F \leq 1 - y$. Then the solution to RP is*

$$(\Delta^*, x^*) = ((1 - \gamma)c_H, y + c_H + c_F).$$

(iii) *Suppose $2c_H + c_F > 1 - y$. Then the solution to RP is*

$$(\Delta^*, x^*) = \left(\frac{1}{2}(1 - y - c_F) - \gamma c_H, \frac{1}{2}(1 + y + c_F) \right).$$

Proof.

(i) The second derivative of $\widetilde{U}$ is

$$\mathbf{D}^2 \widetilde{U}(\Delta, x) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

so $\widetilde{U}$ is strictly concave.

(ii) The unconstrained maximizer of $\widetilde{U}$ solves the first-order conditions

$$x - x - \Delta - \gamma c_H + c_H = 0$$
$$y + (\Delta + \gamma c_H + c_F) - x - \Delta - \gamma c_H + c_H = 0.$$

solve these to get

$$(\Delta^*, x^*) = ((1 - \gamma)c_H, y + c_H + c_F).$$

This is consistent with the constraint just if $x^* \leq 1 - \Delta^* - \gamma c_H$, or

$$2c_H + c_F \leq 1 - y. \tag{5}$$

At this point, all that can go wrong is that x might not be interior. This will happen if $x \geq 1$, or

$$y + c_H + c_F \geq 1.$$

But this is inconsistent with Inequality 5.

(iii) For $2c_H + c_F > 1 - y$, the unconstrained maximizer violates the constraint. This implies the constraint binds, and we can substitute the constraint into the objective to get

$$\max_{x \in [0,1]} (x - c_F)y + (1 - x + c_F)x. \tag{6}$$

Ignoring the boundary constraints on x, the unique maximizer is $x^* = \frac{1}{2}(1 + y + c_F)$. This is consistent with the boundary constraint if and only if $x^* \leq 1$, or

$$c_F \leq 1 - y$$

which is an assumption of the model.

We can then back out Δ^* from the constraint to get

$$(\Delta^*, x^*) = \left(\frac{1}{2} - \frac{1}{2}y - \frac{1}{2}c_F - \gamma c_H, \frac{1}{2}(1 + y + c_F) \right).$$

$\square$

Lemma 7. *Suppose* (Δ^*, x^*) *is a solution to RP. Then, for the retention strategy*

$$\mathbf{r}^* = (1, 1 - (x^* - y), 1 - (x^* - y) - \Delta^*),$$

there exists $\underline{p}^*$ *and* $\overline{p}^*(\cdot)$ *such that* $(\underline{p}^*, \overline{p}^*(\cdot), x^*) \in \mathcal{E}(\mathbf{r}^*)$ *and that profile gives the citizen payoff* $\widetilde{U}(\Delta, x)$.

Proof. We proceed by cases, one for each of the two statements in Lemma 6.

(i) Suppose $2c_H + c_F < 1 - y$.

From Lemmas 1 and 6, the Home leader's acceptance strategy is given by

$$\begin{aligned} \overline{p}^*(x^*) &= \min\langle 1, y + 2c_H + c_F \rangle \\ &= y + 2c_H + c_F \\ &\leq 1. \end{aligned}$$

Lemma 3 and the definition of $\mathbf{r}^*$ imply that any $\underline{p} < \overline{p}^*(x^*)$ is a best response for the Home leader. Choose $\underline{p}^* = y + c_H (< y + 2c_H + c_F = \overline{p}^*(x^*))$. Then $\min\langle \underline{p}^* + c_F, 1 - c_H \rangle = 1 + c_H + c_F = x^*$, and Lemma 2 tells us that x^* is a best response for Foreign.

The Home Citizen's payoff is

$$\underline{p}^* y + \left(\overline{p}^*(x^*) - \underline{p}^*\right) x^* + \int_{\overline{p}^*(x^*)}^{1} (t - c_H)\, dt$$

$$= (y + c_H)y + (x^* + \Delta^* + \gamma c_H - y - c_H)x^* + \int_{y+2c_H+c_F}^{1} (t - c_H)\, dt$$

$$= (x^* - c_F) + (\Delta^* + \gamma c_H + c_F)x^* + \int_{x^*+\Delta^*+\gamma c_H}^{1} (t - c_H)\, dt$$

$$= \widetilde{U}(\Delta^*, x^*),$$

where the second equality uses the definitions of Δ^* and x^*.

(ii) Suppose $2c_H + c_F \geq 1 - y$.

From Lemmas 1 and 6, the Home leader's acceptance strategy is given by $\overline{p}^*(x^*) = 1$.

Lemma 3 and the definition of $\mathbf{r}^*$ imply that any $\underline{p} < \overline{p}^*(x^*)$ is a best response for the Home leader. Choose $\underline{p}^* = y + c_H$. Then $\min\langle \underline{p}^* + c_F, \frac{1}{2}(1 + y + c_F)\rangle = x^*$, and Lemma 2 tells us that x^* is a best response for Foreign.

Lemma 3 and the definition of $\mathbf{r}^*$ imply that any $\underline{p} < \overline{p}^*(x^*)$ is a best response for the Home leader. Choose $\underline{p}^* = x^* - c_F = \frac{1}{2}(1 + y - c_F)$. Then $\min\langle \underline{p}^* + c_F, \frac{1}{2}(1 + y + c_F)\rangle = x^*$, and Lemma 2 tells us that x^* is a best response for Foreign.

The Home Citizen's payoff is

$$\underline{p}^* y + \left(\overline{p}^*(x^*) - \underline{p}^*\right) x^* + \int_{\overline{p}^*(x^*)}^{1} (t - c_H)\, dt$$

$$= (x^* - c_F)y + (1 - x^* + c_F)x^* + \int_{1}^{1} (t - c_H)\, dt$$

$$= (x^* - c_F) + (\Delta^* + \gamma c_H + c_F)x^* + \int_{x^*+\Delta^*+\gamma c_H}^{1} (t - c_H)\, dt$$

$$= \widetilde{U}(\Delta^*, x^*),$$

where the second equality uses the definitions of Δ^* and x^*. $\qquad\square$

Lemma 8. *Suppose $\mathbf{r}$ is a retention strategy and $(\underline{p}, \overline{p}(\cdot), x) \in \mathcal{E}(\mathbf{r})$. Then $(\underline{p}, \overline{p}(\cdot), x)$ gives the citizen payoff $\mathcal{V}^{RP}$ if and only if $r_Q - r_S = x - y$ and $r_S - r_W = \Delta$.*

Proof. Such an **r** induces either a different Δ or different $\underline{p}$, and thus a different x. But $\widetilde{U}$ just is the citizen's payoff as a function of Δ and x. $\qquad\square$

Proof of Proposition 1. Lemma 5 gives $\mathcal{V}^{RP} \geq \mathcal{V}^{EP}$. Since Lemma 7 gives an equilibrium with payoff $\mathcal{V}^{RP}$, we have $\mathcal{V}^{EP} \geq \mathcal{V}^{RP}$. Together, these imply $\mathcal{V}^{EP} = \mathcal{V}^{RP}$. Lemma 8 then implies that $\mathbf{r} \in [0,1]^3$ implements $\mathcal{V}^{EP}$ if and only if $r_Q - r_S = x - y$ and $r_S - r_W = \Delta$. Substituting the solutions from Lemma 6 gives the form of the optimal **r** The inequalities involving $\underline{p}$ and $\overline{p}(x)$ are immediate from the constructions in the proof of Lemma 7. $\qquad\square$

C   Proofs for Section 4

Proof of Proposition 2. Let $x^\dagger = \min\left\langle \frac{1+\underline{p}}{2} + c_F, 1 \right\rangle$, and

$$
\hat{\overline{p}}(x) = \begin{cases} x - c_F & \text{if } x < x^\dagger \\ 1 & \text{if } x \geq x^\dagger \end{cases}
$$

We will proceed in two steps. First, we will show that, if the Home leader adopts $\hat{\overline{p}}$ as her acceptance strategy, then Foreign best responds by offering $x^\dagger$. Second, we construct retention strategies that make $\hat{\overline{p}}$ a best response by the Home leader.

Step 1:   Start with two easy observations.

Claim 1. $x \geq x^\dagger$ *implies* $x \geq c_F$.

Claim 2. *Suppose the Home leader uses strategy* $\hat{\overline{p}}$. *All types accept any offer* $x \geq x^\dagger$.

These set up the main argument.

Claim 3. *If* $x < x^\dagger$ *is a best response to* $\hat{\overline{p}}$, *then it must satisfy* $x \leq \check{x} = \underline{p} + c_F$.

Proof. For any offer x, F's payoff is

$$
Q(x) = (1 - x)\frac{\hat{\overline{p}}(x) - \underline{p}}{1 - \underline{p}} + \int_{\hat{\overline{p}}(x)}^1 \frac{1 - t - c_F}{1 - \underline{p}}\, dt.
$$

31

On $(\check{x}, x^\dagger)$, then $\hat{\bar{p}}(x) = x - c_F$, and $\hat{\bar{p}}'(x) = 1$. Thus, on that interval,

$$Q'(x) = \frac{c_F - x + \underline{p}}{1 - \underline{p}} < 0,$$

so no $x \in (\check{x}, x^\dagger)$ can be a best response. $\qquad\square$

Claim 4. *At offer $x \le \check{x}$, war occurs with probability 1.*

Proof. Since $\hat{\bar{p}}$ is non-decreasing, the acceptance threshold at $x \le \check{x}$ is $\overline{p}^*(x) \le \overline{p}^*(\check{x}) = \underline{p}$. $\qquad\square$

Claims 3, 4, and the definition of $x^\dagger$ imply that offering $x^\dagger$ is a best response to $\hat{\bar{p}}$.

Step 2: We treat the cases $x^\dagger = 1$ and $x^\dagger = \frac{1+p}{2} + c_F < 1$ separately.

(i) Suppose $x^\dagger = 1$, and consider the retention strategy given by:

$$\overline{x} = 1 \quad \overline{r} = c_F \quad \underline{r} = 0 \quad r_W = c_F + \gamma c_H \ .$$

Lemma 4 applied to this strategy yields:

$$\overline{p}^*(x) = \begin{cases} x - c_F & \text{if } x < 1 \\ 1 & \text{if } x = 1 \end{cases}$$

(ii) Suppose $x^\dagger = \frac{1+p}{2} + c_F < 1$, and consider the retention strategy given by:

$$\overline{x} = \frac{1+p}{2} + c_F \quad \overline{r} = \frac{1-p}{2} \quad \underline{r} = 0 \quad r_W = c_F + \gamma c_H \ .$$

Lemma 4 applied to this strategy yields:

$$\overline{p}^*(x) = \begin{cases} x - c_F & \text{if } x < 1 \\ 1 & \text{if } x = 1. \end{cases} \qquad\square$$

Proof of Proposition 3. Recall that $\mathcal{W}_F(\underline{p})$ is the expected payoff that Foreign gets from fighting all types at least $\underline{p}$.

Since (r_S, r_W) is a fully-extractive cutoff reward scheme, r_S has the form

$$r_S(x) = \begin{cases} \overline{r} & \text{if } x \geq x^\dagger \\ \underline{r} & \text{if } x < x^\dagger, \end{cases}$$

where $x^\dagger = 1 - \mathcal{W}_F(\underline{p})$. Moreover, there is an equilibrium of the continuation game in which $x^\dagger$ is offered by Foreign, and is accepted by all types of Home. For $x^\dagger$ to be a best response, it must be that any smaller offer would be rejected by a positive measure of types, else Foreign would have a profitable deviation to a smaller off that was also accepted almost surely. And Lemma 4 implies that the Home Leader's acceptance strategy in such a continuation equilibrium is

$$\tilde{p}(x) = \begin{cases} \iota_{\underline{p}}(x + \gamma c_H - r_W + \overline{r}) & \text{if } x \geq x^\dagger \\ \iota_{\underline{p}}(x + \gamma c_H - r_W + \underline{r}) & \text{if } x < x^\dagger \end{cases}.$$

A slight modification of the proof of Lemma 2 implies that Foreign's best response to $\tilde{p}$ is either the least off that is accepted for sure, $x^\dagger$, or is $x_\dagger = \underline{p} + c_F$. Thus we need to show that Foreign's payoff from offering $x_\dagger$ is not greater than $\mathcal{W}_F(\underline{p})$.

Suppose, seeking a contradiction, that types in $[\underline{p}, \underline{p} + \epsilon)$ accept the offer x. Then Foreign's payoff to offering x, denoted V, satisfies

$$\begin{aligned} \mathsf{V} &\geq \epsilon \frac{1 - \underline{p} - c_F}{1 - \underline{p}} + \int_{\underline{p}+\epsilon}^1 \frac{1 - t - c_F}{1 - \underline{p}} \, dt \\ &> \int_{\underline{p}}^{\underline{p}+\epsilon} \frac{1 - \underline{p} - c_F}{1 - \underline{p}} \, dt + \int_{\underline{p}+\epsilon}^1 \frac{1 - t - c_F}{1 - \underline{p}} \, dt \\ &= \mathcal{W}_F(\underline{p}). \end{aligned}$$

Thus it must be the case that no type accepts $x_\dagger$. In particular, the least type, $\underline{p}$, must reject it. This requires

$$\underline{p} - \gamma c_H + r_W \geq \underline{p} + c_F + \underline{r},$$

or

$$r_W - \underline{r} \geq \gamma c_H + c_F.$$

Since $\gamma c_H + c_F > 0$, we must have $r_W > \underline{r}$. $\qquad\square$

Proof of Proposition 4. The Home citizen's payoff is bounded above by a function that has two terms. The first is the payoff from types who keep the status quo: $\underline{p}y$. The second comes from types who initiate a crisis and (for the upper bound) fully extract: $(1 - \underline{p}) \min \left\langle 1, \frac{1+\underline{p}}{2} + c_F \right\rangle$. All together, the payoff at $\underline{p}$ is

$$\mathcal{F}(\underline{p}) = \underline{p}y + (1 - \underline{p}) \min \left\langle 1, \frac{1+\underline{p}}{2} + c_F \right\rangle.$$

We will proceed in two steps. First, we will find the values of $\underline{p}$ that maximize $\mathcal{F}$. Second, we will extend the construction from the proof of Proposition 2 to a complete retention strategy.

Step 1: Unpacking the minimization gives

$$\mathcal{F}(\underline{p}) = \begin{cases} \underline{p}y + (1 - \underline{p}) & \text{if } \underline{p} \geq 1 - 2c_F \\ \underline{p}y + (1 - \underline{p}) \left(\frac{1+\underline{p}}{2} + c_F \right) & \text{if } \underline{p} < 1 - 2c_F \end{cases}.$$

Consider the first case. We have $\mathcal{F}'(\underline{p}) = y - 1 < 0$, so the payoff is strictly decreasing. This is intuitive—once we are talking about entrants getting the entire pie, we want as many as possible to get it. Thus the optimal $\underline{p}$ lies in $[0, 1 - 2c_F]$.

Consider the second case. An optimal $\underline{p} \in [0, 1 - 2c_F)$ must satisfy the first-order condition:

$$\mathcal{F}'(\underline{p}) = (y - c_F) - \underline{p} \leq 0,$$

with equality if $\underline{p} > 0$.

As an immediate implication of this first-order condition and some calculation, we have:

Lemma 9. *Suppose $y \leq c_F$. Then $\operatorname{argmax}_{\underline{p}} \mathcal{F}(\underline{p}) \subset \{0, 1 - 2c_F\}$. $\underline{p} = 0$ is a solution if and only if $c_F \leq \frac{1}{2}$, while $1 - 2c_F$ is a solution if and only if $c_F \geq \frac{1}{2}$.*

Lemma 10. *Suppose $c_F < y < 1 - c_F$. Then $\operatorname{argmax}_{\underline{p}} \mathcal{F}(\underline{p}) \in \{y - c_F, 1 - 2c_F\}$.*

For calculations of which is optimal in the second Lemma, see the calculations in the file `sage_notes.txt`.

Step 2: From the previous step, we have to consider three cases: (i) $\underline{p} = 0$, (ii) $\underline{p} = 1 - 2c_F$, and (iii) $\underline{p} = y - c_F$ with $c_F < y < 1 - c_F$. In each, we must

34

extend the pair (r_S, r_W) to a complete retention strategy in such a way that the appropriate $\underline{p}$ is part of the equilibrium.

The first case, $\underline{p} = 0$, is easy—set $r_Q = 0$.

For the second and third cases, we will use:

Lemma 11. *Suppose the citizen uses a fully extractive cutoff reward scheme (r_S, r_W) with cutoff given by $x^\dagger$. This strategy can be extended to a complete retention strategy (r_Q, r_S, r_W) that induces an interior entry threshold $\underline{p}$ if and only if $x^\dagger - y + \bar{r} \leq 1$.*

Proof. Notice that type $\underline{p}$ is indifferent between the status quo and initiation followed by acceptance if and only if $y + r_Q = x^\dagger + \bar{r}$, which can be rearranged to give $r_Q = x^\dagger - y + \bar{r}$. This is a feasible retention probability only if $x^\dagger - y + \bar{r} \leq 1$. □

Now we consider the two cases in turn.

(ii) $\underline{p} = 1 - 2c_F$. From the proof of Proposition 2, $\bar{r} = c_F$. Thus the critical condition from Lemma 11 is $1 - y + c_F \leq 1$, or $y > c_F$.

(iii) $\underline{p} = y - c_F$ with $c_F < y < 1 - c_F$. From the proof of Proposition 2, $\bar{r} = \frac{1-\underline{p}}{2}$. Thus the critical condition from Lemma 11 is $\frac{1+\underline{p}}{2} + c_F - y + \frac{1-\underline{p}}{2} \leq 1$, or $y > c_F$.

So in each case, Proposition 2 and Lemma 11 imply that there is a retention strategy that is maximally extractive with the indicated properties. □

References

Austen-Smith, David and Jeffrey Banks. 1989. Electoral Accountability and Incumbency. In *Models of Strategic Choice in Politics*, ed. Peter C. Ordeshook. University of Michigan Press.

Barro, Robert J. 1973. "The Control of Politicians: An Economic Model." *Public Choice* 14:19–42.

Bueno de Mesquita, Bruce and David Lalman. 1992. *War and Reason: Domestic and International Imperative.* New Haven, CT: Yale University Press.

Bueno de Mesquita, Bruce, James D. Morrow, Randolph M. Siverson and Alastair Smith. 1992. *War and Reason: Domestic and International Imperative.* New Haven, CT: Yale University Press.

Bueno de Mesquita, Bruce and Randolph M. Siverson. 1995. "War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability." *American Political Science Review* 89(4):841–855.

Bueno de Mesquita, Bruce, Randolph M. Siverson and Gary Woller. 1992. "War and the Fate of Regimes: A Comparative Analysis." *American Political Science Review* 86(3):638–646.

Busch, Marc L. 2000. "Democracy, Consultation, and the Paneling of Disputes under GATT." *The Journal of Conflict Resolution* 44(4):425–446.

Chiozza, Giacomo and H.E. Goemans. 2004. "International Conflict and the Tenure of Leaders: Is War Still *Ex Post* Inefficient?" *American Journal of Political Science* 48(3):504–619.

Debs, Alexandre and Hein E Goemans. 2010. "Regime type, the fate of leaders, and war." *American Political Science Review* 104(03):430–445.

Downs, George W. and David M. Rocke. 1994. "Conflict, Agency, and Gambling for Resurrection: The Principle-Agent Problem Goes to War." *American Journal of Political Science* 38(2):362–380.

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

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

Fearon, James D. 1999. Electoral Accountability and the Control of Politicians: Selecting Good Types versus Sanctioning Poor Performance. In *Democracy, Accountability, and Representation*, ed. Przeworski, Stokes and Manin. Cambridge University Press.

Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50(1/3):5–25.

Fey, Mark and Kristopher W. Ramsay. 2009. "Uncertainty and Incentives in Crisis Bargaining Games." *University of Rochester, Typescript* .

Goemans, Hein E. 2000. *War and Punishment: The Causes of War termination and the First World War*. Princeton, NJ: Princeton University Press.

Gowa, Joanne. 1999. *Ballots and bullets: The elusive democratic peace*. Princeton, NJ: Princeton University Press.

Hess, Gregory D. and Athanasios Orphanides. 2001. "War and Democracy." *The Journal of Political Economy* 109(4):776–810.

Jackson, Matthew O. and Massimo Morelli. 2007. "Political Bias and War." *American Economic Review* 97(4):1353–1373.

Kant, Immanuel. 1903. *The Perpetual Peace; A Philosophical Essay*. London, UK: Swan Sonnenchein & Co LIM.

Pevehouse, Jon C. 2002. "With a Little Help from My Friends? Regional Organizations and the Consolidation of Democracy." *American Journal of Political Science* 46(3):611–626.

Russet, Bruce. 1993. *Grasping the Democratic Peace: Principles for a Post-Cold War World*. Princeton, NJ: Princeton University Press.

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

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

Schultz, Kenneth. 2001*b*. "Looking for Audience Costs." *Journal of Conflict Resolution* 45(1):32–60.

Slantchev, Branislav L. 2006. "Politicians, the Media, and Domestic Audience Costs." *International Studies Quarterly* 50(2):445–477.

Smith, Alastair. 1998. "International Crises and Domestic Politics." *American Political Science Review* 92(3):623–638.

Snyder, Jack. 1991. *Myths of Empire: Domestic Politics and International Ambition.* Ithica, NY: Cornell University Press.

Snyder, Jack and Erica D Borghard. 2011. "The cost of empty threats: A penny, not a pound." *American Political Science Review* 105(03):437–456.

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