Bargaining under Almost Complete Information

Deepal Basak*

Sept 2022

ABSTRACT

A compromise against public sentiment can hurt a leader in the next primary. We consider the canonical Rubinstein bargaining with frequent offers and binary states (sentiment leans toward one leader or the other). Under complete information, the leaders reach an agreement depending on the sentiment and do so immediately. We perturb this game by introducing a small positive probability ε that the leaders are uninformed about the sentiment. We show that a unique equilibrium emerges that resembles a war of attrition. We study whether, under small ε, in every state, the leaders almost immediately agree to the same policy position as they do under complete information. The answer is yes if we fix the bargaining environment and consider a sequence of ε converging to zero. However, given any ε (however small), we can find bargaining environments in which this is not true.

JEL Classification Numbers: *C78, D82*

Key Words: *Bargaining, Higher-order uncertainty, Public sentiment, Reputation*

Introduction

Legislating in a democratic system often requires policy compromises, and if such a compromise is against the public sentiment, it can be used against a leader in the next primary.

> "Even if a particular deal is the best that can win sufficient support in Congress to pass—and would be an improvement, in their view, over the policy status quo—lawmakers still may conclude that they would be unable to defend it successfully with their constituencies. Lawmakers may well reject 'half a loaf' and settle

*I am grateful to Archishman Chakraborty, Joyee Deb, Jack Fanning, Rick Harbaugh, Urmee Khan, Aaron Kolb, David Pearce, Marilyn Pease, Ennio Stacchetti, and the seminar participants at Stanford Political Economic Theory (SITE), Stony Brook Game Theory meeting, Econometric Society meetings, and SAET for helpful comments and suggestions. I thank Samyak Jain for his excellent research assistance. Basak: Indiana University, Kelley School of Business, Email: dbasak@iu.edu

for nothing, if taking the half would be understood by constituents or denounced by important groups or activists as an unacceptable sellout. Pundits today call this a fear of being 'primaried,' although the electoral imperative to satisfy activist constituencies has deep roots in congressional politics."
— "Negotiating Agreement in Politics," American Political Science Association, Task Force Report, 2013 (p 57)

Consider, for instance, two leaders—left (L) and right (R)—bargaining over a trade deal in which L wants more government interventions or restrictions and R wants fewer. Nature selects a state: the public sentiment, which can be pro or anti free trade. Suppose it is commonly known that the sentiment is anti free trade. In this case, we expect that the leaders will agree on a policy position that imposes significant restrictions on the trade deal. This is because L can credibly commit to not making a large compromise since doing so will hurt her in the next primary. Two rational leaders should reach such an agreement immediately. However, if there is a positive probability ε_j that leader j may not know the underlying public sentiment on this issue, leader $i \neq j$ can exaggerate the expected reaction of her constituents to her compromise. We study whether under small $\varepsilon = (\varepsilon_i, \varepsilon_j)$, the equilibrium outcome in every state is close to that under complete information.

Formally, we consider the canonical Rubinstein (1982) bargaining game with frequent offers. Leaders L and R bargain over a policy position $p \in [0, 1]$. We use the convention that L wants higher p and R wants lower p. Suppose that the leaders are not concerned about the public sentiment at all. Assuming stationary discount rates $r = (r_i, r_j)$, it follows from the standard Rubinstein argument that the leaders will immediately compromise and agree to a policy position in the middle $x^* = r_R/(r_L + r_R)$.

However, when the leaders care about the public sentiment, as the opening quote mentions, a leader may prefer the status quo over agreeing to "half a loaf." To go back to our example, consider, for instance, leader L. Whether compromising and agreeing to x^* is acceptable to L depends on whether the public sentiment is pro or anti free trade. If the public sentiment is pro free trade (leans right), L has no fear that her compromises can be used as an issue to primary her. However, if the public sentiment is anti free trade (leans left), such a compromise can be used to portray her as a sellout. We assume that if the public sentiment leans towards a leader and she makes a sufficiently large compromise (beyond the *tolerance threshold* of her constituents), then she gets 0. We refer to the tolerance threshold for a leader i as $1 - x_i$ (a higher x_i means less tolerance) and assume that anything worse than x^* is not tolerated.

Under complete information, this political bargaining game is equivalent to Binmore et al. (1989). Suppose it is commonly known that the public sentiment leans left on this issue. In that case, L can credibly commit to never agreeing to a policy position $p < x_L$. In this sense, x_L works like an outside option (although not a physical one). Since anything worse than x^* is not tolerated, $x_L \geq x^*$. It then directly follows from Binmore et al. (1989) that under frequent offers, in equilibrium, the leaders immediately agree to x_L, which is just generous enough for L that L takes it rather than saying *deal me out*. We refer to this event as an immediate concession from R. Analogously, when the public sentiment leans right, leader L immediately concedes, and the leaders agree to the policy position $1 - x_R$.

We perturb this game in a simple and easily interpretable way. We introduce a small positive probability ε that the leaders are uninformed about the underlying sentiment. If x^* is acceptable regardless of the public sentiment, then sentiment plays no role and the leaders will always immediately agree at x^* regardless of ε. We assume, however, that this is not the case. We study whether the complete information bargaining result is robust to such perturbation or whether a small ε can make a large difference in the equilibrium outcome.[1]

Formally, the incomplete information bargaining game is as follows. Nature selects a state (left- or right-leaning public sentiment) from a commonly known prior. With probability ε_i, leader i does not learn the state, and with complementary probability, she privately learns the state. We assume that the information structure is commonly known. For convenience, we use θ to represent the primitives of the bargaining environment, which incorporates (1) the impatience, (2) the tolerance, and (3) the prior belief regarding the public sentiment, and $\varepsilon = (\varepsilon_L, \varepsilon_R)$ captures the information environment.

In our perturbed game, a leader i could be an informed type who knows that the public sentiment leans toward her ($\omega_i = I$), an informed type who knows that the public sentiment leans the other way ($\omega_i = I'$), or an uninformed type ($\omega_i = \mathcal{U}$). Notice that the different types have different beliefs about her opponent's types. For instance, an uninformed type believes that if her opponent learns the public sentiment, then her opponent is either the I type or the I' type. In contrast, the I' type believes that if her opponent learns the public sentiment, then her opponent must be the I type and not the I' type. As the bargaining continues, these different types will update their beliefs about their opponent's types, and the uniformed type will also update her belief about the underlying public sentiment. Thus, this simple information perturbation induces higher-order uncertainty and creates an equilibrium with rich bargaining dynamics.

Although we are mostly interested in small ε, in Proposition 1, we describe the equilibrium for any (θ, ε). First, in the spirit of reputational bargaining, we construct the unique equilibrium that resembles a war of attrition. We assume that the I type is *committed* to the same policy position to which she agrees when the public sentiment is commonly known; that is, she always demands the same policy, always accepts a better policy, and always rejects a worse policy. Assuming

[1] It is often difficult to discern the public sentiment for some issues, and accordingly, difficult to predict how a political compromise will play out in a future election. However, in this age of big data, for many issues, the probability that a leader is uninformed about the underlying sentiment is small.

such commitment type Abreu and Gul (2000) (hereafter, AG) shows that under frequent offers, in equilibrium, the non-committed types randomize over when to stop exaggerating and concede to their opponent's demand. This results in a war of attrition. We show later (see Section 3.1) that we do not need to assume the commitment type. Under frequent offers, the commitment of the I type arises endogenously.[2]

Four properties uniquely identify the war of attrition equilibrium. First, it must be that the I' type of a leader concedes before the $\mathcal{U}$ type because, compared to the $\mathcal{U}$ type, the I' type is more reluctant to exaggerate and less optimistic about the opponent conceding. We refer to this property as the two-sided skimming property.

Second, the uninformed types must become convinced about the opponent's commitment simultaneously because once an uninformed leader understands that the opponent will never concede, she stops exaggerating and concedes.

Third, if a non-committed leader believes that her opponent may concede immediately with positive probability, she will exaggerate rather than concede immediately. The second and third properties are standard in reputational bargaining.

Finally, in equilibrium, it must be that at least one of the leaders concedes immediately with probability 1 when she learns that the sentiment leans the other way because the I' type believes that the opponent is either the I type or the $\mathcal{U}$ type. Therefore, if she exaggerates, she must believe that the $\mathcal{U}$ type opponent may concede in the meantime. It then follows from the skimming property of the equilibrium that the I' opponent must have already conceded.

Suppose, as in Property 4, it is leader j who never exaggerates; that is, leader j concedes immediately with probability 1. Property 2 implies that the leaders may continue bargaining until some date T, at which point both uninformed types will become convinced that their opponent is committed. Property 1 implies that bargaining will include two phases: in phase I, $\omega_i = I'$ will mix; in phase II, $\omega_i = \mathcal{U}$ will mix (see Figure 1). Finally, Property 3 implies that both $\omega_i = I'$ and $\omega_j = \mathcal{U}$ cannot concede immediately with positive probability.

This two-phase war of attrition resembles the equilibrium in Abreu et al. (2015) (henceforth, APS). The authors extend the AG framework and introduce two non-committed types on one side who differ in their discount rates. In phase I, the impatient type mixes; in phase II, the patient type mixes. Notice that we have two non-committed types on both sides. Although they are equally patient, they differ in their beliefs, which leads to the skimming property and the two phases.

Thus, when we perturb the complete information bargaining game by introducing a positive probability ε that the leaders are uninformed, in equilibrium, one

[2] The I type becomes committed because she cannot get a policy better than what the leaders agree to when the sentiment is commonly known. On the other hand, by definition, she never accepts a worse policy. This result follows from the feature that the sentiment influences the payoff, provided that a leader makes a sufficiently large compromise against the sentiment.

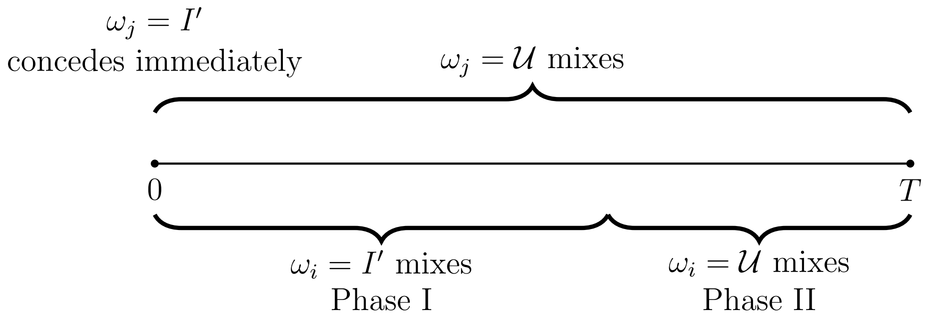

Figure 1: War of attrition with two phases

of the leaders—even after learning that the sentiment leans the other way (the I' type)—continues exaggerating in phase I. We call her the "strong bargainer" and the other the "weak bargainer." The convergence to the complete information result depends on how long the I' strong bargainer continues exaggerating (phase I).

If the I' strong bargainer also concedes immediately with probability 1, then we say that the bargaining strengths are *balanced.* In this case, neither leader exaggerates upon learning that the sentiment leans the other way, and accordingly, there is no phase I. If the I' strong bargainer concedes immediately with probability in $(0, 1)$, then we say that the bargaining strengths are *moderately unbalanced*, and if she never concedes immediately, we say that the bargaining strengths are *extremely unbalanced.*

We show that the degree to which the bargaining strengths are unbalanced is identified by two measures of *relative bargaining strengths.* As in AG, we define the relative bargaining strength of a leader j as $\exp(T_i)/\exp(T_j)$ where T_i is the time leader i takes to convince her uninformed opponent about her commitment. The first measure of relative bargaining strength $B_j^1(\theta, \varepsilon)$ assumes that both I' type leaders concede immediately with probability 1. The second measure $B_j^2(\theta, \varepsilon)$ assumes that only the I' type leader j concedes immediately with probability 1, whereas the I' type leader i never concedes immediately.

If $B_j^1 = 1$, the strengths are balanced. If $B_j^2 > 1 > B_j^1$, then i is strong enough that $\omega_i = I'$ exaggerates, but not so strong that she never concedes immediately; that is, i is moderately strong. If $1 \geq B_j^2$, i is extremely strong. When i is extremely strong, $\omega_i = I'$ never concedes immediately, whereas $\omega_j = \mathcal{U}$ may concede immediately with positive probability.

I show that when we fix the bargaining environment θ and take a sequence of $\varepsilon \to (0, 0)$, the bargaining strengths become almost balanced. Accordingly, the leaders do not exaggerate when they learn that the public sentiment leans the other way. In other words, in every state, they almost immediately agree to the same policy position as they do when the public sentiment is commonly known.

However, this does not mean, given a small ε, that the above probability is close to 1 regardless of the bargaining environment θ. Given ε (however small), we can find bargaining environments (in particular, if one side is sufficiently intolerant) in which the bargaining strengths are far from balanced. Accordingly, one of the leaders will never concede immediately even when she learns that the sentiment leans the other way, and she may keep bargaining for a positive duration (phase I). Thus, in one of the states, it is impossible (probability 0) to reach an immediate agreement on the same policy position as the leaders do when the state is commonly known.

This paper studies (1) a bilateral bargaining game where nature selects which way the public sentiment leans and introduces a small positive probability ε that the leaders are uninformed. (2) This leads to a war of attrition equilibrium in the spirit of AG. (3) However, there is higher-order uncertainty, and the paper shows how to think of the bargaining strengths under such higher-order uncertainty. (4) Finally, it shows that a small ε can make a large difference. Accordingly, this paper is related to four strands of bargaining literature.

Audience cost in political bargaining: Public sentiment influences the audience cost of compromises and accordingly shapes the equilibrium policy in political bargaining. When the public sentiment is commonly known, Muthoo (1992), Levenotoğlu and Tarar (2005) show that a higher audience cost helps a leader because it can be used as credible commitment. The idea that commitment helps in bargaining dates back to at least Schelling (1956). More recently, Basak and Deb (2020) study a political bargaining game in which the public sentiment is unknown and is publicly revealed at a later date. The authors show that in equilibrium, rather than agreeing right away, the leaders wait and see which way the public opinion moves and then agree accordingly. Notice that there is no private information, and hence, a leader cannot exaggerate about her audience cost. In contrast, in this paper, information is private, which induces higher-order uncertainty and creates an equilibrium with much richer equilibrium dynamics in which the leaders exaggerate about their constituents' reaction. We show that even with a small probability of the leaders being uninformed, substantial delays can arise in some bargaining environments.

War of attrition in bargaining: There is a large literature on the political war of attrition. For instance, Alesina and Drazen (1991) considers two groups who fight over a tax burden and cause a delay in adopting a stabilization policy to reduce the large budget deficit.[3] Fearon (1994) considers an international conflict and domestic audience costs. The authors and the literature that follows directly impose a war of attrition structure, whereas we consider a canonical bargaining game, with the war of attrition arising as a unique equilibrium.

As already mentioned, our war of attrition equilibrium is similar to the reputa-

[3] Egorov and Harstad (2017) considers a war of attrition with multiple stages where a firm chooses whether to self-regulate, an activist decides whether to continue boycott and a regulator decides whether to intervene with public regulation.

tional bargaining literature. AG pioneered this literature. For a recent survey of reputational bargaining, see Fanning and Wolitzky (2022). This literature provides an intuitive notion of bargaining strength that determines which agents exaggerate and for how long. This bargaining strength is measured by how quickly an agent can convince her opponent about her commitment. In our setup, since the leaders learn about the sentiment privately, there is higher-order uncertainty. This paper builds the appropriate notion of bargaining strength under such higher-order uncertainty.[4,5]

The war of attrition equilibrium under two-sided asymmetric information was first constructed in Chatterjee and Samuelson (1987, 1988). However, the authors show that other equilibria can be constructed using belief-based threats. AG resolves this multiplicity by assuming commitment types who are immune to belief-based threats. This assumption makes the war of attrition equilibrium unique. In our setup, under frequent offers, this commitment arises endogenously. This follows from the feature that the sentiment influences the payoff, provided a leader makes a sufficiently large compromise against the sentiment. If the public sentiment also affects the payoff for a small compromise, then the I types may become susceptible to belief-based threats. This result can make the convergence to the complete information outcome even weaker.

Higher-order uncertainty in bargaining: Feinberg and Skrzypacz (2005) study a dynamic pricing game in which the buyer knows her private values. The authors show that if there is a positive probability that the seller may know the value as well, contrary to the Coase conjecture, delay must occur.[6] Fanning (2021) considers a reputational bargaining model in which the leaders may privately tell a mediator that they are not committed to their demands and are willing to compromise. The mediator discloses this information when both sides reveal their willingness to compromise. Although the setup is very different, the equilibrium has a property similar to our Property 4. The author shows that at least one leader has to give up immediately after telling the mediator that she is willing to compromise, but the mediator stays silent.

Robustness in bargaining: Weinstein and Yildiz (2013) establish a general result that for any bargaining outcome, one can introduce a small amount of incomplete information in such a way that the resulting type profile has a unique rationalizable action profile that leads to this bargaining outcome. However, the authors mention that *"the types constructed in our article are complicated, and it is not easy to interpret how they are related to economic parameters"* (p 380). In contrast, we introduce an easily interpretable perturbation: a small probability

[4]To understand when a madman strategy is profitable in an international conflict, Acharya and Grillo (2015) introduce commitment types following AG.

[5]Ellingsen and Miettinen (2014) shows that a war of attrition can arise even without any asymmetric information if the agents can commit and such commitment decays over time.

[6]See Tsoy (2018) and Madarász (2021) for more recent developments on dynamic pricing under higher-order uncertainty.

that the leaders are uninformed about the public sentiment. It is important to note that this information perturbation is different from the reputational perturbation. Reputational bargaining studies the limiting environment in which the probability that a player is irrationally committed is close to 0. We study an environment in which one of the players has reasons to be rationally committed, and the probability that a leader does not know whether these reasons are valid is close to 0. Unlike in AG, agreement may not be almost immediate even under small ε.[7]

The rest of the paper is organized as follows. Section 1 describes the model. Section 2 shows the main results. We construct the war of attrition equilibrium assuming the I type is committed (to the same policy position that she agrees to when the sentiment is commonly known). We then study the robustness of the complete information result. In Section 3, we discuss why, under frequent offers, commitment arises endogenously. We also discuss the connections to some existing results. Section 4 concludes. The proofs are relegated to the appendix.

1. Model

The bargaining game: There are two political leaders $N = \{L, R\}$ and a status quo policy that gives 0 to both leaders. An opportunity has arrived to bring about a change, but the two leaders need to agree. We assume that the policy space is $p \in [0, 1]$. While leader L wants a higher p, leader R wants a lower p. An agreement at policy position p is perceived as a compromise from L of size $1-p$ and a compromise from R of size p. The bargaining proceeds as an alternating offer bargaining game 'a la Rubinstein (1982). One of the leaders is picked randomly with probability $\frac{1}{2}$ to make the first policy proposal $p \in [0, 1]$. If the other leader accepts this proposal, an agreement is reached, and the game ends; otherwise, it continues to the next round. The other leader proposes the next round, and the same process continues until an agreement is reached. Thus, the game can potentially go on forever. We assume that offers can be made frequently; that is, the time interval between two rounds $\Delta \to 0$, and a leader $i \in N$ discounts time at a stationary rate r_i. Therefore, the discount factor $e^{-r_i\Delta} \to 1$. Throughout this paper, we only consider bargaining under the continuous-time limit ($\Delta \to 0$).

Public Sentiment and Leaders' Payoff: Before the game begins, nature draws a state $\omega \in \{\mathscr{L}, \mathscr{R}\}$. Let $u_i(p, t, \omega) = e^{-r_i t} u_i(p, \omega)$ be the payoff of leader i if she agrees to policy position p after bargaining for time t, while the state is ω,

[7] Fanning (2018) considers a reputational bargaining game with cost uncertainty that is revealed at a later date. The author shows that delay can arise even when the probability of being committed is close to zero.

where $u_i(p, \omega)$ is as follows:

$$\begin{aligned} u_L(p, \mathscr{L}) &= p \cdot \mathbb{1}(p \geq x_L) \quad \text{and} \quad u_R(p, \mathscr{L}) = (1-p). \\ u_L(p, \mathscr{R}) &= p \qquad\qquad\qquad \text{and} \quad u_R(p, \mathscr{R}) = (1-p) \cdot \mathbb{1}(p \leq 1 - x_R). \end{aligned}$$

We say that in state $\mathscr{L}$ ($\mathscr{R}$), the public sentiment leans left (right). Notice that leader L prefers higher p and leader R prefers lower p. The only difference from standard Rubinstein bargaining is that leader L gets 0 from $p < x_L$ when the state is $\mathscr{L}$, and leader R gets 0 from $p > 1 - x_R$ when the state is $\mathscr{R}$.

This payoff specification captures the feature that (1) if a leader makes a sufficiently large compromise and (2) the public sentiment leans toward her, then it can be used to portray her as a sellout in the next primary. She prefers the status quo over agreeing to such compromise. To go back to our example, if, say, leader L agrees to impose inadequate restrictions on the trade deal, and the sentiment is anti free trade (leans left), then it can be used to primary her. To see this, recall that if L agrees to a policy position p, it is considered as a compromise of size $1 - p$. If $p < x_L$, then the compromise is large enough that it can be used as an issue to turn the constituents against her when the sentiment leans left. Higher x_L means the constituents are less tolerant of compromises (their opinions are easier to turn).

Assumption 1: *When the public sentiment leans toward a leader i, her constituents are sufficiently intolerant: $x_L \geq x^*$, $x_R \geq 1 - x^*$, where $x^* = \frac{r_R}{r_L + r_R}$ is the Rubinstein equilibrium policy position. At least one of the inequalities is strict.*

If the leaders have no concern for public sentiment, then it follows from the standard Rubinstein argument that in equilibrium, the leaders will immediately agree to policy position x^*, which is solely determined by the relative discount rates. If L is relatively more patient, x^* is higher, and when they are equally patient, $x^* = 1/2$. The above assumption says that compromising and agreeing to a policy that is strictly worse than x^* is not acceptable to a leader i when the sentiment leans toward her (the fear of being primaried). Moreover, x^* is not acceptable in at least one of the states.

Complete Information Benchmark Under complete information, our political bargaining game is equivalent to Binmore et al. (1989). The following result directly follows from their work. Under the continuous-time limit ($\Delta \to 0$), given assumption 1, when the public sentiment leans left ($\omega = \mathscr{L}$), the leaders immediately agree to policy policy position x_L, and when the public sentiment leans right ($\omega = \mathscr{R}$), the leaders immediately agree to policy policy position $1 - x_R$.

This result is commonly referred to as *deal me out*. The authors consider a complete information bargaining game in which x_i is an outside option for i, and hence, i never accepts a share lower than x_i. In our setup, x_i is not a physical outside option. Nevertheless, the same is true because of the fear of being primaried. Suppose the sentiment leans left. As in Binmore et al. (1989), L can credibly commit that she will never agree to a policy $p < x_L$. The authors show that the best L can do is to say *deal me out* unless $p \geq x_L$. If $x_L < x^*$, then the standard Rubinstein equilibrium policy is acceptable, and hence, the deal me out result has no impact. However, since $x_L \geq x^*$ (assumption 1), the equilibrium policy is $p = x_L$. We call this event an immediate concession from R. Analogously, if it is commonly known that the public sentiment leans right, then the leaders immediately agree on $p = 1 - x_R$. We call this event an immediate concession from L. Since this is a well-known result, I omit the formal proof.

It is important to note that under the deal me out result, when the sentiment leans toward a leader, she gets the worst policy that she considers acceptable. This result follows from the feature that the sentiment can be used against a leader i provided she makes a sufficiently large compromise (larger than $1 - x_i$). For small compromises, the issue is not salient enough that it can be used against her. This property plays an important role in our commitment argument. In Section 3, we discuss the case in which sentiment can affect the payoff even when the compromise is small.

Incomplete Information We perturb the complete information political bargaining game by introducing a small positive probability that the leaders are uninformed about the underlying public sentiment. Nature draws a state $\omega \in \Omega = \{\mathscr{L}, \mathscr{R}\}$ from a commonly known prior: $\pi_L = P(\omega = \mathscr{L})$, $\pi_R = P(\omega = \mathscr{R})$, and $\pi_L + \pi_R = 1$. Each leader $i \in N$ receives a private signal $\omega_i \in \Omega_i = \Omega \cup \{\mathcal{U}\}$, where $\mathcal{U}$ represents an uninformative signal realization. The signal structure of leader $i \in N = \{L, R\}$ is denoted by the conditional probability distributions $q_i : \Omega \to \mathbb{P}(\Omega_i)$, where

$q_i(\omega_i \mid \omega)$	$\omega_i = \mathscr{L}$	$\omega_i = \mathcal{U}$	$\omega_i = \mathscr{R}$
$\omega = \mathscr{L}$	$1 - \varepsilon_i$	ε_i	0
$\omega = \mathscr{R}$	0	ε_i	$1 - \varepsilon_i$.

We assume that this information structure is commonly known. The information structure can be represented by a vector $\varepsilon = (\varepsilon_L, \varepsilon_R)$, where ε_i captures the probability that leader i does not know the public sentiment. When ε is close to $(0, 0)$, we say that there is almost complete information.

Given this information structure, different types have different beliefs about the public sentiment and the opponent's type. A leader i who receives signal $\omega_i \in \{\mathscr{L}, \mathscr{R}\}$ knows the public sentiment, and when she receives signal $\omega_i = \mathcal{U}$, she believes that with probability π_i the sentiment leans her way and with probability π_j the sentiment leans the other way. For convenience, I relabel the types as I (learns that the public sentiment leans her way), I' (learns that the public sentiment leans the other way), and $\mathcal{U}$ (is uninformed). At the beginning of the game, a leader i, regardless of her type, assigns probability ε_j that her opponent is uninformed. Since the state is either $\mathscr{L}$ or $\mathscr{R}$, the I type does not believe that her opponent can be the I type, and the I' type does not believe that her opponent can be the I' type. However, an $\mathcal{U}$ type believes that her opponent could the I type or the I' type. At the initial node $\emptyset$, let $\alpha_j^{\omega_i}(\emptyset)(\omega_j)$ be the belief of the ω_i type of leader i that her opponent j is the ω_j type, where $i, j \in N$, and $i \neq j$. Then,

$\alpha_j^{\omega_i}(\emptyset)(\omega_j)$	$\omega_j = I$	$\omega_j = I'$	$\omega_j = \mathcal{U}$
$\omega_i = I$	0	$1 - \varepsilon_j$	ε_j
$\omega_i = I'$	$1 - \varepsilon_j$	0	ε_j
$\omega_i = \mathcal{U}$	$\pi_j(1 - \varepsilon_j)$	$\pi_i(1 - \varepsilon_j)$	ε_j.

Strategies, Beliefs, and Solution Concept: For clarity of exposition and to avoid the burden of unnecessary notation, in Section 2, we first construct a war of attrition equilibrium in the spirit of AG. We assume that a leader who knows that the public sentiment leans her way (the I type) is "committed" to demanding the policy position to which she agrees when the public sentiment is commonly known. She always demands the same policy, accepting a better policy with probability 1 and a worse policy with probability 0. Assuming such commitment type, AG shows that under the continuous-time limit ($\Delta \to 0$), in equilibrium, a non-committed type's strategy boils down to the matter of when to concede.

Following AG, we define the strategy of a non-committed type as $F_i^{\omega_i}(t)$, which captures the probability that a non-committed type ω_i of leader i concedes by time t (inclusive). Given $F_j^{\omega_j}(t)$, let $G_j^{\omega_i}(t)$ be the belief of ω_i that her opponent $j \neq i$ will concede by time t. Recall that both $\omega_i = I'$ and $\omega_i = \mathcal{U}$ believe that their opponent is uninformed with probability ε_j. However, while $\omega_i = I'$ believes that the opponent cannot be the I' type, $\omega_i = \mathcal{U}$ assigns probability $\pi_i(1 - \varepsilon_j)$ that the opponent is the I' type. Therefore,

$$G_j^{I'}(t) = \varepsilon_j F_j^{\mathcal{U}}(t) \tag{1}$$

$$G_j^{\mathcal{U}}(t) = \varepsilon_j F_j^{\mathcal{U}}(t) + \pi_i(1 - \varepsilon_j) F_j^{I'}(t). \tag{2}$$

We can see from (1) and (2) that $G_j^{\mathcal{U}}(t) \geq G_j^{I'}(t)$. This means that, conditional on no agreement until time t, the uninformed type is at least as optimistic as the I' type regarding the probability of the opponent conceding in the next instance. The uninformed type is strictly more optimistic than the I' type if either $F_j^{I'}(t) > 0$ or the I' opponent may concede in the next instance.

The uninformed type $\omega_i = \mathcal{U}$ does not know the public sentiment and updates her belief over time that the public sentiment leans the other way with probability

$$\pi_j^{\mathcal{U}}(t) = \frac{\pi_j\left[(1 - \varepsilon_j) + \varepsilon_j(1 - F_j^{\mathcal{U}}(t))\right]}{1 - G_j^{\mathcal{U}}(t)}. \tag{3}$$

The denominator is the probability that the opponent does not concede by time t, and the numerator is the initial probability that the public sentiment leans the other way times the probability that the opponent does not concede by time t, given that the public sentiment leans the other way. Recall that the I and I' types know the public sentiment; that is, they have degenerate beliefs about the state. At any t, $\pi_j^I(t) = 0$ and $\pi_j^{I'}(t) = 1$.

Given the strategy F_j of leader j (committed strategy of $\omega_j = I$ and $F_j^{\omega_j}(t)$ of the non-committed type $\omega_j \in \{I', \mathcal{U}\}$), the expected payoff of a leader i of type ω_i from conceding at time t is

$$U_i^{\omega_i}(t, F_j) := \int_{\tau<t} e^{-r_i\tau} x_i dG_j^{\omega_i}(t) + (1 - G_j^{\omega_i}(t))e^{-r_it}\pi_j^{\omega_i}(t)(1 - x_j). \tag{4}$$

If the opponent concedes in the meantime, at some $\tau < t$, a leader i gets $e^{-r_i\tau}x_i$ (L gets p_L and R gets $1 - (1 - x_R) = x_R$). If the opponent has not conceded in the meantime, she then concedes. Such concession gives her $e^{-r_it}(1 - x_j)$ if the sentiment leans toward her opponent, and 0 otherwise.

Notice that different types assign different probabilities $G_j^{\omega_i}(t)$ to the opponent conceding by time t. Even if these beliefs were the same, the expected payoff in equation (4) would be different across types since their beliefs about the public sentiment $\pi_j^{\omega_i}(t)$ are different. Notice the second part of the expected payoff. Since $\pi_j^I(t) = 0$, $\omega_i = I$ gets zero from such concession. She knows that the sentiment leans in her direction, and such compromise can be used to portray her as a sellout. She is better off sticking with the status quo than making such compromise. On the other hand, since $\pi_j^{I'}(t) = 1$, $\omega_i = I'$ knows that such compromises cannot be used against her. Therefore, she is willing to make such concession. The uninformed type ($\mathcal{U}$) does not know whether or not such compromise can be used against her. She is willing to concede but is more reluctant to concede than the I' type. To see this, note that the expected payoff of the uninformed type is $e^{-r_it}\pi_j^{\mathcal{U}}(t)(1 - x_j)$. Therefore, if she keeps bargaining, she discounts a smaller share than the I' type. Since only the uninformed type can have a degenerate belief about the public sentiment, to simplify notation, we suppress the superscript and simply denote the updated belief of the uninformed type such that the sentiment leans toward her opponent j as $\pi_j(t)$.

We say that (F_i, F_j) constitutes a war of attrition equilibrium if each type or each leader maximizes her expected payoff in (4) given the opponent's strategy.

In AG, there is only one non-committed type on either side. APS builds on this result and introduces two non-committed types on one side. We will build on both of these results and construct the war of attrition equilibrium with two non-committed types on both sides. It is important to note that the two non-committed types have different beliefs about the underlying sentiment and their opponent's type. They will update their beliefs about their opponent differently over time . Moreover, since $\pi_j(t)$ varies over time, unlike in AG, the indifference condition of the uninformed type is not stationary.

In Section 3, we relax the commitment assumption. We consider perfect Bayesian equilibrium (PBE) as our solution concept. A PBE requires that the leaders are sequentially rational (they maximize their expected payoff at any history), and the beliefs are Bayesian consistent on path. In addition, we assume that a leader never proposes a policy that, in any PBE, will be rejected with probability 1 by all types of her opponent at all histories of the game. Given this assumption, under the continuous-time limit ($\Delta \to 0$), we show (see Lemma 2) that in any PBE, the I type chooses to behave like a commitment type in AG. In this sense, it is without loss of generality to look into the above war of attrition equilibrium in our political bargaining game.

Throughout this paper, we maintain the assumption that the time gap between offers $\Delta \to 0$. As is standard in reputational bargaining, we simply refer to this as continuous-time bargaining. All our results are for continuous-time bargaining. It is worth mentioning that AG shows that the equilibrium outcome under discrete-time bargaining converges in distribution to the equilibrium outcome under the continuous-time war of attrition, and the same argument applies here (see Section 3.3).

2. Main Result

2.1. War of Attrition

In this section, building on AG, we construct a unique war of attrition equilibrium (F_i, F_j). A leader who learns that the public sentiment leans her way (the I type) understands that a concession can be used to portray her as a sellout. We assume that she is committed to the same policy position to which she agrees when the sentiment is commonly known and she never concedes. Formally, she always demands the same policy, always accepts a weakly better policy, and never accepts a worse policy. A leader who learns that the public sentiment leans the other way (the I' type) understands that a concession cannot be used against her. An uninformed leader (the $\mathcal{U}$ type), on the other hand, is uncertain regarding whether a concession can or cannot be used against her. In equilibrium, the non-committed types (I' and $\mathcal{U}$) choose $F_i^{\omega_i}(t)$—how long to keep masquerading as the I type—that is, they exaggerate regarding how much their constituents care about their concession.

If $F_i^{\omega_i}(0) = 1$, we say that ω_i always concedes right way; that is, she never exaggerates. If $F_i^{\omega_i}(0) = 0$, we say that ω_i never concedes right away. We define $T_i[\omega_i]$ as the time until which leader i of type ω_i may continue exaggerating and $T_i := \max\{T_i[I'], T_i[\mathcal{U}]\}$ as the time until which leader i may continue exaggerating. In other words, after time $T_i[\omega_i]$, the uninformed leader j is convinced that her opponent is not the type ω_i, and after time T_i, she is convinced that her opponent is committed.

Lemma 1: *In the continuous-time war of attrition with multiple non-committed types on both sides, the following properties must hold true:*

(P1) *For any $i \in N$, $\omega_i = I'$ concedes before $\omega_i = \mathcal{U}$ concedes.*

(P2) *The uninformed types become convinced about their opponent's commitment simultaneously; that is, $T_L = T_R =: T$.*

(P3) *If a leader believes that her opponent may concede immediately with positive probability, then she never concedes immediately. That is, for all $i \in N, \omega_i \in \{I', \mathcal{U}\}$ and $j \neq i$, $F_i^{\omega_i}(0) \cdot G_j^{\omega_i}(0) = 0$.*

(P4) *One of the leaders must concede immediately with probability* 1 *(or never exaggerate) when she learns that the sentiment leans the other way; that is, $T_L[I'] \cdot T_R[I'] = 0$.*

We refer to the first property (P1) as the two-sided skimming property. Unlike the I' type, the uninformed type believes that her opponent could be the I' type. Therefore, if the leaders have not reached an agreement by time t, the uninformed type is at least as optimistic as the I' type regarding the probability of her opponent conceding in the next instance (see equations (1) and (2) and the discussion immediately following). Moreover, unlike the I' type, the uninformed type is uncertain about the public sentiment. Therefore, when the uninformed type exaggerates for a little longer, she discounts a smaller share of the surplus than the I' type. This makes the uninformed type more reluctant to concede than the I' type (see equation (4) and the discussion immediately following). Therefore, if the I' type is indifferent between conceding now and exaggerating for the next instance, the $\mathcal{U}$ type strictly prefers exaggerating. This gives us the skimming property.

The second (P2) and third (P3) properties are standard in reputational bargaining. For the formal argument, see AG. Below, I provide the intuition. Consider (P2). Recall that at time T_i, the uninformed leader j becomes convinced that leader i is committed and will never concede. Suppose, for contradiction, $T_i < T_j$. This means j continues exaggerating even after it is clear that i is committed and will never concede. However, this is not possible since a leader continues exaggerating only if there is a positive probability that the opponent may concede in the meantime. Next, consider (P3). Suppose that a non-committed type ω_i of leader i believes that her opponent concedes right away with positive probability ($G_j^{\omega_i}(0) > 0$). This means she strictly prefers exaggerating over conceding right away. Therefore, $F_i^{\omega_i}(0) = 0$.

Finally, the fourth property (P4) says that at least one leader who learns that the public sentiment leans the other way will concede immediately with probability 1 or never exaggerate. Suppose that $\omega_i = I'$ may exaggerate for some positive duration, that is, $T_i[I'] > 0$. Notice that $\omega_i = I'$ believes that her opponent j is either the $\omega_j = I$ type (and so never concedes) or the uninformed type ($\omega_j = \mathcal{U}$). Therefore, if she does not concede immediately rather continue exaggerating for the next instance, it must be that she believes that $\omega_j = \mathcal{U}$ may concede in the meantime. It then follows from the skimming property (P1) that, in equilibrium, $\omega_j = I'$ must have already conceded; that is, $T_j[I'] = 0$.

Two Phases

Suppose, as in (P4), it is leader j who never exaggerates $T_j[I'] = 0$. The second property (P2) implies that the leaders may continue bargaining until some date T, at which point both uninformed types will simultaneously become convinced that their opponent is committed. The first property (P1) implies that there will be two phases: in phase I, $\omega_i = I'$ will mix; in phase II, $\omega_i = \mathcal{U}$ will mix. However, $\omega_j = \mathcal{U}$ will mix in both phases (see Figure 1). They mix with probabilities that keep the type of opponent who mixes indifferent. Finally, (P3) implies that both $\omega_i = I'$ and $\omega_j = \mathcal{U}$ cannot concede immediately with positive probability: $F_i^{I'}(0) \cdot F_j^{\mathcal{U}}(0) = 0$.

Notice that the presence of uninformed types creates a phase I in which one of the leaders continues exaggerating even when she learns that the sentiment leans the other way (the I' type). We call her the strong bargainer and the other the weak bargainer.

Two Measures of Bargaining Strengths

To figure out which leader is the strong bargainer and which is the weak bargainer, as well as how unbalanced the bargaining strengths are, we define two measures of relative bargaining strengths $B_j^1(\theta, \varepsilon)$ and $B_j^2(\theta, \varepsilon)$. In reputational bargaining, an agent's bargaining strength is measured by how quickly she can convince her opponent about her commitment. Assuming that a non-committed type never concedes immediately, suppose T_i is the time i takes to convince her non-committed opponent about her commitment. Then, the relative bargaining strength of an agent j is $\exp(T_i)/\exp(T_j)$. A relative strength of less than 1 means that she takes longer, and this delay makes her a weak bargainer.

In our political bargaining setup, there are multiple non-committed types with different beliefs. I show that the appropriate notion of bargaining strengths in our setup is captured through two measures. Both of these measures assume that the uninformed types never concede immediately, and T_i is the time leader i takes to convince her uninformed opponent about her commitment. The two measures differ in terms of the assumption regarding what the leaders do when they learn that the sentiment leans the other way (the non-committed I' types).

Bargaining Strength $B_j^1(\theta, \varepsilon)$

Our first measure of relative bargaining strength of a leader $j \in N$ is $B_j^1 = \exp(T_i)/\exp(T_j)$, assuming that both leaders concede immediately when they learn that the sentiment leans the other way; that is,

$$F_i^{I'}(0) = 1 \text{ and } F_j^{I'}(0) = 1.$$

Under the above assumption, there is no phase I. By definition, $\omega_i = \mathcal{U}$ becomes convinced that leader j is committed at time T_j; that is, $G_j^{\mathcal{U}}(T_j) = 1 - \pi_j(1 - \varepsilon_j)$.

Therefore, to find T_i, T_j, we first need to understand $G_j^{\mathcal{U}}(t)$.

Suppose that j has not conceded until time t. We can see from (4) that if $\omega_i = \mathcal{U}$ concedes, she gets $\pi_j(t)(1-x_j)$, and if she exaggerates a little longer and concedes after time Δ, she gets $e^{-r_i\Delta}\pi_j(t+\Delta)(1-x_j)$. However, if the opponent concedes in the meantime, she will get x_i. Accordingly, $\omega_i = \mathcal{U}$ is indifferent if [8]

$$\frac{G_j^{\mathcal{U}}(t+\Delta) - G_j^{\mathcal{U}}(t)}{1-G_j^{\mathcal{U}}(t)}x_i + \left(1 - \frac{G_j^{\mathcal{U}}(t+\Delta) - G_j^{\mathcal{U}}(t)}{1-G_j^{\mathcal{U}}(t)}\right)e^{-r_i\Delta}\pi_j(t+\Delta)(1-x_j)$$
$$= \pi_j(t)(1-x_j).$$

Rearranging and taking $\Delta \to 0$, we get (see the appendix for details)

$$\frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1-G_j^{\mathcal{U}}(t)} = \frac{r_i\pi_j(t)(1-x_j) - \pi_j'(t)(1-x_j)}{x_i - \pi_j(t)(1-x_j)}. \tag{5}$$

Differentiating $\pi_j(t)$ in (3) and substituting $\pi_j(t)$ and $\pi_j'(t)$ in the above, we get (see the appendix for details)

$$\frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1+\pi_i(1-\varepsilon_j) - G_j^{\mathcal{U}}(t)} = \frac{r_i(1-x_j)\pi_j}{x_i - (1-x_j)\pi_j} =: \frac{1}{\eta_j}. \tag{η_j}$$

Recall that at the initial node, the uninformed leader i believes that the public sentiment leans the other way with probability π_j. Thus, the numerator $r_i(1-x_j)\pi_j$ captures the cost of exaggerating a little longer, while the denominator $x_i - (1-x_j)\pi_j$ captures the benefit from the opponent's concession evaluated at the initial node.

Solving this differential equation, we get

$$G_j^{\mathcal{U}}(t) = 1 + \pi_i(1-\varepsilon_j) - \exp\left(-\frac{1}{\eta_j}t\right). \tag{$G_j^{\mathcal{U}}$ - one phase}$$

Finally, solving $G_j^{\mathcal{U}}(T_j) = 1 - \pi_j(1-\varepsilon_j)$, we get

$$T_j = \ln\left((1-\varepsilon_j)^{-\eta_j}\right).$$

Therefore, our first measure of the relative bargaining strength of a leader $j \in N$ is

$$B_j^1(\theta,\varepsilon) := \frac{\exp(T_i)}{\exp(T_j)} = \frac{(1-\varepsilon_i)^{-\eta_i}}{(1-\varepsilon_j)^{-\eta_j}}. \tag{B_j^1}$$

Notice that $B_L^1(\theta,\varepsilon) = 1/B_R^1(\theta,\varepsilon)$. If $B_j^1(\theta,\varepsilon) = 1$, then under $F_i^{I'}(0) = F_j^{I'}(0) = 1$, we have $T_L = T_R$. If $B_j^1(\theta,\varepsilon) < 1$ for some j, then under $F_i^{I'}(0) = F_j^{I'}(0) = 1$, $T_i > T_j$.

[8] Here, the only difference from AG is that $\pi_j(t)$ varies with t, making the indifference conditions non-stationary. If $\pi_j(t) = 1$ for all t and accordingly $\pi_j'(t) = 0$, then this indifference condition coincides with that in AG.

Bargaining Strength $B_j^2(\theta, \varepsilon)$

Our second measure of bargaining strength $B_j^2(\theta, \varepsilon) = \exp(T_i)/\exp(T_j)$, assuming that leader j concedes immediately but leader i never concedes immediately when she learns that the sentiment leans the other way; that is,

$$F_j^{I'}(0) = 1 \text{ and } F_i^{I'}(0) = 0.$$

As in the previous case, we first need to understand $G_i^{\mathcal{U}}(t)$ and $G_j^{\mathcal{U}}(t)$. However, the difference is that under the above assumption, there are two phases:

- Phase I: $\omega_i = I'$ and $\omega_j = \mathcal{U}$ mix in $[0, T_i[I']]$
- Phase II: $\omega_i = \mathcal{U}$ and $\omega_j = \mathcal{U}$ mix in $[T_i[I'], T]$

Phase I: Conditional on no concession from j until time t, we can see from (4) that $\omega_i = I'$ is indifferent between conceding and exaggerating for the next Δ time if

$$\frac{G_j^{I'}(t+\Delta) - G_j^{I'}(t)}{1 - G_j^{I'}(t)} \cdot x_i + \left(1 - \frac{G_j^{I'}(t+\Delta) - G_j^{I'}(t)}{1 - G_j^{I'}(t)}\right) e^{-r_i\Delta}(1 - x_j) = (1 - x_j).$$

Notice that unlike the uninformed type, type I' knows that the public sentiment leans the other way. Therefore, $\pi_j(t)$ does not affect this indifference condition. Rearranging and taking $\Delta \to 0$, we get (see the appendix for details)

$$\frac{\frac{dG_j^{I'}(t)}{dt}}{1 - G_j^{I'}(t)} = \frac{r_i(1 - x_j)}{x_i - (1 - x_j)} =: \frac{1}{\lambda_j}. \tag{λ_j}$$

For $\omega_i = I'$, the cost of delaying the concession is $r_i(1 - x_j)$, while the benefit from the opponent's concession is $x_i - (1 - x_j)$. Solving this differential equation, we get

$$G_j^{I'}(t) = 1 - \exp\left(-\frac{1}{\lambda_j}t\right).$$

On the other hand, $\omega_j = \mathcal{U}$ is indifferent when (5) holds true (interchanging the subscripts i and j). However, unlike before, $\omega_i = \mathcal{U}$ has not yet started conceding ($F_i^{\mathcal{U}}(t) = 0$). Therefore (see equation (3)),

$$\pi_i(t) = \frac{\pi_i}{1 - G_i^{\mathcal{U}}(t)}.$$

Differentiating this and substituting $\pi_i(t)$ and $\pi_i'(t)$ in the indifference condition (5), we get (see the appendix for details)

$$\frac{dG_i^{\mathcal{U}}(t)}{dt} = \frac{r_j(1 - x_i)\pi_i}{x_j} = \frac{1}{\zeta_i}. \tag{ζ_i}$$

Recall that at the initial node, $\omega_j = \mathcal{U}$ believes that the public sentiment leans the other way with probability π_i. Therefore, the cost of exaggerating a little longer is $r_j(1-x_i)\pi_i$. The benefit arises if the opponent i concedes in the meantime. However, in phase I, the type of opponent who may concede is $\omega_i = I'$, and the opponent can be the I' type only if the public sentiment leans toward leader j. Recall that if $\omega_j = \mathcal{U}$ concedes when the public sentiment favors j, she gets 0. Thus, the benefit from the opponent conceding in the meantime is x_j. Solving this differential equation, we get

$$G_i^{\mathcal{U}}(t) = \frac{1}{\zeta_i}t.$$

By definition of $T_i[I']$, $\omega_j = \mathcal{U}$ believes that $\omega_i = I'$ will not exaggerate beyond time $T_i[I']$. Therefore, we have

$$G_i^{\mathcal{U}}(T_i[I']) = \pi_j(1-\varepsilon_i). \tag{6}$$

This gives us

$$T_i[I'] = \zeta_i \pi_j (1-\varepsilon_i). \tag{7}$$

Recall that while the $\omega_i = I'$ mixes in phase I, $\omega_i = \mathcal{U}$ strictly prefers exaggerating over conceding, and $\omega_i = I'$ believes that the opponent will concede by time $T_i[I']$ with probability $G_j^{I'}(T_i[I']) = 1 - \exp\left(-T_i[I']/\lambda_j\right)$. Unlike the $\omega_i = I'$ type, the uninformed type also assigns probability $\pi_i(1-\varepsilon_j)$ that leader j is the I' type and, hence, will never exaggerate and concede immediately. Therefore,

$$G_j^{\mathcal{U}}(T_i[I']) = 1 + \pi_i(1-\varepsilon_j) - \exp\left(-\frac{1}{\lambda_j}T_i[I']\right). \tag{8}$$

Phase II: Starting from $T_i[I']$, the $\mathcal{U}$ types from both sides mix. As in the previous case, solving the differential equation for indifference for any leader $j \in N$ (see equation (η_j)), we get

$$\begin{aligned} G_j^{\mathcal{U}}(t) &= 1 + \pi_i(1-\varepsilon_j) \\ &- \left(1 + \pi_i(1-\varepsilon_j) - G_j^{\mathcal{U}}(T_i[I'])\right) \exp\left(-\frac{1}{\eta_j}(t - T_i[I'])\right) \end{aligned} \tag{$G_j^{\mathcal{U}}$ - two phases}$$

The difference from ($G_j^{\mathcal{U}}$ - one phase) is that there are two phases. When there is no phase I, we have $G_j^{\mathcal{U}}(T_i[I'] = 0) = \pi_i(1-\varepsilon_j)$ for all $i \in N$. However, now we have a phase I, which lasts for $T_i[I']$ (as in (7)). From phase I, we know $G_i^{\mathcal{U}}(T_i[I'])$ and $G_j^{\mathcal{U}}(T_i[I'])$ (see (6) and (8)). Substituting these expressions in ($G_j^{\mathcal{U}}$ - two phases), we get that in the time interval $[T_i[I'], T]$, the belief of the uninformed types must be as follows:

$$G_i^{\mathcal{U}}(t) = 1 + \pi_j(1-\varepsilon_i) - \exp\left(-\frac{1}{\eta_i}(t - T_i[I'])\right)$$

$$G_j^{\mathcal{U}}(t) = 1 + \pi_i(1-\varepsilon_j) - \exp\left(-\frac{1}{\lambda_j}T_i[I'] - \frac{1}{\eta_j}(t - T_i[I'])\right).$$

For any $j \in N$, $\omega_i = \mathcal{U}$ believes that a non-committed leader j will not exaggerate beyond time T_j. Therefore, we have $G_j^{\mathcal{U}}(T_j) = 1 - \pi_j(1-\varepsilon_j)$. Solving this and substituting $T_i[I']$ using (7), we get

$$T_i = \ln((1-\varepsilon_i)^{-\eta_i}) + \zeta_i \pi_j (1-\varepsilon_i).$$

$$T_j = \ln((1-\varepsilon_j)^{-\eta_j}) + \zeta_i \pi_j (1-\varepsilon_i)\left(1 - \frac{\eta_j}{\lambda_j}\right).$$

Define

$$B_j^2(\theta,\varepsilon) := \frac{\exp(T_i)}{\exp(T_j)} = \frac{(1-\varepsilon_i)^{-\eta_i} \cdot \chi_j(\theta,\varepsilon)}{(1-\varepsilon_j)^{-\eta_j}}, \qquad (B_j^2)$$

where

$$\chi_j(\theta,\varepsilon) = \exp\left(\frac{\eta_j}{\lambda_j}\zeta_i \pi_j (1-\varepsilon_i)\right).$$

Notice that $\chi_j = \exp(\frac{\eta_j}{\lambda_j} T_i[I'])$, where $T_i[I']$ is the length of phase I and η_j/λ_j is the rate adjustment since the uninformed weak bargainer j concedes at different rates in the two phases. Since for any (θ,ε), $\chi_j \geq 1$, $B_j^2 \geq B_j^1$ for any $j \in N$.

It is easy to see that if $B_j^2 = 1$, then under $F_i^{I'}(0) = 0, F_j^{I'}(0) = 1$, we have $T_L = T_R$. If $B_j^2 > 1$ for some j, then under $F_i^{I'}(0) = 0, F_j^{I'}(0) = 1$, $T_i > T_j$. If $B_j^2 < 1$ for some j, then under $F_i^{I'}(0) = 0, F_j^{I'}(0) = 1$, $T_i < T_j$.

Degree of Imbalance

Since $B_R^1(\theta,\varepsilon) = 1/B_L^1(\theta,\varepsilon)$, for any (θ,ε), $B_j^1 \leq 1$ for at least one of the leaders. Suppose that $B_j^1 \leq 1$ for leader $j = L$ or R. If both leaders concede immediately when they learn that the sentiment leans the other way, then leader j will take longer than leader i to convince the uninformed opponent about her commitment. This makes leader j the weak bargainer. Below, we divide such unbalanced bargaining strength situations in two categories based on B_j^2.

Moderately Unbalanced Strengths Suppose (θ,ε) is such that

$$B_j^2 \geq 1 \geq B_j^1$$

for some $j \in N$. Then, under $F_i^{I'}(0) = 1, F_j^{I'}(0) = 1$, we have $T_j > T_i$, but under $F_i^{I'}(0) = 0, F_j^{I'}(0) = 1$, we have $T_j < T_i$. Figure 2 depicts the uninformed types' beliefs over time that the opponent is committed. Red curves capture the belief about j, and blue curves capture the belief about i. The dotted curves represent the belief if $F_i^{I'}(0) = 1, F_j^{I'}(0) = 1$ (as in the first measure of relative bargaining strength). If $B_j^1 = 1$, the two dotted lines reach 1 simultaneously. However, given $B_j^1 < 1$, we can see that j takes longer to convince the uninformed opponent. This makes j the weak bargainer. The dashed curves represent the belief if $F_i^{I'}(0) = 0, F_j^{I'}(0) = 1$ (as in the second measure of relative bargaining strength). If $B_j^2 = 1$, the two dashed lines reach 1 simultaneously. However, given $B_j^2 > 1$, we can see that i takes longer to convince the uninformed opponent.

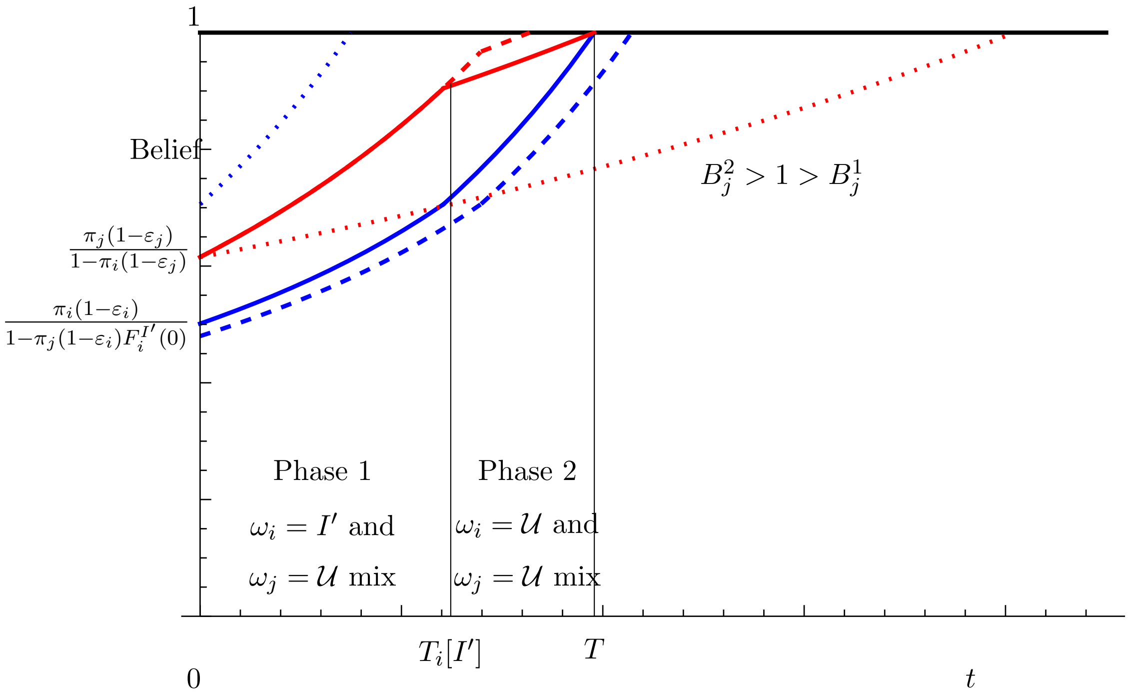

Figure 2: Belief of the uninformed types that the opponent is committed: $\frac{\pi_i(1-\varepsilon_i)}{1-G_i^{\mathcal{U}}(t)}$. See the text for a detailed explanation. Specification: $r = (1, 1), \pi = (0.6, 0.4), x = (0.65, 0.915), \varepsilon = (0.2, 0.2)$.

This means leader j is weak enough that leader i exaggerates after learning that the public sentiment leans the other way, but not so weak that leader i never concedes immediately. That is, it must be that $F_i^{I'}(0) \in (0, 1)$. As $F_i^{I'}(0)$ increases, $T_i[I']$ falls, and accordingly, T_i falls and T_j increases. I show that to reach an equilibrium path of belief, we must have

$$F_i^{I'}(0) = 1 + \left(\frac{\ln B_j^1}{\ln B_j^2 - \ln B_j^1} \right).$$

The solid curves in Figure 2 show the equilibrium belief. Notice that as long as $F_i^{I'}(0) > 0$, the uninformed types will always exaggerate (P3).

Extremely Unbalanced Strengths Suppose (θ, ε) is such that

$$1 \geq B_j^2$$

for some $j \in N$. Then, under $F_i^{I'}(0) = 1$ and $F_j^{I'}(0) = 1$, we have $T_j > T_i$, and under $F_i^{I'}(0) = 0$ and $F_j^{I'}(0) = 1$, we also have $T_j > T_i$. Figure 3 depicts the uninformed types' beliefs over the time that the opponent is committed. The difference compared with Figure 2 is that when $F_i^{I'}(0) = 0, F_j^{I'}(0) = 1$ (as in the second measure of relative bargaining strength), the belief about leader j reaches 1 later than the belief about leader i (the dashed blue curve coincides with the

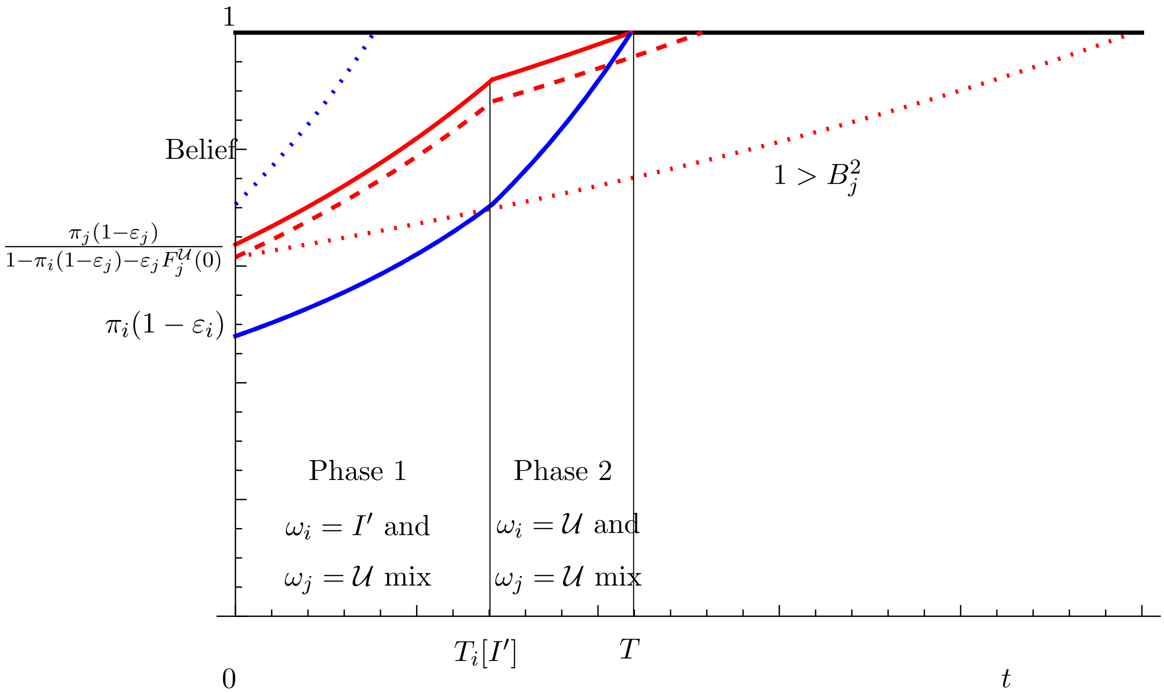

Figure 3: Belief of the uninformed types that the opponent is committed: $\frac{\pi_i(1-\varepsilon_i)}{1-G_i^{\mathcal{U}}(t)}$. See the text for detailed explanation. Specification: $r = (1,1), \pi = (0.6, 0.4), x = (0.65, 0.93), \varepsilon = (0.2, 0.2)$.

solid blue curve). This means that even the uninformed type of leader j must concede immediately with positive probability ($F_j^{\mathcal{U}}(0) > 0$). As $F_j^{\mathcal{U}}(0)$ increases, T_j falls, but $T_i[I']$ or T_i remains unaffected. I show that to reach the equilibrium path of belief, we must have

$$F_j^{\mathcal{U}}(0) = \frac{1}{\varepsilon_j}\left(1 - (B_j^2)^{1/\eta_j}\right).$$

Finally, from $G_j^{\omega_i}(t)$, using (1), (2), we can uniquely identify $F_i^{\omega_i}(t)$ for all $i \in N$ and $\omega_i \in \{I', \mathcal{U}\}$. The following proposition summarizes the war of attrition equilibrium.

Proposition 1: *Under continuous-time bargaining, assuming the I type is committed to the policy position to which they agree when public sentiment is commonly known, a unique equilibrium emerges that resembles a war of attrition with multiple non-commitment types on both sides. In equilibrium, these non-commitment types $\omega_i \in \{I', \mathcal{U}\}$ for each leader $i \in N$ may masquerade as the I type (they exaggerate how their constituents would react to their concession) and concede by time t with probability $F_i^{\omega_i}(t)$, which is specified as follows:*

1. *If (θ, ε) is such that $B_j^2(\theta, \varepsilon) \geq 1 \geq B_j^1(\theta, \varepsilon)$ for some $j \in N$, then the bargaining strengths are moderately unbalanced: leader j is moderately weak, and leader i is moderately strong.*

(a) *if leader j learns that the public sentiment leans the other way, she concedes immediately with probability* 1*:*

$$F_j^{I'}(t) = \mathbb{1}(t \geq 0).$$

(b) *if leader i learns that the public sentiment leans the other way, she mixes between when to concede in the first phase* $[0, T_i[I']]$*, where*

$$F_i^{I'}(t) = F_i^{I'}(0) + t/(\pi_j(1-\varepsilon_i)\zeta_i),$$

$$F_i^{I'}(0) = 1 + \left(\frac{\ln B_j^1}{\ln B_j^2 - \ln B_j^1}\right),$$

$$T_i[I'] = \zeta_i \pi_j (1-\varepsilon_i)(1 - F_i^{I'}(0)).$$

(c) *if leader j is uninformed, she mixes between when to concede in first phase* $[0, T_i[I']]$ *and in the second phase* $[T_i[I'], T]$*, where*

$$F_j^{\mathcal{U}}(t) = \begin{cases} \frac{1}{\varepsilon_j}\left(1 - \exp\left(-\frac{t}{\lambda_j}\right)\right) & \textit{if } t \leq T_i[I'] \\ \frac{1}{\varepsilon_j}\left(1 - \exp\left(-\frac{T_i[I']}{\lambda_j} - \frac{(t - T_i[I'])}{\eta_j}\right)\right) & \textit{if } t \geq T_i[I'], \end{cases}$$

$$T = \ln((1-\varepsilon_i)^{-\eta_i}) + \zeta_i \pi_j (1-\varepsilon_i)(1 - F_i^{I'}(0)).$$

(d) *if leader i is uninformed, she does not concede in the first phase* $[0, T_i[I']]$ *and mixes between when to concede in the second phase* $[T_i[I'], T]$*, where*

$$F_i^{\mathcal{U}}(t) = \frac{1}{\varepsilon_i}\left(1 - \exp\left(-\frac{(t - T_i[I'])}{\eta_i}\right)\right) \mathbb{1}(t \geq T_i[I']).$$

2. *If (θ, ε) is such that $1 \geq B_j^2(\theta, \varepsilon)$ for some $j \in N$, then the bargaining strengths are extremely unbalanced: leader j is extremely weak, and leader i is extremely strong.*

   (a) *if leader j learns that the public sentiment leans the other way, she concedes immediately with probability* 1*:*

$$F_j^{I'}(t) = \mathbb{1}(t \geq 0).$$

   (b) *if leader i learns that the public sentiment leans the other way, she mixes between when to concede in the first phase* $[0, T_i[I']]$*, where*

$$F_i^{I'}(t) = t/(\pi_j(1-\varepsilon_i)\zeta_i),$$

$$T_i[I'] = \zeta_i \pi_j (1-\varepsilon_i).$$

(c) *if leader j is uninformed, she mixes between when to concede in first phase $[0, T_i[I']]$ and in the second phase $[T_i[I'], T]$, where*

$$F_j^{\mathcal{U}}(t) = \begin{cases} \frac{1}{\varepsilon_j}\left(1 - (1 - \varepsilon_j F_j^{\mathcal{U}}(0)) \exp\left(-\frac{t}{\lambda_j}\right)\right) & \text{if } t \leq T_i[I'] \\ \frac{1}{\varepsilon_j}\left(1 - (1 - \varepsilon_j F_j^{\mathcal{U}}(0)) \exp\left(-\frac{T_i[I']}{\lambda_j} - \frac{(t - T_i[I'])}{\eta_j}\right)\right) & \text{if } t \geq T_i[I'], \end{cases}$$

$$F_j^{\mathcal{U}}(0) = \frac{1}{\varepsilon_j}\left(1 - (B_j^2)^{1/\eta_j}\right),$$

$$T = \ln((1 - \varepsilon_i)^{-\eta_i}) + \zeta_i \pi_j (1 - \varepsilon_i).$$

(d) *if leader i is uninformed, she does not concede in the first phase $[0, T_i[I']]$ and mixes between when to concede in the second phase $[T_i[I'], T]$, where*

$$F_i^{\mathcal{U}}(t) = \frac{1}{\varepsilon_i}\left(1 - \exp\left(-\frac{(t - T_i[I'])}{\eta_i}\right)\right) \mathbb{1}(t \geq T_i[I']).$$

The result follows from the discussion preceding the proposition. For the formal proof, see the appendix. When $B_j^1 = 1$, $F_i^{I'}(t) = F_j^{I'}(t) = \mathbb{1}(t \geq 0)$. That is, a leader concedes immediately when she learns that the public sentiment leans the other way. In this case, we say that the bargaining strengths are exactly balanced. On the other hand, if $B_j^2 = 1$, then $F_i^{I'}(0) = F_j^{\mathcal{U}}(0) = 0$. In this case, leader i never concedes immediately, and leader j always concedes immediately when she learns that the public sentiment leans the other way.

It is important to note the equilibrium strategy of a leader when she learns that the sentiment leans the other way. Consider, for instance, $\omega_L = I'$. It follows from the above proposition that if (θ, ε) is such that (1) $B_R^1 \geq 1$, then L is the weak bargainer and concedes immediately with probability 1 (never exaggerates). If (θ, ε) is such that (2) $B_R^2 > 1 > B_R^1$, then L is a moderately strong bargainer and concedes immediately with probability $F_i^{I'}(0) < 1$. She may continue exaggerating until $T_i[I'] = \zeta_i \pi_j (1 - \varepsilon_i)(1 - F_i^{I'}(0))$. If (θ, ε) is such that (3) $1 \geq B_R^2$, then L is an extremely strong bargainer and never concedes immediately; that is, $F_i^{I'}(0) = 0$. She may continue exaggerating until $T_i[I'] = \zeta_i \pi_j (1 - \varepsilon_i)$.

Effect of Tolerance Readers may wonder whether having less tolerant constituents (say, higher x_L) helps a leader L in bargaining. When the constituents of leader L are less tolerant, it is easier to portray her as a sellout when the sentiment leans toward L (state $\mathscr{L}$). When it is commonly known that the state is $\mathscr{L}$, a higher x_L gives more leverage to leader L, and in equilibrium, she gets a more preferred policy and, accordingly, a larger payoff. However, under incomplete information, the effect of higher intolerance is more nuanced. A higher x_L means that when a non-committed leader R concedes, L gets a higher share and R a lower share. However, this may adversely affect the bargaining strengths of leader L, and thus, having more intolerant constituents does not necessarily help a leader. For a numerical example, notice that the only difference in the parametric specifications in Figures 2 and 3 is that x_j is higher in the second case, which turns the moderately weak bargainer leader j into an extremely weak bargainer.[9]

The effect of tolerance on bargaining strength is not monotonic. We can see from (4) that when x_L is higher, the benefit of exaggerating a little longer is higher for leader L, and the cost of exaggerating a little longer is lower for leader R. Thus, both the leaders become more reluctant to concede when they are not committed. In reputational bargaining, when the opponent becomes more reluctant to concede, in equilibrium, a leader must concede at a slower rate to keep the opponent indifferent (otherwise, the opponent would prefer exaggerating over conceding), which means she will take longer to convince her opponent regarding her commitment. Since both leaders concede at slower rates, a higher x_L can have an ambiguous effect on bargaining strengths. However, when x_L is sufficiently high, the cost effect dominates (since cost becomes close to zero), which turns L very weak.

2.2. Almost Complete Information

A leader concedes immediately when it is commonly known that the public sentiment leans the other way. However, when there is a positive probability that the opponent may not know this, the leader may exaggerate the possible reaction of her constituents to her concession. In this section, we study whether the leaders continue doing so even when the probability that a leader is uninformed about the public sentiment is close to zero. For any (θ, ε), under the continuous-time limit and commitment of the I type, Proposition 1 characterizes the resulting unique equilibrium, which resembles a war of attrition. In this section, we consider the case of almost complete information: the probability that the leaders do not know the public sentiment $\varepsilon = (\varepsilon_L, \varepsilon_R)$ is close to $(0, 0)$. We investigate whether, in every state, the distribution of the equilibrium outcome (p, t) is close to that when the public sentiment is commonly known.

Put differently, we perturb the political bargaining game in which public sentiment is commonly known by introducing a small probability that the leaders are uninformed about the underlying sentiment. This means that under ε close to $(0, 0)$, when the sentiment leans toward a leader, it is likely that she is the I type (and committed) and her opponent is the I' type. This is in contrast to the reputational perturbation where in the limit, neither agent is committed. AG shows that in the limit, an agreement is reached immediately with probability 1.[10]

[9] When the strengths are only moderately unbalanced, an uninformed weak bargainer, say, L, expects that the opponent may concede immediately if the opponent learns that public sentiment leans toward L. However, if L turns into a very weak bargainer, she no longer expects this, and her expected payoff falls.

[10] Reputational bargaining treats committed types as irrational and studies the limit where both players are likely non-committed. In contrast, in our political bargaining setup, one of the leaders has reasons to be committed (which comes from the public sentiment). We study the limit where the leaders are likely to know who has legitimate reasons for being committed.

Since it is likely that a leader is the I type when the sentiment leans toward her, and the I types always insist on the same policy p to which they agree when the public sentiment is commonly known, the probability that they will reach an agreement based on a different policy position is negligible. However, this does not mean they will reach such an agreement immediately. It depends on how long a leader exaggerates even after learning that the public sentiment is unfavorable.

Recall (from Proposition 1) that a weak bargainer never exaggerates after learning that the public sentiment leans the other way, whereas a strong bargainer may exaggerate in phase I. So, we need to understand what happens to phase I when ε is small. To see this, we first fix a bargaining environment θ and then consider a sequence of $\{\varepsilon_n\}$ that converges to $(0,0)$ as $n \to \infty$. Suppose that the corresponding equilibrium strategies are $F_{in}(t), F_{jn}(t)$.

Proposition 2: *Given any bargaining environment θ, as $\{\varepsilon_n\} \to (0,0)$, in the continuous-time war of attrition equilibrium, the leaders almost never exaggerate when they learn that the public sentiment leans the other way; that is, for all $i \in N$,*

$$\lim_{n\to\infty} F_{in}^{I'}(t) = \mathbb{1}(t \geq 0).$$

This means that the probability that the leaders will immediately agree on policy position x_L in state $\mathscr{L}$ and $1 - x_R$ in state $\mathscr{R}$ converges to 1.

The argument involves three simple steps. First, I show that for any θ, when n is sufficiently large, the bargaining strengths cannot be extremely unbalanced. That is, for sufficiently large n, for any $j \in N$, $B_j^2(\theta, \varepsilon_n) > 1$.

Second, under such large n, I show that the bargaining strengths are almost balanced; that is, $B_j^1(\theta, \varepsilon_n)$ is close to 1 for any j. This means that even the strong bargainer—say, leader i—almost never exaggerates when she learns that the public sentiment leans the other way, that is, $\lim_{n\to\infty} F_{in}^{I'}(0) = 1$. Recall that the weak bargainer never exaggerates when she learns that the public sentiment leans the other way. Therefore, $\lim_{n\to\infty} F_{in}^{I'}(t) = \mathbb{1}(t \geq 0)$ for all $i \in N$.

Third, notice that when ε is small, both leaders are likely to know the public sentiment (that is, either type I or type I'). If the I' type leaders almost never exaggerate after learning that the public sentiment leans the other way, then the probability that they will immediately agree on the same policy as they do when the public sentiment is commonly known is close to 1.

It is important to note that in the above robustness argument, we consider the bargaining environment θ as fixed. It is not necessary that for a given small ε, the bargaining strengths will be nearly balanced regardless of the bargaining environment θ. In fact, as the following proposition shows, there exists θ where the bargaining strengths can be extremely unbalanced.

Proposition 3: *Given any ε (however small), there exists bargaining environment $\theta = (r, x, \pi)$ such that the bargaining strengths are extremely unbalanced. Accordingly, in the continuous-time war of attrition equilibrium, the strong bargainer—*

say, leader i—never concedes immediately when she learns that the public sentiment leans the other way ($F_i^{I'}(0) = 0$) and may continue exaggerating for a positive duration (phase I), where

$$F_i^{I'}(t) = \frac{r_j \pi_i (1 - x_i)}{x_j \pi_j (1 - \varepsilon_i)} \cdot t.$$

This means that when the public sentiment leans toward the weak bargainer, the probability of immediate agreement at the complete information policy position is 0.

We prove this result by constructing a bargaining environment $\theta = (r, x, \pi)$ such that given ε, $B_j^2(\theta, \varepsilon) < 1$. For this construction, suppose without loss of generality that ε is such that $\varepsilon_L \geq \varepsilon_R > 0$. Consider a bargaining environment θ in which the L constituents are extremely intolerant when $\omega = \mathscr{L}$; that is, x_L is close to 1. This means that if leader R concedes, she gets almost nothing. This makes leader R very reluctant to concede. Therefore, in equilibrium, to keep leader R indifferent between conceding and exaggerating for a little longer, leader L must concede at a very slow rate. Otherwise, R will always exaggerate rather than concede. As we have mentioned before, higher x_L can have an ambiguous effect on the bargaining strengths. However, when x_L becomes sufficiently large, it turns leader L into an extremely weak bargainer. To see this, note that (see (η_j),(ζ_i),(λ_j)) as $x_L \to 1$, $\lambda_L, \zeta_L, \eta_L \to \infty$, while λ_R, ζ_R, and η_R are finite. I show that for any ε, such that $\varepsilon_L \geq \varepsilon_R > 0$, we can find x_L sufficiently close to 1 such that

$$B_L^2(\theta, \varepsilon) = \frac{(1 - \varepsilon_R)^{-\eta_R}}{(1 - \varepsilon_L)^{-\eta_L}} \cdot \exp\left(\frac{\eta_L}{\lambda_L} \cdot \zeta_R \pi_L (1 - \varepsilon_R)\right) < 1.$$

Since L is extremely weak, in equilibrium, when leader R learns that the public sentiment leans toward L, she never concedes immediately ($F_R^{I'}(0) = 0$) and may continue exaggerating in phase I. It follows from Proposition 1 that

$$F_R^{I'}(t) = \frac{1}{\pi_L (1 - \varepsilon_R) \zeta_R} \cdot t = \frac{r_L \pi_R (1 - x_R)}{x_L \pi_L (1 - \varepsilon_R)} \cdot t,$$

which reaches 1 at $T_R[I']$, where $T_R[I'] = \zeta_R \pi_L (1 - \varepsilon_R) = \frac{x_L \pi_L (1 - \varepsilon_R)}{r_L \pi_R (1 - x_R)} > 0$. Recall that when it is commonly known that the sentiment leans toward L ($\omega = \mathscr{L}$), leader R immediately concedes. In sharp contrast, when there is a small positive probability ε that the leaders are uninformed, in state $\omega = \mathscr{L}$, there is zero probability that R will concede immediately.

3. Discussion

In this section, I elaborate on the relation to the existing literature and discuss some existing results (or their minor variations) that play important roles behind our result.

3.1. Endogenous Commitment

In Section 2, we assume that the I types are committed to the same policy to which they agree when the public sentiment is commonly known. In this section, we show that this commitment arises endogenously and discuss the features of our setup that drive this result.

Lemma 2: *As $\Delta \to 0$, in any PBE, where the leaders never make a demand that is rejected by all opponent types at all histories of the game, the following must be true: (1) $\omega_L = I$ always demands the policy position x_L, agrees to any policy $p \geq x_L$, and rejects any policy $p < x_L$. (2) $\omega_R = I$ always demands the policy position $1 - x_R$, agrees to any policy $p \leq 1 - x_R$, and rejects any policy $p > 1 - x_R$.*

Watson (1998) shows that under incomplete information in bargaining, in any PBE, regardless of their type, agents accept an offer above the maximum share they can get under complete information with probability 1 and accept an offer lower than the minimum share they can get under complete information with probability 0. In our political bargaining setup, under the complete information benchmark, when $\Delta \to 0$, the maximum share leader i can get is x_i, and the minimum share she can get is $1 - x_i$. Therefore, when $\Delta \to 0$, in any PBE, leader L will accept any policy position $p \geq x_L$ with probability 1 and reject any policy position $p < 1 - x_R$ with probability 1. Analogously, leader R will accept any policy position $p \leq 1 - x_R$ with probability 1 and reject any policy position $p > x_L$ with probability 1. Consider, say, leader L. Recall that the $\omega_L = I$ accepts $p < x_L$ with probability 0. Moreover, if she ever demands a policy position $p > x_L$, it will never be accepted by any opponent type at any history of the game. We assume that in equilibrium, the leaders never make such demands. Thus, $\omega_L = I$ becomes endogenously committed to x_L.

It is important to note that this endogenous commitment comes from two features of our political bargaining setup. First, the I type knows the public sentiment. Otherwise, she will be willing to accept a policy that is worse than what she agrees to when the sentiment is commonly known. Second, even when it is commonly known that the sentiment leans toward a leader i, she cannot get a better policy than the worst policy she considers acceptable (deal me out). This is because a compromise can be used to portray a leader as a sellout in the next primary only when the compromise is sufficiently large. In the next section, we discuss an alternative setup in which even small compromises can have adverse electoral effects, and accordingly, commitment may not arise endogenously.

3.2. Uniqueness and Alternative Setup

The early literature on bargaining with two-sided asymmetric information (see, for instance, Chatterjee and Samuelson (1987, 1988)) shows that many equilibria can be constructed using belief-based threats where the leaders are harshly punished through beliefs for deviating from a proposed equilibrium strategy—identified as the weakest type and given a low continuation value. In contrast, in our political bargaining setup, a unique equilibrium emerges. This difference comes from the fact that the I types are committed, which makes them immune to belief-based threats. AG first established the uniqueness result assuming commitment types. However, notice that unlike in AG, the leaders are perfectly rational. Nevertheless, the I types become endogenously committed as in AG (see Lemma 2).

To understand the crucial difference between our political bargaining setup and the setup in Chatterjee and Samuelson (1987) that drives this endogenous commitment, consider the following alternative payoff specification:

$$\begin{array}{lll} u_L(p,\mathscr{L}) = \frac{p-x_L}{1-x_L} & \text{and} & u_R(p,\mathscr{L}) = (1-p). \\ u_L(p,\mathscr{R}) = p & \text{and} & u_R(p,\mathscr{R}) = \frac{1-x_R-p}{1-x_R}. \end{array}$$

In our original setup, the sentiment is used to portray a leader as a sellout provided she makes a sufficiently large compromise. A small compromise is not salient, and sentiment makes no difference. However, under the above specification, sentiment matters even for small compromises. To see this, consider leader L and $\omega = \mathscr{L}$. As in our original setup, leader L will never agree to $p < x_L$. However, notice that even when $p > x_L$, the payoff depends on the degree of tolerance: having more intolerant constituents (higher x_L) means that L gets a lower payoff even from such small compromises. Therefore, when x_L is higher, if L continues bargaining, she discounts a lower share $(p - x_L)/(1 - x_L)$. This makes her more willing to continue bargaining. Accordingly, under complete information, in equilibrium, she gets

$$x_L + \frac{r_R}{r_L + r_R}(1 - x_L).$$

This result is commonly referred to as *split the difference*. Each side takes the surplus she can get outside (x_L for leader L and 0 for leader R), plus her Rubinstein share of the net surplus $(1 - x_L)$.

Notice that, unlike in our original setup, when the sentiment is commonly known to lean toward L, the leaders immediately agree to a policy $p > x_L$, which means that L gets a better policy than the worst policy she considers acceptable. Binmore et al. (1986) first pointed out the conceptual difference between *deal me out* and *split the difference* solutions while making the connection between the Rubinstein bargaining and Nash bargaining solution. This makes the I type susceptible to belief-based threats. We may construct a PBE in which leader L is committed to demanding the above share because off-path the opponent assumes that she is the I' type (or the $\mathcal{U}$ type if the opponent is the I' type). Chatterjee and Samuelson (1988) construct a similar war of attrition equilibrium but show that other equilibria could be constructed as well using such belief-based threats. This means that for a given ε (however small), we may find other ways in which substantial delays may arise in equilibrium (other than as specified in Proposition 3).

3.3. Reputational Bargaining

The war of attrition style equilibrium is a common feature of the reputational bargaining literature. See, for instance, Kambe (1999), Wolitzky (2012), Atakan and Ekmekci (2014), Özyurt (2015), Fanning (2016, 2018, 2021), and Ekmekci and Zhang (2021). This literature is agnostic about what drives the commitment, however. Instead, it allows for many commitment types and finds out which commitment types the rational agent will mimic. In contrast, this paper considers a political bargaining problem with a small probability that the leaders are uninformed about the underlying state—the public sentiment. Under binary states, commitment arises endogenously (see Section 3.1), which leads to a war of attrition similar in spirit to reputational bargaining.

Throughout this paper, we focus on the continuous-time limit. AG establishes that as the time interval between offers $\Delta \to 0$, the outcome under a discrete-time bargaining equilibrium converges in distribution to the war of attrition equilibrium under the continuous-time limit (Proposition 4 in AG). Assuming commitment types, AG shows that this distributional convergence follows from the Coasian result (see Proposition 8.4 in Myerson (1991) or Lemma 1 in AG), which shows that under $\Delta \to 0$, when a leader reveals that she is not committed and the opponent has not done so, she immediately concedes.[11] In our setup, when $\Delta \to 0$, the I types behave like the commitment type in AG (see Lemma 2). Although, unlike in AG, the non-committed uninformed type updates her belief about the state, the same Coasian argument can easily be extended.

Lemma 3: *As $\Delta \to 0$, if a non-committed type I' or $\mathcal{U}$ of leader i ever reveals that she is not the committed I type, while the opponent (leader j) has not done so, then she immediately concedes.*

Since the proof is mostly standard in the literature, I omit the formal proof from the main paper and relegate it to the online appendix. The idea is as follows. Under discrete-time bargaining, a leader may gradually reduce her demand rather than concede immediately. However, when $\Delta \to 0$, (1) the bargaining must end in finite time, and (2) the non-committed leader always prefers to concede now rather than continue bargaining if she knows that the bargaining will end in the next ϵ time for some small positive ϵ. Thus, the effect of asymmetric information overwhelms the effect of impatience. Given Lemmas 2 and 3, one can use the same argument as in AG to show that the outcome under discrete-time bargaining converges in distribution to the outcome under the continuous-time war of attrition equilibrium.

It is important to note that the information perturbation is different from the reputational perturbation. Under reputational bargaining, commitment types are irrational, and AG studies the limit when neither leader is committed. The authors

[11] It is worth mentioning that AG shows that an alternate offer bargaining protocol is not required for this result.

show that when the probability that the agents are committed is close to zero, they reach an agreement almost immediately. In contrast, in our setup, a rational leader could be committed, and when $\varepsilon \to (0,0)$, in each state, one of the leaders is a committed I type and the other is a non-committed I' type. As we saw in Proposition 3, even under a small ε, it is possible that in one state, the leaders never reach an immediate agreement.

It is also worth pointing out some differences between reputational bargaining and the nature of the war of attrition in our political bargaining setup. In reputational bargaining, both sides could be committed, so they never reach an agreement. However, in our setup, the public sentiment leans toward one leader or the other. Therefore, only one of the leaders can be the committed type. Thus, regardless of the state, an agreement will eventually be reached. Perhaps somewhat surprisingly, under reputational bargaining, the committed types get a lower expected payoff relative to their rational counterparts (see Sanktjohanser (2018)). This payoff loss comes from the fact that, unlike the rational agent, the committed type remains committed even in the end (at time T) when the opponent does not give up. However, the committed leader is committed in our setup because she privately learns that the public sentiment leans toward her. Unlike the uninformed leader, she knows that the opponent cannot actually be committed. Therefore, in equilibrium, the committed leader's expected payoff is higher than that of the uninformed leader. For a formal comparison of the equilibrium payoff, see the online appendix.

3.4. One-sided uncertainty

Feinberg and Skrzypacz (2005) studies a bargaining game in which a seller offers a price over time and a buyer decides whether to accept the price. While the buyer knows the value, the seller privately learns the value with positive probability. In this section, we consider a special case that is close to such setup. Consider a sequence $\{\varepsilon_{in}\} \to 0$ and $\{\varepsilon_{jn}\} \to \varepsilon_j > 0$. That is, leader i is an informed type (I or I') with a probability of almost 1, but leader j could be uninformed. We check how the equilibrium in Proposition 1 would behave in this limit.

Notice that $\lim_{n\to\infty} B^1_{jn} = (1-\varepsilon_j)^{\eta_j} < 1$; that is, in the limit, leader j is the weak bargainer. Therefore, leader j never exaggerates when she learns that the public sentiment leans the other way $\lim_{n\to\infty} F^{I'}_{jn}(t) = \mathbb{1}(t \geq 0)$. Regardless of the bargaining strengths, $\lim_{n\to\infty} T_{in} = \lim_{n\to\infty} T_{in}[I']$; that is, in the limit, the game only lasts during phase I.

Recall that in phase I, $\omega_i = I'$ and $\omega_j = \mathcal{U}$ mix between exaggerating and conceding. How long phase I lasts depends on ε_j.

Case 1: If $\varepsilon_j \in (1-\exp(-\zeta_i\pi_j/\lambda_j), 1)$, then

$$\lim_{n\to\infty} B^2_{jn} = (1-\varepsilon_j)^{\eta_j} \exp\left(\frac{\eta_j}{\lambda_j}\zeta_i\pi_j\right) < 1.$$

Therefore, it follows from Proposition 1 that in the limit, leader j is extremely weak, and

$$\lim_{n\to\infty} F_{in}^{I'}(t) = \frac{t}{\zeta_i \pi_j}$$

$$\lim_{n\to\infty} F_{jn}^{\mathcal{U}}(t) = \frac{1}{\varepsilon_j}\left(1-(1-\varepsilon_j)\exp\left(-\frac{(t-\zeta_i\pi_j)}{\lambda_j}\right)\right).$$

Notice that leader i never concedes immediately, whereas even the uninformed leader j may concede immediately with positive probability. In the limit, phase I may last until $\lim_{n\to\infty} T_{in}[I'] = \zeta_i\pi_j$.

Case 2: If $\varepsilon_j \in (0, 1-\exp(-\zeta_i\pi_j/\lambda_j))$, then

$$\lim_{n\to\infty} B_{jn}^2 = (1-\varepsilon_j)^{\eta_j}\exp\left(\frac{\eta_j}{\lambda_j}\zeta_i\pi_j\right) > 1.$$

Therefore, it follows from Proposition 1 that in the limit, leader j is moderately weak, and

$$\lim_{n\to\infty} F_{in}^{I'}(t) = 1 + \frac{\lambda_j}{\zeta_i\pi_j}\ln(1-\varepsilon_j) + \frac{t}{\zeta_i\pi_j}$$

$$\lim_{n\to\infty} F_{jn}^{\mathcal{U}}(t) = \frac{1}{\varepsilon_j}\left(1-\exp\left(-\frac{t}{\lambda_j}\right)\right).$$

Notice that leader i may concede immediately with positive probability when she learns that the public sentiment leans the other way, whereas the uninformed leader j never concedes immediately. In the limit, phase I can last until $\lim_{n\to\infty} T_{in}[I'] = -\lambda_j\ln(1-\varepsilon_j)$.

4. Conclusion

This paper considers a canonical bargaining game between two political leaders over a policy issue while public sentiment leans toward one leader or the other. Under complete information, the bargaining game is the same as in Binmore et al. (1989). We introduce a simple, easily interpretable information perturbation ε that the leaders are uninformed about the public sentiment. While under complete information, a leader concedes immediately whenever she learns that the sentiment leans the other way, she may not do so if her opponent could be uninformed. She could exaggerate how her constituents would react to her concession. We construct a war of attrition equilibrium in the spirit of AG.[12]

We show that in equilibrium, one of the leaders will concede immediately after learning that the sentiment leans the other way, as under complete information. However, the other leader, namely, the strong bargainer, can continue exaggerating

[12] We show that in our setup, commitment arises endogenously, making this equilibrium unique (as in AG). We provide a variation of our model to illustrate how other equilibria can arise (as in the early bargaining literature).

for some time even when she knows that the sentiment leans the other way. How long she continues exaggerating depends on the relative bargaining strength. If the strengths are moderately unbalanced, she concedes immediately with positive probability, but if the strengths are extremely unbalanced, she never concedes immediately.

Given a bargaining environment, as $\varepsilon \to (0,0)$, the bargaining strengths become almost balanced, and so she concedes immediately with a probability close to 1. In other words, as in the case under complete information, she almost never exaggerates. However, given ε (however small), we can find bargaining environments in which the strengths are extremely unbalanced. Therefore, in sharp contrast to the outcome under complete information, when the sentiment leans toward the weak bargainer, the strong bargainer never concedes immediately. Thus, this paper provides a simple natural bargaining setup where a small probability ε of being uninformed can make a large difference.

Appendix

Proof of Proposition 1

It follows from Lemma 1 that in equilibrium there are two phases. In phase I (if it exists), the I' type of the strong bargainer and the $\mathcal{U}$ type of the weak bargainer randomize. In phase II, the $\mathcal{U}$ type from both sides randomize. The types who randomize must be indifferent between conceding and exaggerating for a little longer.

Step I: Indifference Conditions

[I A] Indifference of $\omega_i = I'$:

Conditional on no agreement until time t, the $\omega_i = I'$ is indifferent between conceding now and conceding after Δ time if

$$\frac{G_j^{I'}(t+\Delta) - G_j^{I'}(t)}{1 - G_j^{I'}(t)} \cdot x_i + \left(1 - \frac{G_j^{I'}(t+\Delta) - G_j^{I'}(t)}{1 - G_j^{I'}(t)}\right) e^{-r_i \Delta}(1 - x_j) = (1 - x_j).$$

When $\Delta \to 0$, the indifference condition boils down to

$$\frac{\frac{dG_j^{I'}(t)}{dt}}{1 - G_j^{I'}(t)} = r_i \frac{(1 - x_j)}{x_i - (1 - x_j)} = \frac{1}{\lambda_j}. \tag{A.1}$$

Notice that the hazard rate is constant as in AG. Solving differential equation (A.1) with some starting point t_0, we get

$$G_j^{I'}(t) = 1 - \left(1 - G_j^{I'}(t_0)\right) \exp\left(-\frac{1}{\lambda_j}(t - t_0)\right). \tag{A.2}$$

[I B] Indifference of $\omega_i = \mathcal{U}$:

While $\omega_i = I'$ knows that the state favors the opponent, $\omega_i = \mathcal{U}$ is uncertain about the state. After seeing no concession from leader j, she updates her belief that the state favors the opponent with probability

$$\pi_j(t) = \frac{\pi_j\left[(1-\varepsilon_j) + \varepsilon_j(1 - F_j^{\mathcal{U}}(t))\right]}{1 - G_j^{\mathcal{U}}(t)} = \pi_j + \frac{\pi_j \pi_i (1-\varepsilon_j) F_j^{I'}(t)}{1 - G_j^{\mathcal{U}}(t)}. \tag{A.3}$$

Unlike the I' type, the $\mathcal{U}$ type gets $\pi_j(t)(1-x_j)$ if she concedes now and $\pi_j(t+\Delta)(1-x_j)$ if she concedes after time Δ. Accordingly, the $\mathcal{U}$ type is indifferent if

$$\frac{G_j^{\mathcal{U}}(t+\Delta) - G_j^{\mathcal{U}}(t)}{1 - G_j^{\mathcal{U}}(t)} \cdot x_i + \left(1 - \frac{G_j^{\mathcal{U}}(t+\Delta) - G_j^{\mathcal{U}}(t)}{1 - G_j^{\mathcal{U}}(t)}\right) e^{-r_i\Delta} \pi_j(t+\Delta)(1-x_j) = \pi_j(t)(1-x_j).$$

Rearranging and taking $\Delta \to 0$, we get

$$\lim_{\Delta \to 0} \frac{\frac{G_j^{\mathcal{U}}(t+\Delta) - G_j^{\mathcal{U}}(t)}{\Delta}\left(x_i - e^{-r_i\Delta}\pi_j(t+\Delta)(1-x_j)\right)}{1 - G_j^{\mathcal{U}}(t)}$$

$$= \lim_{\Delta \to 0}(1-x_j)\left[\frac{\pi_j(t) - e^{-r_i\Delta}\pi_j(t)}{\Delta} - e^{-r_i\Delta}\frac{\pi_j(t+\Delta) - \pi_j(t)}{\Delta}\right],$$

$$\frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1 - G_j^{\mathcal{U}}(t)} = \frac{r_i \pi_j(t)(1-x_j) - \pi_j'(t)(1-x_j)}{x_i - \pi_j(t)(1-x_j)} = \frac{1}{\lambda_j(t)}. \tag{A.4}$$

Since $\pi_j'(t) \geq 0$ and $\pi_j(t) \leq 1$, $\frac{1}{\lambda_j} > \frac{1}{\lambda_j(t)}$. Moreover, since $\mathcal{U}$ is always more optimistic than I' that the opponent may concede in the next instance (See equations (1) and (2)), the LHS in equation (A.1) is lower than the LHS in (A.4). Thus, when the I' type is indifferent between conceding and exaggerating, the $\mathcal{U}$ type strictly prefers exaggerating (consistent with (P1) in lemma 1).

Recall that if $\omega_i = \mathcal{U}$ randomizes in the first phase, then it is the I' opponent who randomizes, while if $\omega_i = \mathcal{U}$ randomizes in the second phase, then it is the $\mathcal{U}$ opponent who randomizes. Thus, depending on the phase, $\omega_i = \mathcal{U}$ updates her belief $\pi_j(t)$ differently. Next, we look into these two cases separately.

[I B.1] I' opponent randomizes:

Consider the case where the opponent of type $\omega_j = I'$ randomizes. Then, $\omega_j = \mathcal{U}$ has not started conceding yet; that is, $F_j^{\mathcal{U}}(t) = 0$. Accordingly, from equation (A.3), we get $\pi_j(t) = \frac{\pi_j}{1 - G_j^{\mathcal{U}}(t)}$. Differentiating w.r.t t, we get $\pi_j'(t) = \pi_j \frac{dG_j^{\mathcal{U}}(t)}{dt} / \left(1 - G_j^{\mathcal{U}}(t)\right)^2$.

Substituting $\pi_j(t)$ and $\pi'_j(t)$ in the equation (A.4) and then simplifying, we get

$$\frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1-G_j^{\mathcal{U}}(t)} = \frac{r_i \frac{\pi_j(1-x_j)}{1-G_j^{\mathcal{U}}(t)} - \frac{\pi_j \frac{dG_j^{\mathcal{U}}(t)}{dt}(1-x_j)}{\left(1-G_j^{\mathcal{U}}(t)\right)^2}}{x_i - \frac{\pi_j(1-x_j)}{1-G_j^{\mathcal{U}}(t)}}$$

$$= \frac{r_i\pi_j(1-x_j)}{x_i(1-G_j^{\mathcal{U}}(t)) - \pi_j(1-x_j)} - \frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1-G_j^{\mathcal{U}}(t)}\left[\frac{\pi_j(1-x_j)}{x_i(1-G_j^{\mathcal{U}}(t)) - \pi_j(1-x_j)}\right]$$

$$\Rightarrow \frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1-G_j^{\mathcal{U}}(t)}\left[1+\frac{\pi_j(1-x_j)}{x_i(1-G_j^{\mathcal{U}}(t)) - \pi_j(1-x_j)}\right] = \frac{r_i\pi_j(1-x_j)}{x_i(1-G_j^{\mathcal{U}}(t)) - \pi_j(1-x_j)}$$

$$\Rightarrow \frac{dG_j^{\mathcal{U}}(t)}{dt} = \frac{r_i(1-x_j)\pi_j}{x_i} = \frac{1}{\zeta_j}.$$

Solving this differential equation with some starting point t_0, we get

$$G_j^{\mathcal{U}}(t) = G_j^{\mathcal{U}}(t_0) + \frac{1}{\zeta_j}(t-t_0). \tag{A.5}$$

[I B.2] $\mathcal{U}$ opponent randomizes:

Consider the case where $\omega_j = \mathcal{U}$ randomizes. Then, $\omega_j = I'$ has already conceded; that is, $F_j^{I'}(t) = 1$. Accordingly, from equation (A.3), we have $\pi_j(t) = \pi_j + \frac{\pi_j\pi_i(1-\varepsilon_j)}{1-G_j^{\mathcal{U}}(t)}$. Differentiating w.r.t t, we get $\pi'_j(t) = \pi_j\pi_i(1-\varepsilon_j)\frac{dG_j^{\mathcal{U}}(t)}{dt} / \left(1-G_j^{\mathcal{U}}(t)\right)^2$. Substituting $\pi_j(t)$ and $\pi'_j(t)$ in the equation (A.4) and then simplifying, we get

$$\frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1-G_j^{\mathcal{U}}(t)} = \frac{r_i(1-x_j)\frac{\pi_j\left[1+\pi_i(1-\varepsilon_j)-G_j^{\mathcal{U}}(t)\right]}{1-G_j^{\mathcal{U}}(t)} - \frac{\pi_j\pi_i(1-\varepsilon_j)\frac{dG_j^{\mathcal{U}}(t)}{dt}(1-x_j)}{\left(1-G_j^{\mathcal{U}}(t)\right)^2}}{x_i - (1-x_j)\frac{\pi_j\left[1+\pi_i(1-\varepsilon_j)-G_j^{\mathcal{U}}(t)\right]}{1-G_j^{\mathcal{U}}(t)}}$$

$$= \frac{r_i(1-x_j)\pi_j\left[1+\pi_i(1-\varepsilon_j)-G_j^{\mathcal{U}}(t)\right]}{x_i(1-G_j^{\mathcal{U}}(t)) - (1-x_j)\pi_j\left[1+\pi_i(1-\varepsilon_j)-G_j^{\mathcal{U}}(t)\right]}$$
$$- \frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1-G_j^{\mathcal{U}}(t)}\left[\frac{\pi_j\pi_i(1-\varepsilon_j)(1-x_j)}{x_i(1-G_j^{\mathcal{U}}(t)) - (1-x_j)\pi_j\left[1+\pi_i(1-\varepsilon_j)-G_j^{\mathcal{U}}(t)\right]}\right]$$

$$\Rightarrow \frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1-G_j^{\mathcal{U}}(t)}\left[1+\frac{\pi_j\pi_i(1-\varepsilon_j)(1-x_j)}{x_i-(1-x_j)\pi_j(1+\pi_i(1-\varepsilon_j)) - (x_i-(1-x_j)\pi_j)G_j^{\mathcal{U}}(t)}\right]$$
$$= \frac{r_i(1-x_j)\pi_j\left[1+\pi_i(1-\varepsilon_j)-G_j^{\mathcal{U}}(t)\right]}{x_i-(1-x_j)\pi_j(1+\pi_i(1-\varepsilon_j)) - (x_i-(1-x_j)\pi_j)G_j^{\mathcal{U}}(t)}$$

$$\Rightarrow \frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1 - G_j^{\mathcal{U}}(t)} =$$

$$\frac{r_i(1-x_j)\pi_j\left[1+\pi_i(1-\varepsilon_j) - G_j^{\mathcal{U}}(t)\right]}{x_i - (1-x_j)\pi_j(1+\pi_i(1-\varepsilon_j)) + \pi_j\pi_i(1-\varepsilon_j)(1-x_j) - (x_i - (1-x_j)\pi_j)G_j^{\mathcal{U}}(t)}$$

$$\Rightarrow \frac{\frac{dG_j^{\mathcal{U}}(t)}{dt}}{1+\pi_i(1-\varepsilon_j) - G_j^{\mathcal{U}}(t)} = \frac{r_i(1-x_j)\pi_j}{x_i - (1-x_j)\pi_j} = \frac{1}{\eta_j}.$$

Solving this differential equation with starting some point t_0, we get

$$\int_{t_0}^{t} \frac{dG_j^{\mathcal{U}}}{1+\pi_i(1-\varepsilon_j) - G_j^{\mathcal{U}}} = \frac{1}{\eta_j}\int_{t_0}^{t} dt$$

$$\text{or,} \quad -\ln\left(1+\pi_i(1-\varepsilon_j) - G_j^{\mathcal{U}}(t)\right) + \ln\left(1+\pi_i(1-\varepsilon_j) - G_j^{\mathcal{U}}(t_0)\right) = \frac{1}{\eta_j}(t-t_0)$$

$$\text{or,} \quad \left(1+\pi_i(1-\varepsilon_j) - G_j^{\mathcal{U}}(t)\right) = \left(1+\pi_i(1-\varepsilon_j) - G_j^{\mathcal{U}}(t_0)\right)\exp\left(-\frac{1}{\eta_j}(t-t_0)\right)$$

This gives us

$$G_j^{\mathcal{U}}(t) = 1+\pi_i(1-\varepsilon_j) - \left(1+\pi_i(1-\varepsilon_j) - G_j^{\mathcal{U}}(t_0)\right)\exp\left(-\frac{1}{\eta_j}(t-t_0)\right). \tag{A.6}$$

Step II: Balanced Strengths

Recall that $T_i[I'] \cdot T_j[I'] = 0$ (Lemma 1 (P4)). Let us fist consider the case where $T_i[I'] = T_j[I'] = 0$, that is, a leader never exaggerates when she learns that the public sentiment leans the other way. This means $F_i^{I'}(t) = \mathbb{1}(t \geq 0)$ for all $i \in N$. We will look for (θ, ε) such that this will hold true.

Note that since $F_i^{I'}(0) \cdot G_j^{I'}(0) = 0$ (Lemma 1 (P3)). Since $F_i^{I'}(0) > 0$ for all $i \in N$, we must have $G_j^{I'}(0) = \varepsilon_j F_j^{\mathcal{U}}(0) > 0$ for all $j \in N$. Therefore, $F_j^{\mathcal{U}}(0) = 0$ for all $j \in N$. Accordingly, $G_j^{\mathcal{U}}(0) = \pi_i(1-\varepsilon_j)$ and $G_i^{\mathcal{U}}(0) = \pi_j(1-\varepsilon_i)$. Only the uninformed types randomizes after $t = 0$. Starting from the initial time $t_0 = 0$, the uninformed types on both sides will update their beliefs that the opponent will exaggerate until time t according to (A.6). This gives us

$$G_j^{\mathcal{U}}(t) = 1+\pi_i(1-\varepsilon_j) - \exp\left(-\frac{1}{\eta_j}t\right).$$

This means that leader j may continue exaggerating until time T_j where

$$G_j^{\mathcal{U}}(T_j) = 1+\pi_i(1-\varepsilon_j) - \exp\left(-\frac{1}{\eta_j}T_j\right) = 1-\pi_j(1-\varepsilon_j).$$

Solving this, we get

$$\exp(T_j) = (1-\varepsilon_j)^{-\eta_j}.$$

Since in equilibrium, $T_i = T_j$ (Lemma 1 (P2)), (θ, ε) must be such that

$$B_j^1(\theta, \varepsilon) = \frac{(1-\varepsilon_i)^{-\eta_i}}{(1-\varepsilon_j)^{-\eta_j}} = 1. \tag{A.7}$$

Notice that for any (θ, ε), $B_L^1 = 1/B_R^1$. When (θ, ε) is such that $B_L^1 = B_R^1 = 1$, we say that the bargaining strengths are balanced.

Step III: Unbalanced Strengths

Suppose (θ, ε) is such that $B_j^1(\theta, \varepsilon) < 1$. Then, if both the I' types do not exaggerate, we have $T_i < T_j$ and Lemma 1 (P2) does not hold. We say that the bargaining strengths are unbalanced — leader j is the weak bargainer and leader i is the strong bargainer. To makes $T_i = T_j$, it must be that $F_i^{I'}(0) < 1$. This is because Lemma 1 (P4) means $F_i^{I'}(0) = 1$ for at least one $i \in N$, and $F_j^{I'}(0) < 1$ will widen the gap between T_i and T_j. Since $F_i^{I'}(0) < 1$, in the first phase $[0, T_i[I']]$, $\omega_i = I'$ and $\omega_j = \mathcal{U}$ randomize, and in the second phase $[T_i[I'], T]$, $\omega_i = \mathcal{U}$ and $\omega_j = \mathcal{U}$ randomize. Since $F_j^{I'}(0) = 1$, we must have $F_i^{\mathcal{U}}(0) = 0$ (Lemma 1 (P3)). Notice that $F_j^{\mathcal{U}}(0)$ can be positive. However, it follows from Lemma 1 (P3) that $F_i^{I'}(0) \cdot F_j^{\mathcal{U}}(0) = 0$.

[III A] Belief about leader i:

$\omega_j = \mathcal{U}$ believes that leader i will concede immediately with probability $G_i^{\mathcal{U}}(0) = \pi_j(1-\varepsilon_i)F_i^{I'}(0)$, and starting from $t_0 = 0$, in phase I, she updates her belief that the opponent will exaggerate until time t according to (A.5) (interchange i and j). This gives us

$$G_i^{\mathcal{U}}(t) = \pi_j(1-\varepsilon_i)F_i^{I'}(0) + \frac{1}{\zeta_i}t. \tag{A.8}$$

By time $T_i[I']$, the I' type opponent finish exaggerating. Therefore,

$$G_i^{\mathcal{U}}(T_i[I']) = \pi_j(1-\varepsilon_i)F_i^{I'}(0) + \frac{1}{\zeta_i}T_i[I'] = \pi_j(1-\varepsilon_i).$$

Solving this, we get

$$T_i[I'] = \zeta_i \pi_j (1-\varepsilon_i)(1-F_i^{I'}(0)) \tag{A.9}$$

Starting at time $t_0 = T_i[I']$, and given $G_i^{\mathcal{U}}(T_i[I']) = \pi_j(1-\varepsilon_i)$, $\omega_j = \mathcal{U}$ update her belief in phase II that the opponent will exaggerate until time t according to (A.6) (interchange i and j). This gives us

$$G_i^{\mathcal{U}}(t) = 1 + \pi_j(1-\varepsilon_i) - \exp\left(-\frac{1}{\eta_i}(t - T_i[I'])\right) \tag{A.10}$$

This means that leader i will keep exaggerating until time T_i, where

$$G_i^{\mathcal{U}}(T_i) = 1 + \pi_j(1-\varepsilon_i) - \exp\left(-\frac{1}{\eta_i}(T_i - T_i[I'])\right) = 1 - \pi_i(1-\varepsilon_i).$$

Substituting (A.9), we get

$$\exp\left(-\frac{1}{\eta_i}(T_i - \zeta_i\pi_j(1-\varepsilon_i)(1-F_i^{I'}(0)))\right) = (1-\varepsilon_i).$$

This gives us

$$T_i = \ln((1-\varepsilon_i)^{-\eta_i}) + \zeta_i\pi_j(1-\varepsilon_i)(1-F_i^{I'}(0)). \tag{A.11}$$

Note that if $F_i^{I'}(0) = 1$, then $\exp(T_i) = (1-\varepsilon_i)^{-\eta_i}$ as in the balanced strength situation. As $F_i^{I'}(0)$ falls, it take leader i more time to finish building her reputation.

[III B] Belief about leader j:

$\omega_i = I'$ believes that leader j will concede immediately with probability $G_j^{I'}(0) = \varepsilon_j F_j^{\mathcal{U}}(0)$, and starting from $t_0 = 0$, in phase I, she updates her belief that the opponent will exaggerate until time t according to (A.2). This gives us

$$G_j^{I'}(t) = 1 - (1-\varepsilon_j F_j^{\mathcal{U}}(0))\exp\left(-\frac{1}{\lambda_j}t\right). \tag{A.12}$$

Since, $F_j^{I'}(0) = 1$, we have

$$G_j^{\mathcal{U}}(t) = \pi_i(1-\varepsilon_j) + 1 - (1-\varepsilon_j F_j^{\mathcal{U}}(0))\exp\left(-\frac{1}{\lambda_j}t\right). \tag{A.13}$$

Therefore, $\omega_i = \mathcal{U}$ believes that leader j will exaggerate until time $t = T_i[I']$ with probability

$$G_j^{\mathcal{U}}(T_i[I']) = \pi_i(1-\varepsilon_j) + 1 - (1-\varepsilon_j F_j^{\mathcal{U}}(0))\exp\left(-\frac{1}{\lambda_j}T_i[I']\right).$$

Starting at $t_0 = T_i[I']$, and given $G_j^{\mathcal{U}}(T_i[I'])$, $\omega_i = \mathcal{U}$ believes that in phase II leader j will exaggerate until time t according to (A.6). This gives us

$$G_j^{\mathcal{U}}(t) = 1 + \pi_i(1-\varepsilon_j) - (1-\varepsilon_j F_j^{\mathcal{U}}(0))\exp\left(-\frac{1}{\lambda_j}T_i[I']\right)\exp\left(-\frac{1}{\eta_j}(t - T_i[I'])\right). \tag{A.14}$$

This means that leader j will continue exaggerating until time T_j, where

$$G_j^{\mathcal{U}}(T_j) = 1 + \pi_i(1-\varepsilon_j) - (1-\varepsilon_j F_j^{\mathcal{U}}(0))\exp\left(-\frac{1}{\lambda_j}T_i[I']\right)\exp\left(-\frac{1}{\eta_j}(T_j - T_i[I'])\right)$$
$$= 1 - \pi_j(1-\varepsilon_j)$$

This simplifies to

$$\exp\left(-\frac{1}{\eta_j}(T_j - T_i[I'])\right) = \frac{(1-\varepsilon_j)\exp\left(T_i[I']\right)^{\frac{1}{\lambda_j}}}{(1-\varepsilon_j F_j^{\mathcal{U}}(0))}$$

$$\implies \exp(T_j) = (1-\varepsilon_j)^{-\eta_j} \cdot (1-\varepsilon_j F_j^{\mathcal{U}}(0))^{\eta_j} \cdot \exp(T_i[I'])^{1-\frac{\eta_j}{\lambda_j}}$$

Substituting (A.9), we get

$$\begin{aligned} T_j &= \ln((1-\varepsilon_j)^{-\eta_j}) + \ln((1-\varepsilon_j F_j^{\mathcal{U}}(0))^{\eta_j}) \\ &+ \left(\zeta_i \pi_j (1-\varepsilon_i)(1-F_i^{I'}(0))\right)\left(1-\frac{\eta_j}{\lambda_j}\right). \end{aligned} \tag{A.15}$$

Note that as $F_j^{\mathcal{U}}(0)$ increases T_j falls. Also, since $\lambda_j < \eta_j$, as $F_i^{I'}(0)$ increases T_j increases.

Since in equilibrium $T_i = T_j$ (Lemma 1 (P2)), it follows from equations (A.11) and (A.15) that $F_i^{I'}(0)$ and $F_j^{\mathcal{U}}(0)$ must be such that

$$(1-\varepsilon_i)^{-\eta_i} \cdot \exp\left(\frac{\eta_j}{\lambda_j}\zeta_i \pi_j (1-\varepsilon_i)(1-F_i^{I'}(0))\right) = (1-\varepsilon_j)^{-\eta_j}(1-\varepsilon_j F_j^{\mathcal{U}}(0))^{\eta_j} \tag{A.16}$$

It follows from Lemma 1 (P3) that $F_i^{I'}(0) \cdot F_j^{\mathcal{U}}(0) = 0$. If in equilibrium, $F_i^{I'}(0) = F_j^{\mathcal{U}}(0) = 0$, then it must be that

$$(1-\varepsilon_i)^{-\eta_i} \cdot \exp\left(\frac{\eta_j}{\lambda_j}\zeta_i \pi_j (1-\varepsilon_i)\right) = (1-\varepsilon_j)^{-\eta_j}$$

$$\text{or, } B_j^2(\theta, \varepsilon) = \frac{(1-\varepsilon_i)^{-\eta_i}}{(1-\varepsilon_j)^{-\eta_j}} \cdot \exp\left(\frac{\eta_j}{\lambda_j}\zeta_i \pi_j (1-\varepsilon_i)\right) = 1. \tag{A.17}$$

[III C] Moderately Unbalanced Strengths:

Leader j is weak enough ($1 > B_j^1$) that leader i may exaggerate even when she learns that the public sentiment leans the other way. However, leader j is not so weak ($B_j^2 > 1$) that leader i never concedes immediately when she learns that the public sentiment is unfavorable . That is, $F_i^{I'}(0) \in (0,1)$. This implies $F_j^{\mathcal{U}}(0) = 0$. Therefore, solving (A.16), we get

$$F_i^{I'}(0) = 1 + \left(\frac{\ln B_j^1}{\ln B_j^2 - \ln B_j^1}\right).$$

Consider the strong bargainer i. In the first phase $[0, T_i[I']]$, $\omega_i = \mathcal{U}$ always exaggerates ($F_i^{\mathcal{U}}(t) = 0$), and $\omega_i = I'$ exaggerates exactly so much that $G_i^{\mathcal{U}}(t)$ is as in equation (A.8) ($\omega_j = \mathcal{U}$ is indifferent). Recall that $G_i^{\mathcal{U}}(t) = \varepsilon_i F_i^{\mathcal{U}}(t) + \pi_j(1-\varepsilon_i)F_i^{I'}(t)$ (See equation (2)). Substituting ($F_i^{\mathcal{U}}(t) = 0$), we get

$$F_i^{I'}(t) = \frac{G_j^{\mathcal{U}}(t)}{\pi_j(1-\varepsilon_i)} = F_i^{I'}(0) + t/(\pi_j(1-\varepsilon_i)\zeta_i).$$

Since $\omega_i = I'$ finish exaggerating by time $T_i[I']$, $F_i^{I'}(T_i[I']) = 1$. In the second phase $[T_i[I'], T]$, $\omega_i = \mathcal{U}$ exaggerates exactly so much that $G_i^{\mathcal{U}}(t)$ is as in equation

(A.10) ($\omega_j = \mathcal{U}$ is indifferent). Substituting $F_i^{I'}(t) = 1$ in $G_i^{\mathcal{U}}(t) = \varepsilon_i F_i^{\mathcal{U}}(t) + \pi_j(1-\varepsilon_i)F_i^{I'}(t)$, we get

$$F_i^{\mathcal{U}}(t) = \frac{1}{\varepsilon_i}\left(G_i^{\mathcal{U}}(t) - \pi_j(1-\varepsilon_i)\right) = \frac{1}{\varepsilon_i}\left(1 - \exp\left(-\frac{(t - T_i[I'])}{\eta_i}\right)\right).$$

Next, consider the weak bargainer j. $\omega_i = I'$ never exaggerates — that is,

$$F_j^{I'}(t) = \mathbb{1}(t \geq 0).$$

In the first phase $[0, T_i[I']]$, $\omega_j = \mathcal{U}$ exaggerates exactly so much that $G_j^{I'}(t)$ is as in equation (A.12) ($\omega_i = I'$ is indifferent). Recall that $G_j^{I'}(t) = \varepsilon_j F_j^{\mathcal{U}}(t)$ (See equation (1)). Therefore, for $t \in [0, T_i[I']]$,

$$F_j^{\mathcal{U}}(t) = \frac{G_j^{I'}(t)}{\varepsilon_j} = \frac{1}{\varepsilon_j}\left(1 - \exp\left(-\frac{t}{\lambda_j}\right)\right)$$

In the second phase $[T_i[I'], T]$, $\omega_j = \mathcal{U}$ exaggerates exactly so much that $G_j^{\mathcal{U}}(t)$ is as in equation (A.14) ($\omega_i = \mathcal{U}$ is indifferent). Substituting $F_j^{I'}(t) = 1$ in $G_j^{\mathcal{U}}(t) = \varepsilon_j F_j^{\mathcal{U}}(t) + \pi_i(1-\varepsilon_j)F_j^{I'}(t)$ (See equation (2)), we get that for $t \in [T_i[I'], T]$,

$$F_j^{\mathcal{U}}(t) = \frac{G_j^{\mathcal{U}}(t) - \pi_i(1-\varepsilon_j)}{\varepsilon_j} = \frac{1}{\varepsilon_j}\left(1 - \exp\left(-\frac{T_i[I']}{\lambda_j} - \frac{(t - T_i[I'])}{\eta_j}\right)\right).$$

[III D] Extremely Unbalanced Strengths:

Leader j is so weak ($1 > B_j^2$) that leader i never concedes immediately when she learns that the public sentiment leans the other way. That is, $F_i^{I'}(0) = 0$. Since $1 > B_j^2$, if $F_i^{I'}(0) = F_j^{\mathcal{U}}(0) = 0$, we have $T_j > T_i$. which violates Lemma 1 (P2). To make $T_j = T_i$, $F_j^{\mathcal{U}}(0)$ must be such that equation (A.16) holds. Solving this, we get

$$F_j^{\mathcal{U}}(0) = \frac{1}{\varepsilon_j}\left(1 - (B_j^2)^{1/\eta_j}\right).$$

Consider the strong bargainer i. In the first phase $[0, T_i[I']]$, $\omega_i = \mathcal{U}$ always exaggerates ($F_i^{\mathcal{U}}(t) = 0$), and $\omega_i = I'$ exaggerates exactly so much that $G_i^{\mathcal{U}}(t)$ is as in equation (A.8), which gives us

$$F_i^{I'}(t) = \frac{G_j^{\mathcal{U}}(t)}{\pi_j(1-\varepsilon_i)} = t/(\pi_j(1-\varepsilon_i)\zeta_i).$$

In the second phase $[T_i[I'], T]$, $\omega_i = \mathcal{U}$ exaggerates exactly so much that $G_i^{\mathcal{U}}(t)$ is as in equation (A.10), which gives us

$$F_i^{\mathcal{U}}(t) = \frac{1}{\varepsilon_i}\left(G_i^{\mathcal{U}}(t) - \pi_j(1-\varepsilon_i)\right) = \frac{1}{\varepsilon_i}\left(1 - \exp\left(-\frac{(t - T_i[I'])}{\eta_i}\right)\right).$$

Next, consider the weak bargainer j. $\omega_i = I'$ never exaggerates — that is,

$$F_j^{I'}(t) = \mathbb{1}(t \geq 0).$$

In the first phase $[0, T_i[I']]$, $\omega_j = \mathcal{U}$ exaggerates exactly so much that $G_j^{I'}(t)$ is as in equation (A.12), which gives us

$$F_j^{\mathcal{U}}(t) = \frac{G_j^{I'}(t)}{\varepsilon_j} = \frac{1}{\varepsilon_j}\left(1 - (1 - \varepsilon_j F_j^{\mathcal{U}}(0)) \exp\left(-\frac{t}{\lambda_j}\right)\right)$$

for $t \in [0, T_i[I']]$. In the second phase $[T_i[I'], T]$, $\omega_j = \mathcal{U}$ exaggerates exactly so much that $G_j^{\mathcal{U}}(t)$ is as in equation (A.14), which give us

$$F_j^{\mathcal{U}}(t) = \frac{1}{\varepsilon_j}\left(1 - (1 - \varepsilon_j F_j^{\mathcal{U}}(0)) \exp\left(-\frac{T_i[I']}{\lambda_j} - \frac{(t - T_i[I'])}{\eta_j}\right)\right)$$

for $t \in [T_i[I'], T]$. $\square$

Proof of Proposition 2

Step I:

Given θ, for sufficiently large n, the bargaining strengths cannot be extremely unbalanced.

Consider $\varepsilon_i \in (0, \tilde{\varepsilon})$ for all $i \in N$. Suppose that leader L is the strong bargainer. Then, $B_L^1(\theta, \varepsilon) = (1 - \varepsilon_R)^{-\eta_R}/(1 - \varepsilon_L)^{-\eta_L} \leq (1 - \tilde{\varepsilon})^{-\eta_R}$. The inequality follows from setting $\varepsilon_L = 0$ and $\varepsilon_R = \tilde{\varepsilon}$. Similarly, if leader R is the strong bargainer, then $B_R^1(\theta, \varepsilon) = (1 - \varepsilon_L)^{-\eta_L}/(1 - \varepsilon_R)^{-\eta_R} \leq (1 - \tilde{\varepsilon})^{-\eta_L}$. Therefore,

$$\max_{i \in \{L,R\}} B_i^1(\theta, \varepsilon) \leq (1 - \tilde{\varepsilon})^{-\max\{\eta_L, \eta_R\}}.$$

Suppose that L is very weak — that is, $1 \geq B_L^2(\theta, \varepsilon)$. Then,

$$1 \geq \frac{(1 - \varepsilon_R)^{-\eta_R} \cdot \chi_L}{(1 - \varepsilon_L)^{-\eta_L}},$$

$$\text{or,} \quad \left(\frac{(1 - \varepsilon_R)^{-\eta_R}}{(1 - \varepsilon_L)^{-\eta_L}}\right)^{\frac{1}{(1-\varepsilon_L)}} \geq \exp\left(\frac{\eta_R}{\lambda_R} \cdot \zeta_L \pi_R\right).$$

This is possible only if

$$(1 - \tilde{\varepsilon})^{-\frac{1}{1-\tilde{\varepsilon}}} \geq \exp\left(\frac{1}{\max\{\eta_L, \eta_R\}} \cdot \frac{\eta_R}{\lambda_R} \cdot \zeta_L \pi_R\right).$$

Similarly, R is very weak — that is, $1 \geq B_R^2(\theta, \varepsilon)$ only if

$$(1 - \tilde{\varepsilon})^{-\frac{1}{1-\tilde{\varepsilon}}} \geq \exp\left(\frac{1}{\max\{\eta_L, \eta_R\}} \cdot \frac{\eta_L}{\lambda_L} \cdot \zeta_R \pi_L\right).$$

Given the bargaining environment θ, let us define

$$k^1 := \frac{1}{\max\{\eta_L, \eta_R\}} \cdot \min\left\{\frac{\eta_R \zeta_L \pi_R}{\lambda_R}, \frac{\eta_L \zeta_R \pi_L}{\lambda_L}\right\}$$

$$\text{and} \quad \phi(\tilde{\varepsilon}) := -\frac{\ln(1-\tilde{\varepsilon})}{1-\tilde{\varepsilon}}.$$

Therefore, given θ, the bargaining strengths are extremely unbalanced only if

$$\phi(\tilde{\varepsilon}) \geq k^1.$$

Note that for a given θ, k^1 is a positive constant, while $\phi(\tilde{\varepsilon})$ is increasing in $\tilde{\varepsilon}$, and $\phi(\tilde{\varepsilon}) \to 0$ as $\tilde{\varepsilon} \to 0$. Define ε^1 such that $\phi(\varepsilon^1) = k^1$. Then, for $\tilde{\varepsilon} < \varepsilon^1$, $\phi(\tilde{\varepsilon}) < k^1$. There exists n^1 such that when $n > n^1$, $\varepsilon_{in} < \tilde{\varepsilon} < \varepsilon^1$ for all $i \in \{L, R\}$, which implies the bargaining strengths cannot be extremely unbalanced.

Step II:

Assume that n is sufficiently large such that the bargaining strengths are not extremely unbalanced. Given (θ, ε_n), let X_{in} be the random variable that captures when, in equilibrium, leader i stops exaggerating after she learns that the public sentiment leans the other way. We have $Pr(X_{in} \leq t) = F_{in}^{I'}(t)$. We show that the sequence of random variable X_n converges to the constant 0 — that is,

$$\lim_{n\to\infty} F_{in}^{I'}(t) = \mathbb{1}(t \geq 0).$$

Consider $\varepsilon_i < \tilde{\varepsilon}$ for all $i \in \{L, R\}$. If leader i is the weak bargainer, then $F_i^{I'}(t) = \mathbb{1}(t \geq 0)$. Therefore, the above claim is trivially true. Suppose that leader i is the strong bargainer. Then we have

$$|1 - F_{in}^{I'}(t)| \leq -\left(\frac{\ln B_j^1}{\ln B_j^2 - \ln B_j^1}\right)$$

$$= \frac{1}{\pi_j(1-\varepsilon_i)}\left(\frac{\lambda_j}{\zeta_i \eta_j}\right)\left[\eta_i \ln(1-\varepsilon_i) - \eta_j \ln(1-\varepsilon_j)\right]$$

$$\leq \frac{1}{\pi_j}\left(\frac{\lambda_j}{\zeta_i}\right)\left[-\frac{\ln(1-\tilde{\varepsilon})}{1-\tilde{\varepsilon}}\right].$$

The last inequality follows since $\frac{1}{1-\varepsilon_i} \leq \frac{1}{1-\tilde{\varepsilon}}$ and $\eta_i \ln(1-\varepsilon_i) - \eta_j \ln(1-\varepsilon_j) \leq -\eta_j \ln(1-\tilde{\varepsilon})$. Define

$$k^2 := \max\{\frac{\lambda_R}{\pi_R \zeta_L}, \frac{\lambda_L}{\pi_L \zeta_R}\}.$$

Recall that $\phi(\tilde{\varepsilon})$ in increasing, and as $\tilde{\varepsilon} \to 0$, $k^2\phi(\tilde{\varepsilon}) \to 0$. For any $\delta > 0$, define ε^2 such that $k^2\phi(\varepsilon^2) := \delta$. This means for $\tilde{\varepsilon} < \varepsilon^2$,

$$|1 - F_{in}^{I'}(t)| \leq k^2\phi(\tilde{\varepsilon}) < \delta.$$

Let n^2 be such that when $n > n^2$, $\varepsilon_{in} < \tilde{\varepsilon} < \varepsilon^2$ for all $i \in \{L, R\}$. Therefore, for any $\delta > 0$, when $n > \max\{n^1, n^2\}$, $\varepsilon_{in} < \tilde{\varepsilon} < \min\{\varepsilon^1, \varepsilon^2\}$ for all $i \in \{L, R\}$ and accordingly,

$$|1 - F_{in}^{I'}(t)| < \delta.$$

Step III:

Given θ, consider $n > \max\{n^1, n^2\}$ such that the probability that the I' type exaggerates is at most δ. We show that as $n \to \infty$, the probability that the leader will not immediately agree on the same policy position as they do when public sentiment is commonly known converges to 0.

Suppose that the state is $\mathscr{L}$, then the probability that the leaders will not immediately agree on policy x_L is at most $\varepsilon_{Rn} + (1-\varepsilon_{Rn})(1-F^{I'}_{Rn}(0)) \leq \varepsilon_{Rn} + (1-\varepsilon_{Rn})\delta$. As $n \to \infty$, this upper bound converges to 0. Similarly if the state is $\mathscr{R}$, the probability that the leaders will not immediately agree on policy $1-x_R$ is at most $\varepsilon_{Ln} + (1-\varepsilon_{Ln})(1-F^{I'}_{Ln}(0)) \leq \varepsilon_{Ln} + (1-\varepsilon_{Ln})\delta$. As $n \to \infty$, this upper bound converges to 0. □

Proof of Proposition 3

Suppose, without loss of generality, $\varepsilon_j \geq \varepsilon_i$. Consider a bargaining environment θ where x_j is close to 1. It follows from (η_j),(ζ_i),(λ_j) that $\lambda_j, \zeta_j, \eta_j \to \infty$ and $\lambda_i, \zeta_i, \eta_i$ converge to some finite values. Given continuity w.r.t x_j, we can find x_j close to 1 such that $\eta_j - \eta_i > 0$. Since $\varepsilon_j \geq \varepsilon_i$, and $\eta_j > \eta_i$, $B^1_j(\theta,\varepsilon) = \frac{(1-\varepsilon_i)^{-\eta_i}}{(1-\varepsilon_j)^{-\eta_j}} < 1$. That is, leader j is the weak bargainer. Leader j is a very weak bargainer when $1 > B^2_j(\theta,\varepsilon)$,

$$\text{or, } (1-\varepsilon_j)^{-\eta_j} > (1-\varepsilon_i)^{-\eta_i} \cdot \exp\left(\frac{\eta_j}{\lambda_j}\zeta_i\pi_j(1-\varepsilon_i)\right),$$

$$\text{or, } \frac{\eta_j}{\lambda_j}\zeta_i\pi_j < \frac{1}{1-\varepsilon_i}(\eta_i \ln(1-\varepsilon_i) - \eta_j \ln(1-\varepsilon_j)).$$

Rearranging this and using $\varepsilon_j \geq \varepsilon_i$, we get that if

$$\left(\frac{\eta_j}{\eta_j - \eta_i}\right)\frac{\zeta_i}{\lambda_j}\pi_j < \frac{-\ln(1-\varepsilon_i)}{1-\varepsilon_i} = \phi(\varepsilon_i),$$

then leader j is a very weak bargainer. Note that

$$\lim_{x_j \to 1}\frac{\eta_i}{\eta_j} = 0, \text{ and } \lim_{x_j \to 1}\frac{\zeta_i}{\lambda_j} = 0.$$

Thus, when $x_j \to 1$, the LHS converges to 0. Therefore, for any $\varepsilon_i > 0$, it follows from continuity that when x_j is sufficiently close to 1, the above inequality holds. Since leader j is a very weak bargainer, in equilibrium, leader i never concedes immediately when she learns that the sentiment leans the other way — that is $F^{I'}_i(0) = 0$. We can see from proposition 1 that for $t \in [0, T_i[I']]$, where $T_i[I'] = \zeta_i\pi_j(1-\varepsilon_i) > 0$,

$$F^{I'}_i(t) = \frac{r_j\pi_i(1-x_i)}{x_j\pi_j(1-\varepsilon_i)} \cdot t.$$

Suppose that leader L is the very weak bargainer. When $\omega = \mathscr{L}$, the probability that the leaders will immediately agree on x_L is 0. Analogously, when leader R is the very weak bargainer, and $\omega = \mathscr{R}$, the probability that the leader will immediately agree on $1 - x_R$ is 0. $\square$

Proof of Lemma 2

A pure strategy for leader i of type ω_i is a map $\sigma_i^{\omega_i}$ that specifies at any history h the action she takes. If at history h, leader i gets the chance to make an offer, she chooses a policy proposal $p \in [0,1]$; and if at history h, the opponent gets the chance to make an offer, and offers p, we say the history is (h,p), and she chooses whether to accept $(\mathcal{A})$ the current offer p or not.

Given the strategy of her opponent j (all types ω_j), leader i of type ω_i updates her belief at any history h about the state (public sentiment ω) and the opponent's type (ω_j). A strategy profile induces a distributions over the outcome (p,t), which determines the expected payoff. PBE requires that each type of each leader maximizes her expected payoff at any history, and the beliefs are Bayesian consistent on path. Let R_i be the set of strategies leader $i \in N$ may play in a PBE.

Let Z_i be the set of policy offers leader $i \in N$ rejects with positive probability after some history while playing a strategy in R_i, that is, $Z_i := \{p \in [0,1] | \exists h \in H, \sigma_i \in R_i : \sigma_i((h,p))(\mathcal{A}) < 1\}$. Let us define

$$\overline{p}_L := \sup Z_L \text{ and } \underline{p}_R := \inf Z_R.$$

Let Y_i be the set of offers any $i \in N$ accepts with positive probability after some history while playing a strategy in R_i, that is, $Y_i := \{p \in [0,1] | \exists h \in H, \sigma_i \in R_i : \sigma_i((h,p))(\mathcal{A}) > 0\}$. Let us define

$$\underline{p}_L := \inf Y_L \text{ and } \overline{p}_R := \sup Y_R.$$

This means that leader L will accept any policy $p \geq \overline{p}_L$ with probability 1 and reject any policy $p < \underline{p}_L$ with probability 1. Similarly, leader R will accept any policy $p \leq \underline{p}_R$ with probability 1 and reject any policy $p > \overline{p}_R$ with probability 1. Since, there is a type of leader L who never accepts an offer $p < x_L$, it must be that $\overline{p}_L \geq x_L$. Also, since there is a type of a leader R who never accepts an offer $p > 1 - x_R$, it must be that $\underline{p}_R \leq 1 - x_R$.

If leader L rejects an offer and makes a counter offer $\min\{1 - x_R, \underline{p}_R\}$, it will be accepted. On the other hand, if leader R rejects an offer and makes a counter offer $\underline{p}_L$, it will be definitely rejected. Therefore,

$$\underline{p}_L \geq e^{-r_L \Delta} \min\{1 - x_R, \underline{p}_R\} \text{ and } 1 - \underline{p}_R \leq e^{-r_R \Delta}(1 - \underline{p}_L). \tag{A.18}$$

Analogously, if leader R rejects an offer and makes a counter offer $\max\{x_L, \overline{p}_L\}$, it will be accepted. On the other hand, if leader L rejects an offer and makes a counter offer $\overline{p}_R$, it will be definitely rejected. Therefore,

$$1 - \overline{p}_R \geq e^{-r_R\Delta}(1 - \max\{x_L, \overline{p}_L\}) \text{ and } \overline{p}_L \leq e^{-r_L\Delta}\overline{p}_R. \tag{A.19}$$

Assume, for contradiction, that $\underline{p}_R < 1 - x_R$. Then from (A.18)

$$\underline{p}_L \geq e^{-r_L\Delta}\underline{p}_R \geq e^{-r_L\Delta}(1 - e^{-r_R\Delta}(1 - \underline{p}_L))$$

$$\text{or, } \underline{p}_L(1 - e^{-(r_L+r_R)\Delta}) \geq e^{-r_L\Delta}(1 - e^{-r_R\Delta})$$

$$\text{or, } \underline{p}_L \geq \frac{e^{-r_L\Delta}(1 - e^{-r_R\Delta})}{(1 - e^{-(r_L+r_R)\Delta})}.$$

Taking $\Delta \to 0$, we get

$$\underline{p}_L \geq \frac{r_R}{r_L + r_R}$$

Similarly, $\underline{p}_R \geq 1 - e^{-r_R\Delta}(1 - \underline{p}_L) \geq 1 - e^{-r_R\Delta}(1 - e^{-r_L\Delta}\underline{p}_R)$. Taking $\Delta \to 0$, we get

$$\underline{p}_R \geq \frac{r_R}{r_L + r_R}.$$

However, assumption 1 gives us $1 - x_R \leq \frac{r_R}{r_L+r_R}$, which contradicts $\underline{p}_R < 1 - x_R$. Therefore, $\underline{p}_R = 1 - x_R$. This implies $1 - e^{r_R\Delta}x_R \geq \underline{p}_L \geq e^{-r_L\Delta}(1 - x_R)$. As $\Delta \to 0$, we have $\underline{p}_L = 1 - x_R$.

Similarly, if $\overline{p}_L > x_L$, then from (A.19) and taking $\Delta \to 0$, we get

$$\overline{p}_L \leq \frac{r_R}{r_L + r_R}, \ \overline{p}_R \leq \frac{r_R}{r_L + r_R}.$$

However, assumption 1 gives us $x_L \geq \frac{r_R}{r_L+r_R}$, which contradicts $\overline{p}_L > x_L$. Therefore, $\overline{p}_L = x_L$. This implies $1 - e^{-r_R\Delta}(1 - x_L) \geq \overline{p}_R \geq e^{r_L\Delta}x_L$. As $\Delta \to 0$, we have $\overline{p}_R = p^L$.

Thus, in any PBE, leader L always accept policy $p \geq x_L$ and always rejects policy $p < 1 - x_R$, and leader R always accept policy $p \leq 1 - x_R$ and always rejects policy $p > x_L$. If leader L proposes a policy $p > x_L$, then in any PBE, at any history, the proposal will be rejected by all types of leader R. Analogously, if leader R proposes a policy $p < 1 - x_R$, then in any PBE, at any history, the proposal will be rejected by all types of leader L. We only look into PBE where leaders do not make policy proposals which are rejected by all types of her opponent, at all histories in any PBE. Finally, note that, by definition $\omega_L = I$ never accepts a policy $p < x_L$, and $\omega_R = I$ never accepts a policy $p > 1 - x_R$. Therefore the I types endogenously choose to behave like the commitment types — always demand the policy they agree on when the sentiment is commonly known, accepts a better policy with probability 1 and never accepts a worse policy. □

References

D. Abreu and F. Gul. Bargaining and reputation. *Econometrica*, 68(1):85–117, 2000. 4

D. Abreu, D. Pearce, and E. Stacchetti. One-sided uncertainty and delay in reputational bargaining. *Theoretical Economics*, 10(3):719–773, 2015. 4

A. Acharya and E. Grillo. War with crazy types. *Political Science Research and Methods*, 3(2):281–307, 2015. 7

A. Alesina and A. Drazen. Why are stabilizations delayed? *American Economic Review*, 81(5):1170–1188, 1991. 6

A. E. Atakan and M. Ekmekci. Bargaining and reputation in search markets. *Review of Economic studies*, 81(1):1–29, 2014. 29

D. Basak and J. Deb. Gambling over public opinion. *The American Economic Review*, 110.11:3492–3521, 2020. 6

K. Binmore, A. Rubinstein, and A. Wolinsky. The nash bargaining solution in economic modelling. *The RAND Journal of Economics*, pages 176–188, 1986. 28

K. Binmore, A. Shaked, and J. Sutton. An outside option experiment. *The Quarterly Journal of Economics*, pages 753–770, 1989. 2, 9, 31

K. Chatterjee and L. Samuelson. Bargaining with two-sided incomplete information: An infinite horizon model with alternating offers. *The Review of Economic Studies*, 54(2):175–192, 1987. 7, 27, 28

K. Chatterjee and L. Samuelson. Bargaining under two-sided incomplete information: The unrestricted offers case. *Operations Research*, 36(4):605–618, 1988. 7, 27, 28

G. Egorov and B. Harstad. Private politics and public regulation. *The Review of Economic Studies*, 84(4):1652–1682, 2017. 6

M. Ekmekci and H. Zhang. Reputational bargaining with ultimatum opportunities. arXiv preprint arXiv:2105.01581, 2021. 29

T. Ellingsen and T. Miettinen. Tough negotiations: Bilateral bargaining with durable commitments. *Games and Economic Behavior*, 87:353–366, 2014. 7

J. Fanning. Reputational bargaining and deadline. *Econometrica*, 84(3):1131–1179, 2016. 29

J. Fanning. No compromise: Uncertain costs in reputational bargaining. *Journal of Economic Theory*, 175:518–555, 2018. 8, 29

J. Fanning. Mediation in reputational bargaining. *The American Economic Review,*, 111.8:2444–72, 2021. 7, 29

J. Fanning and A. Wolitzky. Reputational bargaining. In *Bargaining*, pages 35–60. Palgrave Macmillan, Cham, 2022. http://economics.mit.edu/files/21210. 7

J. D. Fearon. Domestic political audiences and the escalation of international disputes. *American political science review*, 88(3):577–592, 1994. 6

Y. Feinberg and A. Skrzypacz. Uncertainty about uncertainty and delay in bargaining. *Econometrica*, 73(1):69–91, 2005. 7, 30

S. Kambe. Bargaining with imperfect commitment. *Games and Economic Behavior*, 28(2):217–237, 1999. 29

B. Levenotoğlu and A. Tarar. Prenegotiation public commitment in domestic and international bargaining. *American Political Science Review*, 99(3):419–433, 2005. 6

K. Madarász. Bargaining under the illusion of transparency. *American Economic Review*, 111(11):3500–3539, 2021. 7

A. Muthoo. Revocable commitment and sequential bargaining. *The Economic Journal*, 102(411):378–387, 1992. 6

R. B. Myerson. Game theory: analysis of conflict. *Cambridge: Mass, Harvard University*, 1991. 29

S. Özyurt. Searching for a bargain: power of strategic commitment. *American Economic Journal: Microeconomics*, 7(1):320–53., 2015. 29

A. Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica*, 50(1): 97–109, 1982. 2, 8

A. Sanktjohanser. Optimally stubborn. *Working Paper*, 2018. 30

T. C. Schelling. An essay on bargaining. *The American Economic Review*, pages 281–306, 1956. 6

A. Tsoy. Alternating-offer bargaining with the global games information structure. *Theoretical Economics*, 13(2):869–931, 2018. 7

J. Watson. Alternating-offer bargaining with two-sided incomplete information. *The Review of Economic Studies*, 65(3):573–594, 1998. 27

J. Weinstein and M. Yildiz. Robust predictions in infinite-horizon games—an unrefinable folk theorem. *Review of Economic Studies*, 80.1:365–394, 2013. 7

A. Wolitzky. Reputational bargaining with minimal knowledge of rationality. *Econometrica*, 80(5):2047–2087, 2012. 29

Online Appendix

Proof of Lemma 3 This proof is a simple extension of Lemma 1 in AG. Suppose that the leader i has revealed that she is not the I type. However, the I' and $\mathcal{U}$ type of leader i assign strictly positive probabilities that leader j could be the I type. It follows from Lemma 1 in AG that the $\omega_i = I'$ will concede immediately. However, unlike $\omega_i = I'$, the $\omega_i = \mathcal{U}$ is uncertain about the state, and updates her belief about the state over time. This affects her incentive to continue bargaining. The following proof accommodates this difference and follows the same steps of Lemma 1 in AG to show that the Coasian result holds.

Step I: The game must end in finite time with probability 1.

$\omega_j = I$ is committed to x_j. The non-committed types of leader j can always pretend to be the I type. Let $\psi_j^{\omega_j}(t)$ be the probability that type ω_j of leader j insists on getting x_j until some time t. Then, $\omega_i = \mathcal{U}$ believes that leader j will insist on getting x_j with probability $\psi_j(t) = \pi_j(1-\varepsilon_j)+\pi_i(1-\varepsilon_j)\psi_j^{I'}(t)+\varepsilon_j\psi_j^{\mathcal{U}}(t)$. After seeing such insistence until time t, she believes that the sentiment leans the other way with probability

$$\pi_j(t) = \frac{\pi_j((1-\varepsilon_j)+\varepsilon_j\psi_j^{\mathcal{U}}(t))}{\psi_j(t)}.$$

Consider $t_2 > t_1$. Suppose leader j has been insisting on x_j until time t_1. If the $\omega_i = \mathcal{U}$ accepts leader j's offer then she gets $e^{-r_i t_1}\pi_j(t_1)(1-x_j)$. If she continues bargaining until t_2, then the most she can get is

$$\frac{\psi_j(t_2)}{\psi_j(t_1)}e^{-r_i t_2}\pi_j(t_2)(1-x_j) + \left(1-\frac{\psi_j(t_2)}{\psi_j(t_1)}\right)e^{-r_i t_1}x_i.$$

Therefore, after seeing leader j insisting until time t_1, the $\omega_i = \mathcal{U}$ will continue bargaining until t_2 only if

$$\frac{\psi_j(t_2)}{\psi_j(t_1)}e^{-r_i(t_2-t_1)}\pi_j(t_2)(1-x_j) + \left(1-\frac{\psi_j(t_2)}{\psi_j(t_1)}\right)x_i \geq \pi_j(t_1)(1-x_j).$$

This simplifies to

$$\frac{\psi_j(t_2)}{\psi_j(t_1)} \leq \frac{x_i - \pi_j(t_1)(1-x_j)}{x_i - e^{-r_i(t_2-t_1)}\pi_j(t_2)(1-x_j)} = \frac{x_i - \frac{\pi_j((1-\varepsilon_j)+\varepsilon_j\psi_j^{\mathcal{U}}(t_1))}{\psi_j(t_1)}(1-x_j)}{x_i - e^{-r_i(t_2-t_1)}\frac{\pi_j((1-\varepsilon_j)+\varepsilon_j\psi_j^{\mathcal{U}}(t_2))}{\psi_j(t_2)}(1-x_j)}$$

$$\begin{aligned}
\implies \psi_j(t_2)x_i - e^{-r_i(t_2-t_1)}\pi_j((1-\varepsilon_j)+\varepsilon_j\psi_j^{\mathcal{U}}(t_2))(1-x_j) \\
\leq \psi_j(t_1)x_i - \pi_j((1-\varepsilon_j)+\varepsilon_j\psi_j^{\mathcal{U}}(t_1))(1-x_j) \\
\implies \psi_j(t_1) - \psi_j(t_2) \geq
\end{aligned}$$

$$\frac{\pi_j(1-x_j)}{x_i}\cdot\big((1-\varepsilon_j)(1-e^{-r_i(t_1-t_2)})+\varepsilon_j(\psi_j^{\mathcal{U}}(t_1)-e^{-r_i(t_2-t_1)}\psi_j^{\mathcal{U}}(t_2))\big)$$

$$\implies \psi_j(t_1)-\psi_j(t_2)\geq\frac{\pi_j(1-x_j)}{x_i}\cdot\big((1-\varepsilon_j)(1-e^{-r_i(t_2-t_1)})\big)=:\kappa.$$

Consider $t_2-t_1=\tau$. Starting at time 0, the $\omega_i=\mathcal{U}$ will continue bargaining for next τ time only if she believes that leader j will insists on getting x_j with probability less than $1-\kappa$. Repeating the same argument, she will continue bargaining for 2τ time only if she believes that leader j will insists on getting x_j with probability less than $1-2\kappa$, and so on. There exists K such that $1-K\kappa<\pi_j(1-\varepsilon_j)$. Therefore, $\omega_i=\mathcal{U}$ will keep bargaining until time $K\tau$ only if she believes that leader j will insist on getting x_j with probability less than $\pi_j(1-\varepsilon_j)$. This contradicts the fact that leader j could be type I and always insists on getting x_j.

Step II: $\omega_i=\mathcal{U}$ must concede immediately.

Suppose, for contradiction, that this is not true. Let $\bar{t}>0$ be the supremum of the time such that the $\omega_i=\mathcal{U}$ has not conceded and accepted leader j's offer. Consider the last ϵ time interval, $(\bar{t}-\epsilon,\bar{t})$. Let x be the sup of $\omega_i=\mathcal{U}$'s payoff if leader j agrees to take less than x_j in $(\bar{t}-\epsilon,\bar{t}-(1-\beta)\epsilon)$, where $\beta\in(0,1)$. Let y be the sup of $\omega_i=\mathcal{U}$'s payoff if leader j does not do so. Let $\xi=\frac{\psi_j(\bar{t}-(1-\beta)\epsilon)}{\psi_j(\bar{t}-\epsilon)}$ be the probability $\omega_i=\mathcal{U}$ assigns to leader j not accepting anything below x_j in $(\bar{t}-\epsilon,\bar{t}-(1-\beta)\epsilon)$.

At any t, leader j can behave like the $\omega_j=I$ and insist on getting x_j, and thus, guarantee herself $e^{-r_j(\bar{t}-t)}x_j$. So the maximum share leader j can get is $(1-e^{-r_j(\bar{t}-t)}x_j)$. Therefore, $\omega_i=\mathcal{U}$'s expected payoff cannot be higher than $\pi_j(t)(1-e^{-r_j(\bar{t}-t)}x_j)$. This gives us

$$x\leq\pi_j(\bar{t}-\epsilon)(1-e^{-r_j\epsilon}x_j),$$
$$y\leq e^{-r_j\beta\epsilon}\pi_j(\bar{t}-(1-\beta)\epsilon)(1-e^{-r_j(1-\beta)\epsilon}x_j).$$

If the $\omega_i=\mathcal{U}$ accepts leader j's offer she gets $\pi_j(\bar{t}-\epsilon)(1-x_j)$ and if she continues bargaining until $\bar{t}$, then she gets at most

$$(1-\xi)\pi_j(\bar{t}-\epsilon)(1-e^{-r_j\epsilon}x_j)+\xi e^{-r_j\beta\epsilon}\pi_j(\bar{t}-(1-\beta)\epsilon)(1-e^{-r_j(1-\beta)\epsilon}x_j).$$

Therefore, after no agreement until time $\bar{t}-\epsilon$, $\omega_i=\mathcal{U}$ keeps bargaining until $\bar{t}$ only if

$$\xi\leq\frac{\pi_j(\bar{t}-\epsilon)(1-e^{-r_j\epsilon}x_j-(1-x_j))}{\pi_j(\bar{t}-\epsilon)(1-e^{-r_j\epsilon}x_j)-e^{-r_j\beta\epsilon}\pi_j(\bar{t}-(1-\beta)\epsilon)(1-e^{-r_j(1-\beta)\epsilon}x_j)}$$

$$\text{or, }\frac{\psi_j(\bar{t}-(1-\beta)\epsilon)}{\psi_j(\bar{t}-\epsilon)}\leq\frac{1-e^{-r_j\epsilon}x_j-(1-x_j)}{1-e^{-r_j\epsilon}x_j-e^{-r_j\beta\epsilon}\frac{\pi_j(\bar{t}-(1-\beta)\epsilon)}{\pi_j(\bar{t}-\epsilon)}(1-e^{-r_j(1-\beta)\epsilon}x_j)}$$

After substituting $\pi_j(.)$, this simplifies to

$$(1-e^{-r_j\epsilon}x_j)\frac{\psi_j(\bar{t}-(1-\beta)\epsilon)}{\psi_j(\bar{t}-\epsilon)}-e^{-r_j\beta\epsilon}(1-e^{-r_j(1-\beta)\epsilon}x_j)\frac{(1-\varepsilon_j)+\varepsilon_j\psi_j^{\mathcal{U}}(\bar{t}-(1-\beta)\epsilon)}{(1-\varepsilon_j)+\varepsilon_j\psi_j^{\mathcal{U}}(\bar{t}-\epsilon)}$$

$$\leq 1-e^{-r_j\epsilon}x_j-(1-x_j).$$

$$\text{or, } A\xi - B\xi^j \leq \frac{1-e^{-r_j\epsilon}x_j-(1-x_j)}{(1-e^{-r_j\epsilon}x_j)-e^{-r_j\beta\epsilon}(1-e^{-r_j(1-\beta)\epsilon}x_j)}, \tag{B.1}$$

where,

$$\xi^j = \frac{(1-\varepsilon_j)+\varepsilon_j\psi_j^{\mathcal{U}}(\bar{t}-(1-\beta)\epsilon)}{(1-\varepsilon_j)+\varepsilon_j\psi_j^{\mathcal{U}}(\bar{t}-\epsilon)}$$

is the conditional probability of insistence by leader j until $\bar{t}-(1-\beta)\epsilon$ from $\bar{t}-\epsilon$ given the state leans toward j, and

$$A = \frac{(1-e^{-r_j\epsilon}x_j)}{(1-e^{-r_j\epsilon}x_j)-e^{-r_j\beta\epsilon}(1-e^{-r_j(1-\beta)\epsilon}x_j)}$$

$$B = \frac{e^{-r_j\beta\epsilon}(1-e^{-r_j(1-\beta)\epsilon}x_j)}{(1-e^{-r_j\epsilon}x_j)-e^{-r_j\beta\epsilon}(1-e^{-r_j(1-\beta)\epsilon}x_j)}.$$

Note that

$$A-B=1.$$

It is easy to check that $0 \leq \xi \leq \xi^j \leq 1$.

Claim 1: *For $\beta \in (x_j, 1)$, $\exists \delta_\beta < 1$ such that when $\epsilon \to 0$, $\xi < \delta_\beta$.*

Proof. Assume for contraction that the above claim does not hold true. Then, there exists a subsequence of $\xi(\epsilon_n)$ that converges to 1 while $\epsilon_n \to 0$. If $\xi = 1$, then it must be that $\xi^j = 1$. This implies that the LHS of equation (B.1) is $A - B = 1$. Taking $\epsilon \to 0$ and using L'Hospital rule, the right hand side of equation (B.1) becomes

$$\frac{r_jx_j}{r_jx_j + r_j\beta(1-x_j) - r_j(1-\beta)x_j} = \frac{x_j}{\beta}.$$

Therefore, when $\beta \in (x_j, 1)$, for any sequence of $\epsilon_n \to 0$, the RHS of equation (B.1) is strictly less than 1. This contradicts the inequality in equation (B.1). □

Therefore, $\omega_i = \mathcal{U}$ will play a strategy that can continue the bargaining from $(\bar{t}-\epsilon)$ to $(\bar{t}-(1-\beta)\epsilon)$ only if $\psi_j(\bar{t}-(1-\beta)\epsilon) < \delta_\beta\psi_j(\bar{t}-\epsilon)$. Repeating the same argument, the $\omega_i = \mathcal{U}$ will continue bargaining from $\bar{t}-(1-\beta)\epsilon$ to $\bar{t}-(1-\beta)^2\epsilon$ only if $\psi_j(\bar{t}-(1-\beta)^2\epsilon) < \delta_\beta\psi_j(\bar{t}-(1-\beta)\epsilon) < \delta_\beta^2\psi_j(\bar{t}-\epsilon)$. Repeating the argument K times we get, $\psi_j(\bar{t}-(1-\beta)^K\epsilon) < \delta_\beta^K\psi_j(\bar{t}-\epsilon)$. Since leader j could be the I type and always insists on getting x_j, $\psi_j(.) \geq \pi_j(1-\varepsilon_j)$. However, for K such that $\delta_\beta^K < \pi_j(1-\varepsilon_j)$, the above inequality cannot hold true. Note that leader i always get the chance to make offers sufficiently close to $\bar{t}-(1-\beta)^m\epsilon$ for all $m = 1, 2 \ldots K$. Therefore, $\omega_i = \mathcal{U}$ will never play a strategy that will continue the bargaining until $\bar{t}$. This contradicts the definition of $\bar{t}$. □

Equilibrium payoff comparison

Let us define $V_i^{\omega_i}$ as the expected payoff of type ω_i of leader i in equilibrium. Suppose that the leader j is a weak bargainer. $\omega_j = \mathcal{U}$ plays the following strategy with positive probability — wait until T and then concede if leader i has not conceded already. Therefore,

$$V_j^{\mathcal{U}} = \left[\pi_j(1-\varepsilon_i)\int_0^{T_i[I']} e^{-r_j t} dF_i^{I'}(t) + \varepsilon_i \int_{T_i[I']}^{T} e^{-r_j t} dF_i^{\mathcal{U}}(t)\right] x_j$$

$$+\pi_i(1-\varepsilon_i)e^{-r_j T}(1-x_i).$$

In contrast, $\omega_j = I$ always insists, and hence her payoff is

$$V_j^{I} = \left[(1-\varepsilon_i)\int_0^{T_i[I']} e^{-r_j t} dF_i^{I'}(t) + \varepsilon_i \int_{T_i[I']}^{T} e^{-r_j t} dF_i^{\mathcal{U}}(t)\right] x_j.$$

$$= V_j^{\mathcal{U}} + \pi_i(1-\varepsilon_i)\left(\int_0^{T_i[I']} e^{-r_j t} dF_i^{I'}(t) x_j - e^{-r_j T}(1-x_i)\right) > V_j^{\mathcal{U}}.$$

Next, consider the strong bargainer i. Unlike $\omega_j = \mathcal{U}$, the $\omega_i = \mathcal{U}$ keeps waiting after seeing no immediate concession from her opponent. She waits until $T_i[I']$ before starting to concede. She plays the following strategy with positive probability — wait until T and then concede if leader j has not conceded already. Therefore,

$$V_i^{\mathcal{U}} = \left[\pi_i(1-\varepsilon_j) + \varepsilon_j \int_{T_i[I']}^{T} e^{-r_i t} dF_j^{\mathcal{U}}(t)\right] x_i + \pi_j(1-\varepsilon_j)e^{-r_i T}(1-x_j).$$

$\omega_i = I$ always insists, and hence her payoff is

$$V_i^{I} = \left[(1-\varepsilon_j) + \varepsilon_j \int_{T_i[I']}^{T} e^{-r_i t} dF_j^{\mathcal{U}}(t)\right] x_i$$

$$= V_i^{\mathcal{U}} + \pi_j(1-\varepsilon_j)\left(x_i - e^{-r_j T}(1-x_j)\right) > V_i^{\mathcal{U}}.$$

Thus, the stubborn type gets a higher expected payoff compared to the uninformed type. An uninformed leader j believes that that leader i could be stubborn with probability $\pi_i(1-\varepsilon_i)$. Unlike the stubborn type, the uninformed leader j can raise her offer at the last minute and recovers $e^{-r_j T}(1-x_i)$. In contrast, the stubborn type knows that the opponent cannot be stubborn and assigns this probability $\pi_i(1-\varepsilon_i)$ to leader i being the I' type. Therefore, with probability $\pi_i(1-\varepsilon_i)$ she gets $\int_0^{T_i[I']} e^{-r_j t} dF_i^{I'}(t) x_j$, which is strictly higher than what the uninformed type can recover by giving up at the last minute. □