Bargaining Power in Crisis Bargaining

Bahar Leventoğlu*

September 2022

Abstract

A large body of game-theoretic work examines the process by which uncertainty can lead to inefficient war. In a typical crisis bargaining model, players negotiate according to a pre-specified bargaining protocol and no player has the ability to change the rules of the game. However, when one of the parties has full bargaining power and is able to set the bargaining protocol on her own, the protocol itself becomes an endogenous decision variable. I formulate this problem in a principal-agent framework. I show that both the likelihood of costly war and the exact mechanism that yields it depend on the nature of the informational problem as well as the identity of the informed player.

*Duke University, Department of Political Science, Durham, NC 27708. E-mail: bahar.leventoglu@duke.edu ORCID ID: 0000-0001-5687-8451 An earlier version of this manuscript was circulated under the title "Crisis Bargaining without Mediation"

1 Introduction

The international relations literature has long argued that some kind of incomplete information, uncertainty and misperception between more-or-less rationally led states make them go into war (Blainey, 1988; Fearon, 1995; Jervis, 1976; Wittman, 1979; Van Evera, 1999). There is a large body of game-theoretic work that examines the process through which uncertainty can lead to inefficient war. A central result is that incentives to misrepresent privately held information play a key role in shaping the behavior of the participants and thereby the likelihood of war and the nature of a peaceful settlement (Fearon, 1995; Powell, 1996a, 1999; Schultz, 1998; Slantchev, 2003; Smith and Stam, 2006; Wagner, 2000).

A typical crisis bargaining game assigns a particular player as first mover in the game to make an offer for a peaceful settlement. Although such assumptions yield some bargaining power for the first mover, the impact of bargaining power on crisis resolution has not been studied systematically. I develop a systematic approach to study the consequences of information problem when one of the states possesses full bargaining power. I refer to that state as the powerful state.

As pointed out by Banks (1990) and Fey and Ramsay (2011), the formal crisis bargaining literature hardly agrees on the practice of game-theoretic modeling. Who will make the first offer? Will it be a take-it-or-leave-it offer model, or can a second offer be made when the first one is rejected? In the latter case, will one side make all the offers or will it be an alternating offers model? When, in the game, do players make an offer and when do they fight?

Such assumptions have significant implications on predictions. For example, Fearon (1995) identifies a risk-return trade-off in a take-it-or-leave-it bargaining game with uncertainty about the opponent's resolve. Powell (1996a) generalizes this risk-return trade-off argument in an alternating offers model in which a player may reject an offer to make a counter offer. Leventoğlu and Tarar (2008) show that a small and intuitive change to the timing of when players can engage in war in Powell's model can generate a peaceful outcome with no risk-return trade-off.

Banks (1990) and Fey and Ramsay (2011) adopt a mechanism design approach (Myerson, 1979) and study a general information revelation game. This approach allows to make predictions that are robust to details of a game that that actors engage in. If a prediction holds in all incentive compatible direct mechanisms, it holds in every equilibrium of every game that parties may be playing. If a property does not hold in any incentive compatible direct mechanisms, then it does not hold in any equilibrium of any game (Fey and Ramsay, 2011).

Despite robust predictions, the mechanism design model does not explain when a prediction would prevail if it did not hold in all incentive compatible direct mechanisms. Rationality implies that a state with full bargaining power will choose a bargaining protocol that will benefit her the most. The principal-agent framework deals with information and incentive issues between a principal and an agent when the principal has full bargaining power. In contrast to the mechanism design approach, the principal-agent framework provides unique predictions under the full bargaining power assumption.

Informational asymmetry in crisis bargaining models typically concerns either the cost of war players incur (Fearon, 1994; Powell, 1996a; Schultz, 1999) or the power distribution between them, i.e. the probability of victory in war (Reed, 2003; Smith and Stam, 2006; Wagner, 2000; Wittman, 1979). I study these two special types of private information and refer to a player's private information as that player's type.

These two types of informational problems are fundamentally different. A player's cost of war determines only his or her war payoff, but it does not say anything about the adversary's war payoff. This is known as the private values case. In contrast, private information about the probability of victory in war determines one's and her adversary's war payoffs. This is known as the common values case.

When the adversary agent holds private information, the standard principal-agent framework (Salanie, 1999; Bolton and Dewatripont, 2005) applies in both cases of private values and common values. If the powerful state has private information, then the problem turns into a principal-agent problem with informed principal (Maskin and Tirole, 1990, 1992), because the powerful state's choice of bargaining protocol may signal her private information to the adversary, which in turn may affect the adversary's beliefs and incentives.

Maskin and Tirole (1990) develop a method to solve the problem of an informed principal in the case of private values. They assume that some types of principals may violate the individual rationality constraint for the agent, however the individual rationality constraint has to hold in expectations.[1] In contrast, a fundamental assumption in crisis bargaining is that no player can be forced to accept any deal that is worse than their war payoff. The individual rationality constraint cannot be violated by any type of principal.[2] Then, Maskin and Tirole approach is reduced to the standard principal-agent framework in my crisis bargaining model.

I defer the analysis of these three cases to an appendix. I study the case in which the state with full bargaining power holds information about the probability of victory in a war. The probability of victory determines war payoffs for both sides. The choice of the bargaining protocol by the powerful state may reveal information. This turns the game into a signaling game where the standard principal-agent framework does not apply. Unlike standard signaling games in which the game form is exogenously given, the contract (or the game form) is endogenously determined.

I extend the informed principal with common values method of Maskin and Tirole (1992). In Maskin and Tirole, once players sign a contract, they commit to the terms of the contract even after information is revealed. This relaxes the individual rationality constraint for the uninformed state by forcing him to accept terms that are worse than his war payoff once information

---

[1] see their problem F_*^i on page 392

[2] Technically, this means that the only feasible value for r^i is zero in their problems F_*^i and V_I^i on page 392 of Maskin and Tirole (1990).

is revealed. In a crisis bargaining game, no player can be forced to accept a deal that is worse than his or her expected war payoff. I extend Maskin and Tirole by incorporating this fundamental assumption in individual rationality constraints.

I showcase the methodology in a take-it-or-leave-it offer game with common values in which the state with private information makes the offer. The take-it-or-leave-it bargaining protocol is commonly applied in crisis bargaining. I study a bluffing equilibrium of the game. In a bluffing equilibrium, a weak type may rationally imitate a strong type in a semi separating equilibrium. Such bluffing may increase the likelihood of war and the weak type's share from a peaceful settlement (de Mesquita and Lalman, 1992). But this equilibrium does not survive as the optimal resolution of the conflict from the point of view of the powerful state. Therefore, information asymmetry alone is not sufficient for bluffing to cause war when the informed state has full bargaining power.[3] I use the same game to implement the optimal solution.

The paper proceeds as follows: Section 2 introduces model and discusses methodology. Section 3 introduces a take-it-or-leave-it offer game and studies two equilibria of the game. Section 4 extents and solves the common values with informed principal problem. Section 5 discusses the results and Section

[3]Similarly, Fey and Ramsay (2011) find peace as one of the robust predictions when private information is about cost of war. They show that there always exists a game form in which parties peacefully settle and sharing resources proportional to their respective military strengths is a necessary condition for a game form to always have a peaceful equilibrium. When one of the parties has full bargaining power, Fearon's risk-return trade-off argument (Fearon 1995) prevails as the unique prediction.

6 concludes. I defer all technical proofs to an appendix.

2 The Model and the methodology

Two states, D (he) and S (she) have a dispute over a divisible good of size 1. D and S's status quo shares of the disputed good are q and $1-q$, respectively, where $0 \leq q \leq 1$. They can live with status quo, reach a peaceful agreement to reallocate the good between themselves, or they can go to war. If parties go to war, the state that wins the war obtains the entire good. D wins the war with probability p and S wins the war with probability $1-p$. Fighting is costly, D and S pay a cost of $c_D, c_S > 0$, respectively, if they go to war. The players are risk neutral. Therefore, the expected payoff from war to D is $EU_D(war) = p1 + (1-p)0 - c_D = p - c_D$ and the expected payoff from war to S is $EU_S(war) = (1-p)1 + p0 - c_S = 1 - p - c_S$. If the two states do not reallocate the good peacefully and do not go to war, then the status quo prevails.

A state is "satisfied" if it receives from its status quo share as much utility as it receives from war and it is "dissatisfied" if it strictly prefers to go to war rather than living with its status quo share (Powell, 1996a,b). At most one state can be dissatisfied. I assume that S is satisfied and D is dissatisfied. That is, $1 - q > 1 - p - c_S$ and $p - c_D > q$ for all values of p, c_S and c_D.

An information structure for the values of p, c_S and c_D and a bargaining protocol, i.e. game form, for a peaceful resolution of the conflict complete a typical crisis bargaining model. In the main text, I assume that c_D and c_S are common knowledge, but p is equal to p_h with probability π and to $p_l < p_h$ with probability $1 - \pi$. Although the distribution of p is common knowledge, only S knows the true value of p.

I depart from the literature by not assuming any specific bargaining protocol. Instead, I assume that S has full bargaining power and the ability to set the bargaining protocol on her own. Therefore, the protocol is determined endogenously in the model. I adopt a principal-agent framework to study the problem, in which S is the principal and D is the agent.

The informational assumptions of the crisis bargaining literature fall into two broad categories: Information about cost of war, c_S and c_D, and information about p. In the former case, p is common knowledge and each player privately knows their cost of war. Then D's private information does not affect S's war payoff $1 - p - c_S$ and vice versa. This is a private values set up. In the latter case, players' cost of war are common knowledge, one of the players privately knows the true value of p and both players' war payoffs are determined by this privately held information. This case is called common values.

The identity of the player that holds private information matters. If D has private information, the nature of the information does not matter for S. Regardless of whether there is a private values or a common values situation, S has no information that might affect D's payoffs, so her choice of the contract will not have any impact on the set of individually rational and incentive compatible contracts. This is a standard principal-agent problem (Salanie, 1999; Bolton and Dewatripont, 2005). I defer the analysis of this case to an appendix.

However, if S holds private information, her choice of contract can potentially signal her private information to D and thereby affect the terms of a peaceful bargain. In effect, the problem turns into a signaling game and the contract that S chooses becomes a signal for D. When the private information concerns cost of war, it becomes a principal-agent problem with informed principal and private values (Maskin and Tirole, 1990). I defer the analysis of this case to the appendix. Below, I study the case in which S knows the value of p. This is a principal-agent problem with informed principal and common values (Maskin and Tirole, 1992). I extend the methodology by incorporating individual rationality constraints imposed by the very nature of a crisis bargaining problem.

First, I analyse the crisis bargaining scenario in a take-it-or-leave-it offer game below. Then I study it in the principal-agent framework.

3 Take-It-Or-Leave-It Offer Game

A take-it-or-leave-it offer (TILI) is one of the canonical crisis bargaining games in the international relations literature (Fearon, 1995). In this game, S offers $t \in [0, 1]$. If D accepts the offer, he gets t, S gets $1 - t$ and the game ends. If D rejects the offer, the plays go to war.

Suppose that $c_S + c_D \geq \pi(p_h - p_l)$. Refer to S as S_h if $p = p_h$ and S_l if $p = p_l$.

I will study two Perfect Bayesian equilibria of this game. In the first equilibrium, S_l signals her type by making a lower offer, which carries risk of war. In the second equilibrium, S_h bluffs by imitating S_l. The former is a separating equilibrium and the latter is a semi-separating equilibrium. I will revisit these equilibria later and defer the analysis to an appendix.

Signaling Equilibrium

Let t_h and t_l be the equilibrium offers by S_h and S_l, respectively, and α be the probability that D rejects the low offer t_l, where

$$t_h = p_h - c_D$$

$$t_l = p_l - c_D < t_h$$

$$\alpha = \frac{p_h - p_l}{(p_h - p_l) + (c_D + c_S)}$$

The following strategy and belief profile forms a PBE:

- S_l offers t_l and S_h offers t_h.
- D's beliefs and strategy are given as follows:
  - If D receives an offer of $t > t_l$, he believes that $p = p_h$ with probability 1, otherwise he believes that $p = p_l$ with probability

1.

- If D receives an offer of $t \geq t_h$, he accepts the offer
- If D receives an offer of $t \in (t_l, t_h)$ or $t < t_l$, he rejects the offer
- If D receives an offer of t_l, he accepts the offer with probability α and rejects it with probability $1 - \alpha$

Bluffing Equilibrium

Let t_h and t_l be the equilibrium offers by S_h and S_l, respectively, and α be the probability that D rejects the low offer t_l and ϕ be D's posterior belief for p after receiving an offer of t_l, $Prob(p = p_h | t_l)$. S_h bluffs in a semi separating equilibrium when she plays the same strategy with S_l. Let $\beta \in [0, 1]$ be the probability that S_h bluffs in equilibrium. Define

$$\phi = \frac{\pi\beta}{\pi\beta + (1 - \pi)} \in [0, 1)$$

$$t_h = p_h - c_D$$

$$t_l = \phi(p_h - c_D) + (1 - \phi)(p_l - c_D) = p_l - c_D + \phi(p_h - p_l) < t_h$$

and

$$\alpha = \frac{c_D + c_S}{c_D + c_S + (1 - \phi)(p_h - p_l)} \in (0, 1)$$

The following strategy and belief profile forms a PBE:

- S_l always offers t_l,

- S_h offers t_l with probability β and t_h with probability $1 - \beta$,

- D's beliefs and strategy are given as follows:

  - If D receives an offer of $t > t_l$, he believes that $p = p_h$ with probability 1; if he receives an offer of t_l, he believes that $p = p_h$ with probability ϕ; if he receives an offer of $t < t_l$, he believes that $p = p_l$ with probability 1.

  - D rejects any offer $t \in [0, p_h - c_D] \backslash \{t_l, t_h\}$, accepts any offer $t \geq t_h$.

  - If D receives an offer of t_l, he accepts it with probability α and rejects it with probability $1 - \alpha$.

4 Common values with Informed Principal

When S holds private information, S's choice of contract can signal her private information to D. In a common values set up, information transmission changes D's payoffs so S's problem turns into a signaling game without a pre-specified game form. Neither the standard principal-agent framework nor typical signaling games apply in this case. The solution of S's problem obtains by extending Maskin and Tirole (1992). The first step is to find an optimal separating contract. The second step looks for a pooling contract which is better than the optimal separating contract for both types of S. In the latter step, information revelation is still possible after the players commit to a pooling contract. This allows for a relaxation of the individual rationality constraint ex ante. The crisis bargaining model differs at this stage since no state can be enforced to accept a payoff less than their war payoff. I incorporate this fundamental feature in the framework below.

4.1 Optimal Separating Contracts

A contract specifies probability of war and a peaceful settlement contingent on type. Let α_τ be the probability of war and t_τ be D's share in a peaceful settlement when $p = p_\tau$, $\tau \in \{h, l\}$. Therefore, a contract is characterized by four values $\{\alpha_h, t_h, \alpha_l, t_l\}$. The contract also satisfies individual rationality and incentive compatibility constraints, as described below.

Let the low-resolve type S choose the contract (α_h, t_h) and the high-resolve type S choose the contract (α_l, t_l) in a separating equilibrium. Since D will find out about the true value of p when he observes S's choice of contract, each of these contracts must satisfy the associated individual rationality constraint for D:

$$t_\tau \geq p_\tau - c_D \text{ for } \tau \in \{h, l\}.$$

Since D does not hold private information, S is not constrained by incentive compatibility for D. However, each type of S must make sure that, given the contract for the other type, she chooses a contract that will separate her from the other type. That induces an incentive compatibility constraint for the other type. For example, a low-resolve type S collects an expected payoff of

$$\alpha_h(1 - p_h - c_S) + (1 - \alpha_h)(1 - t_h)$$

from her own contract and

$$\alpha_l(1 - p_h - c_S) + (1 - \alpha_l)(1 - t_l)$$

from choosing the contract of the high-resolve type. The former must be at least as big as the latter for the low-resolve type not to imitate the high-resolve type by choosing the contract of the high-resolve type.

A separating contract is optimal if, given the expected payoff for a type, it maximizes the expected payoff for the other type. Thus, a pair of optimal separating contracts $(\alpha_h^{sep}, t_h^{sep})$ and $(\alpha_l^{sep}, t_l^{sep})$ is a solution to the following two problems:

The problem of the high-resolve type S: Given $(\alpha_h^{sep}, t_h^{sep})$,

$$(\alpha_l^{sep}, t_l^{sep}) \text{ solves } \max_{(\alpha_l, t_l)} \ \alpha_l(1 - p_l - c_S) + (1 - \alpha_l)(1 - t_l)$$

$$\text{subject to}$$

$$D_l - IR : t_l \geq p_l - c_D$$

$$S_h - IC : \alpha_h^{sep}(1 - p_h - c_S) + (1 - \alpha_h^{sep})(1 - t_h^{sep}) \geq$$

$$\alpha_l(1 - p_h - c_S) + (1 - \alpha_l)(1 - t_l)$$

The problem of the low-resolve type S: Given $(\alpha_l^{sep}, t_l^{sep})$,

$$(\alpha_h^{sep}, t_h^{sep}) \text{ solves } \max_{(\alpha_h, t_h)} \; \alpha_h(1 - p_h - c_S) + (1 - \alpha_h)(1 - t_h)$$

$$\text{subject to}$$

$$D_h - IR : t_h \geq p_l - c_D$$

$$S_l - IC : \alpha_l^{sep}(1 - p_l - c_S) + (1 - \alpha_l^{sep})(1 - t_l^{sep}) \geq$$
$$\alpha_h(1 - p_l - c_S) + (1 - \alpha_h)(1 - t_h)$$

$S_h - IC$ ensures that the low-resolve type S will not have an incentive to choose the contract of the high resolve type. This is the incentive compatibility constraint that the high-resolve type S is subject to when she is choosing her contract. Similarly, the low-resolve type S is constrained by $S_l - IC$ when she chooses her own contract.

The following result states that only the high-resolve type S is constrained by incentive compatibility at the optimal separating contract.

Proposition 1 *$S_l - IC$ is slack and $S_h - IC$ is binding at the optimal separating contract.*

At the optimal solution, the high-resolve type S separates herself from the low-resolve type by choosing a contract that does not give the low resolve type S any incentive to mimic her. Therefore, the high-resolve type S pays the cost of separation at the optimal solution.

The following summarizes the unique optimal separating contract.

Proposition 2 *The pair of optimal separating contracts is unique and is given by* $(\alpha_h^{sep} = 0, t_h^{sep} = p_h - c_D)$ *and* $(\alpha_l^{sep} = \frac{p_h - p_l}{(p_h - p_l) + (c_S + c_D)} \in (0,1), t_l^{sep} = p_l - c_D)$.

Substantively, the high-resolve type S separates herself from the low-resolve type by being aggressive. She faces risk of war with positive probability in order to secure a bigger share of the pie in peace time.

The expected payoffs for the two types of S are given as follows.

$$V_h^{sep} = 1 - p_h + c_D$$
$$V_l^{sep} = 1 - p_l + c_D - \alpha_l^{sep}(c_D + c_S)$$

These payoffs set the lower bound for what each type of S can achieve in crisis bargaining with D. The next step searches for contracts that provide each type of S with a better payoff.

4.2 The best contract

The second step of Maskin and Tirole (1992) looks for a contract that is better than the optimal separating contract for both types of S and potentially involves separation of types of S and commitment by both D and S.

Formally, let both types of S offer the same pair of contracts $\{(\alpha_h^*, t_h^*), (\alpha_l^*, t_l^*)\}$, and if D agrees, and only after he agrees, S reveals her type and both D and S have to go with the terms of the contract associated with S's type.

Let V_h^* be the expected payoffs of S_h from (α_h^*, t_h^*) and V_l^* be the expected payoffs of S_l from (α_l^*, t_l^*). Then

$$(Better) : V_h^* \geq V_h^{sep} \text{ and } V_l^* \geq V_l^{sep}$$

must hold. Also, $S_h - IC$ and $S_l - IC$ must hold with (α_h^*, t_h^*) and (α_l^*, t_l^*).

Finally, an individual rationality constraint for D must hold. Since S's type is revealed only after both D and S commit to the terms of the contract, this induces the following ex ante individual rationality constraint for D:

$$D - IR^{MaskinTirole} : \pi \left[\alpha_l^*(p_l - c_D) + (1 - \alpha_l^*)t_l^*\right] + (1 - \pi)\left[\alpha_h^*(p_h - c_D) + (1 - \alpha_h^*)t_h^*\right]$$
$$\geq \pi(p_l - c_D) + (1 - \pi)(p_h - c_D)$$

When D decides to accept or reject the contract pair, he does not know the type of S he faces. He will learn S's type only after accepting the contract pair. Then both players will play the contract associated with S's type. The left hand side of $D - IR^{MaskinTirole}$ is D's expected payoff from accepting and committing to the terms of these contracts. Alternatively, he can reject the offer and fight, which provides him with the expected payoff on the right hand side.

The second step of Maskin and Tirole (1992) requires the constraints $(Better)$, $S_h - IC$, $S_l - IC$ and $D - IR^{MaskinTirole}$.

In a crisis bargaining game, no player can be forced to accept a deal that is worse than his or her expected war payoff. This fundamental assumption requires modification in Maskin and Tirole's individual rationality constraint. $D - IR^{MaskinTirole}$ does not ensure $t_\tau \geq p_\tau - c_D$ after S's type is revealed. In Maskin and Tirole (1992), D commits to living with t_τ even if $t_\tau < p_\tau - c_D$. In other words, the individual rationality constraint is relaxed ex ante and that is how (α_h^*, t_h^*) and (α_l^*, t_l^*) can potentially provide better payoffs than the optimal separating contracts in Maskin and Tirole (1992).

If $(\alpha_h^*, t_h^*) \neq (\alpha_l^*, t_l^*)$, then D will learn S's type and S will not be able to force D to accept anything less than his war payoff. Then $\{(\alpha_h^*, t_h^*), (\alpha_l^*, t_l^*)\}$ will constitute a separating equilibrium with the usual individual rationality constraints for D. Since the optimal separating contract is already characterized with these rationality constraints and it is unique, any such $\{(\alpha_h^*, t_h^*), (\alpha_l^*, t_l^*)\}$ can not be strictly better for S than the optimal separating contract.

This implies that the only other potentially better contract is a pooling contract that does not reveal any information, that is $(\alpha_h^*, t_h^*) = (\alpha_l^*, t_l^*) = (\alpha^*, t^*)$. Then $S_h - IC$ and $S_l - IC$ trivially hold and the individual rationality constraint becomes

$$D - IR^{Pooling} : t^* \geq \pi(p_l - c_D) + (1 - \pi)(p_h - c_D)$$

Also

$$V_h^* = \alpha^*(1 - p_h - c_S) + (1 - \alpha^*)(1 - t^*) \text{ and}$$
$$V_l^* = \alpha^*(1 - p_l - c_S) + (1 - \alpha^*)(1 - t^*)$$

I summarize this result in the following:

Proposition 3 *If the pair of optimal separating contracts is not the best choice for* S*, then the optimal contract is a pooling contract and it satisfies* $D - IR^{Pooling}$, $V_h^* \geq V_h^{sep}$ *and* $V_l^* \geq V_l^{sep}$.

The constraints $V_h^* \geq V_h^{sep}$ and $V_l^* \geq V_l^{sep}$ are required for optimality. If either of them is violated, then the associated type will have an incentive to separate herself by offering the optimal separating equilibrium contract.

The individual rationality constraint $D - IR^{Pooling}$ has an important substantive interpretation. In a separating contract, the individual rationality constraint holds for each type: $t_h^{sep} \geq p_h - c_D$ and $t_l^{sep} \geq p_l - c_D$, which also imply $D - IR^{Pooling}$. However, $D - IR^{Pooling}$ does not imply the former. In other words, $D - IR^{Pooling}$ relaxes the individual rationality constraint imposed by a separating contract on S. This is because when D is offered a pooling contract, he does not find out about the true value of p so his individual rationality constraint must only hold in expectation. This relaxation opens up the possibility of better bargains for S.

Then the solution for the best contract can be found by comparing the optimal pooling contract with the optimal separating contract. Optimality and $D - IR^{Pooling}$ imply that

$$\alpha^* = 0 \text{ and } t^* = \pi(p_l - c_D) + (1-\pi)(p_h - c_D)$$

must hold with an optimal pooling contract. That is, there is no fighting at the optimal pooling contract, which yields

$$V_h^* = V_l^* = 1 - t^* = 1 - p_h + c_D + \pi(p_h - p_l)$$

Finally, $V_h^* \geq V_h^{sep}$ holds and

$$V_l^* \geq V_l^{sep} \Leftrightarrow \pi \geq \pi^{**} = \frac{p_h - p_l}{(p_h - p_l) + (c_D + c_S)}$$

so the following main result follows:

Theorem 4 *Assume that S privately knows the true value of p.*

(i) If $\pi \geq \pi^{**}$, then the conflict is resolved efficiently at the optimal pooling equilibrium: Both types of S offer $t^* = \pi(p_l - c_D) + (1-\pi)(p_h - c_D)$ without risking war and D accepts the offer.

(ii) If $\pi < \pi^{**}$, then the conflict is resolved inefficiently at the optimal separating equilibrium: The low-resolve type S offers $p_h - c_D$ and avoids war; high-resolve type S fights with probability π^{**} and offers $p_l - c_D$ otherwise. Thus, the ex ante probability of war is $\pi\pi^{**}$.

Now let us contrast the optimal contract of an informed principal to that of an uninformed principal. The following theorem summarizes Result 6 for the uninformed principal case:

Theorem 5 *Assume that D privately knows the true value of p.*

(i) If $\pi \geq \pi^{**}$, then the conflict is resolved efficiently: S offers $p_h - c_D$ without risking war and D accepts the offer.

(ii) If $\pi < \pi^{**}$, then the conflict is resolved inefficiently: S offers $p_l - c_D$, which the low-resolve type D accepts and the high-resolve type D rejects and fights. So, the ex ante probability of war is π.

5 Discussion

Risk-return trade-off vs Separating-Pooling trade-off

When one of the parties has full bargaining power, inefficient war may break out for two reasons: The first one is Fearon's risk-return trade-off explanation (Fearon, 1995). Accordingly, the powerful party may risk war in order to obtain a bigger share of the pie in a peaceful settlement. The findings in the appendix imply that this prediction is robust to the nature of informational problem when the powerful state is not informed. Regardless of whether the informational asymmetry is about parties' individual costs of war or power distribution between them, war may break out as a consequence of a risk-return trade-off calculation by the powerful party when she is not informed.

If the powerful state knows that power distribution favors her, then she can signal her high-resolve by aggression. In order to separate herself from a low-resolve type, she can make an offer that carries risk of war. That can happen when the ex ante probability of her being a high-resolve type is low. In this case, a low-resolve type avoids war by offering her adversary a high share of $p_h - c_D$. The high-resolve type offers the adversary a low share of $p_l - c_D$ for a peaceful settlement that carries a positive probability of war.

When the probability of being a high-resolve type is high for the powerful state, then she can do better by pooling with the low-resolve type and avoiding risk of war. In this case, D does not find out about the true value of p after receiving a "pooling" offer. Thus, both types of S make a moderate offer between $p_l - c_D$ and $p_h - c_D$ and D's individual rationality constraint only holds in expectations.

In general, both the likelihood of war and the players' shares in a peaceful settlement depend on the identity of the informed player.

Take-it-or-leave-it offer

In his seminal paper, Fearon (1995) models crisis bargaining process as a take-it-or-leave-it offer game. Although this take-it-or-leave-it offer game provides most of the ingredients for modeling full bargaining power, it is still not a complete description. For example, a bluffing equilibrium may also arise in a crisis take-it-or-leave-it offer game. which can also explain war (de Mesquita and Lalman, 1992). Bluffing may increase the likelihood of war and the low-resolve type's share from a peaceful settlement. However, my findings imply that such rational behavior may emerge only if neither party possesses full bargaining power.

In contrast, the separating equilibrium implements the optimal solution for S. Therefore, unlike in a take-it-or leave-it crisis bargaining game, one needs to characterize the set of all equilibria and make additional assumptions to select from the set of multiple equilibria.

Mechanism Design vs Principal-Agent Framework

Banks (1990) and Fey and Ramsay (2011) address the following question: Given that there are many game forms and bargaining protocols, which predictions are robust to variations in the underlying game structure? These scholars adopt a mechanism design approach and provide "game-free" results without reference to a specific game form.

Their question is fundamentally different from the question I address in this paper. Mechanism design theory does not provide an answer for the question about which crisis bargaining game is going to be played. Moreover, mechanism design theory does not tell which prediction will survive when there are multiple robust predictions regarding the outcome. When one of the players can set the rules of the crisis bargaining game, the bargaining protocol itself becomes endogenous. In this case, the rationality assumption provides unique answers for both: The player chooses the protocol that benefits her the most, which induces the best outcome for her.

To illustrate, consider Fey and Ramsay's findings that if negotiating parties know each other's military strength but each party privately knows its cost of war, there always exists a game form in which parties peacefully settle (Proposition 2, Fey and Ramsay (2011)) and sharing resources proportional to their military strengths is a necessary condition for a game form to always have a peaceful equilibrium (Proposition 3, Fey and Ramsay (2011)). That is, peace is a robust prediction according to Fey and Ramsay (2011).

But this does not mean that peace will prevail. For example, consider two countries D and S engaged in a crisis bargaining over a pie of size 20. If they fight, the victor is determined with equal probability and obtains the entire pie. They also pay a cost for fighting. D's cost of war is either 5 or 1 with equal probability and S's cost of war is 2. D knows his cost but S cannot observe it and S's cost is commonly known.

If the players agree to play the game form with a peaceful equilibrium, each collects a payoff of 10. If they fight, the expected payoffs of a high-cost (low-resolve type) D, a low-cost (high-resolve type) D and S are given by 5, 9 and 8 respectively.

If S has the power to choose the bargaining protocol, then she can make a take-it-or-leave it offer of 6 to D. A high-cost D accepts this offer since it is higher than his expected war payoff of 5, and S collects a payoff of 14 (= 20 - 6). On the other hand, a low-cost D rejects it, because it is lower than his war payoff of 8, and S collects her war payoff of 8 as a result. S's expected payoff from this risky deal is 0.5x14 + 0.5x8 = 11 which is greater than her peace payoff of 10. Thus, S would risk war if she could make that offer.

In other words, existence of a game form with a peaceful equilibrium, even when its existence is robustly predicted as in Fey and Ramsay (2011), does not mean that peace will prevail. When one of the players has full bargaining power, the principal-agent approach predicts that peaceful bargaining does not always exist.

When D also knows something about p

The true value of p may be determined by private information of both S and D (Fey and Ramsay, 2011). Maskin and Tirole (1992) can be applied to this case as well. This requires two types of changes.

The first one concerns D's incentives to reveal his private information, so the associated problems are appended with incentive compatibility constraints for D. In the problem of optimal separating contract, D will learn S's type. Then the incentive compatibility constraint will hold for each type of D given that he knows S's type. In contrast, in a pooling contract, D will not learn S's type so the incentive compatibility constraint will only hold in expectations for each type of D.

The second concerns S's payoffs from contracts. Since S will not learn D's type until after offering a contract, S's payoff will be in expectations.

I conjecture that both risk-return trade-off and separating-pooling trade-off will play a role in inefficient fighting in this case. I leave this for future research.

6 Conclusion

The formal international conflict literature makes, but does not explicitly state, an important assumption: Players negotiate according to a pre-specified bargaining protocol in a crisis bargaining game and no player has the ability to change the rules of the game. However, this assumption is hardly satisfied in crisis situations where one of the parties has the ability to set the bargaining protocol on her own.

If one party in a crisis bargaining situation has full bargaining power, then she chooses the bargaining protocol that would benefit her the most. However, there are infinitely many forms of bargaining protocols and moreover a particular protocol may give rise to multiple equilibria. Then what particular bargaining protocol and which equilibrium of this protocol will predict the outcome of such a crisis bargaining game? Game theory is silent on this question except when one of the negotiating parties has full bargaining power. I produce unique predictions for this empirically plausible crisis bargaining scenario by formulating it as a principal-agent problem.

The methodology and the predictions depend crucially on the nature of the informational problem and the identity of the informed party. The two classical assumptions in the existing literature regarding the nature of the informational problem fall into two broad categories: When the informational asymmetry is about a player's cost of war, then a player's private information determines only his own payoff from war. This is the case of private values and the standard principal-agent framework applies. When a player's private information concerns power distribution between players, this information determines his and his opponent's payoffs. This is the case of common values and the problem becomes a principal-agent problem with informed principal.

In the analysis, Fearon's celebrated risk-return trade-off argument arises as a robust prediction when the party that has full bargaining power is uninformed. This finding is independent of the nature of the informational problem. That is, regardless of whether the informational asymmetry concerns the cost of war or power distribution, if the powerful state is not informed, then her risk-return trade-off may cause inefficient fighting. On the contrary, if the informational asymmetry concerns power distribution between parties and the powerful state is informed, then war may break out not as a consequence of a risk-return trade-off but because the possibility of war signals the type of the high-resolve type principal.

The standard principal-agent framework is applied widely in social sciences. However only few applications of the principal-agent problem with informed principal exist. Crisis bargaining is a natural application for both the standard principal-agent framework and the one with informed principal.

References

Banks, J. S. (1990). Equilibrium behavior in crisis bargaining games. *American Journal of Political Science*, 34(3):599–614.

Blainey, G. (1988). *The Causes of War*. New York: The Free Press.

Bolton, P. and Dewatripont, M. (2005). *Contract Theory*. Cambridge, MA: The MIT Press.

de Mesquita, B. B. and Lalman, D. (1992). *War and Reason*. New Haven, CT: Yale University Press.

Fearon, J. D. (1994). Domestic political audiences and the escalation of international disputes. *American Political Science Review*, 88:577–592.

Fearon, J. D. (1995). Rationalist explanations for war. *International Organization*, 49(3):379–414.

Fey, M. and Ramsay, K. W. (2011). Incentives in crisis bargaining: Game-free analysis of international conflict. *American Political Science Review*, 55(1):149–169.

Jervis, R. (1976). *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.

Leventoğlu, B. and Tarar, A. (2008). Does private information lead to delay or war in crisis bargaining? *International Studies Quarterly*, 52(3):533–553.

Maskin, E. S. and Tirole, J. (1990). The principal-agent relationship with an informed principal, i: Private values. *Econometrica*, 58:379–410.

Maskin, E. S. and Tirole, J. (1992). The principal-agent relationship with an informed principal, ii: Common values. *Econometrica*, 60:1–42.

Myerson, R. B. (1979). Incentive compatibility and the bargaining problem. *Econometrica*, 47(1):61–73.

Powell, R. (1996a). Bargaining in the shadow of power. *Games and Economic Behavior*, 15(2):255–289.

Powell, R. (1996b). Stability and the distribution of power. *World Politics*, 48:239–267.

Powell, R. (1999). *In the Shadow of Power: States and Strategies in International Politics*. Princeton, NJ: Princeton University Press.

Reed, W. (2003). Information, power and war. *American Political Science Review*, 97(4):633–641.

Salanie, B. (1999). *The Economics of Contracts*. Cambridge, MA. The MIT Press.

Schultz, K. (1998). Domestic opposition and signaling in international crisis. *American Political Science Review*, 94(4):829–844.

Schultz, K. (1999). Do domestic institutions constrain or inform? contrasting two institutional perspectives on democracy and war. *International Organization*, 53(2):233–266.

Slantchev, B. (2003). The principle of convergence in wartime negotiation. *American Journal of Political Science*, 97(4):621–632.

Smith, A. and Stam, A. C. (2006). Divergent beliefs in 'bargaining and nature of war. *Journal of Conflict Resolutione*, 50(4):614–618.

Van Evera, S. (1999). *Causes of War: Power and the Roots of Conflict*. Ithaca, NY: Cornell University Press.

Wagner, H. (2000). Bargaining and war. *American Journal of Political Science*, 44(3):469–484.

Wittman, D. (1979). How wars end. *Journal of Conflict Resolution*, 23(4):743–763.

Appendix

In this appendix, I will study private values with uninformed principal, private values with informed principal, common values with uninformed principal and common values with informed principal. The first three are straightforward applications. I include them for the sake of completeness.

A Private Values: Cost of War

Assume that D's cost of war is his private information. It is common knowledge that it is either low $c_D = c_l$ with probability π, or high $c_D = c_h > c_l$ with probability $1 - \pi$. Both types of D are dissatisfied with the status quo. That is, $p - c_l > p - c_h > q$. I will refer to D with c_h as "high-cost" or "low-resolve" type and to D with c_l as "low-cost" or "high-resolve" type.

A.1 Private Values with Uniformed Principal

For the time being, I will assume that S's cost of war is common knowledge. However, I will argue later that the analysis and the results remain the same when S's cost of war is her private information. Therefore, the identity of the informed player does not affect either the methodology or the results in the case of private values.

Let $\hat{c}$ denote D's report about his type and $\{\alpha_h, t_h, \alpha_l, t_l\}$ denote a contract. The optimal contract maximizes S's expected payoff subject to individual rationality and incentive compatibility constraints, as described below.

Since D can unilaterally guarantee his war payoff, individual rationality requires that each type of D is offered at least his outside option in a peaceful deal:

$$D_l - IR : t_l \geq p - c_l, \text{ and} \tag{1}$$

$$D_h - IR : t_h \geq p - c_h \tag{2}$$

Incentive compatibility ensures that D will report his type truthfully:

$$D_l - IC : \alpha_l(p - c_l) + (1 - \alpha_l)t_l \geq \alpha_h(p - c_l) + (1 - \alpha_h)\max\{t_h, p - c_l\}$$

$$D_h - IC : \alpha_h(p - c_h) + (1 - \alpha_h)t_h \geq \alpha_l(p - c_h) + (1 - \alpha_l)\max\{t_l, p - c_h\}.$$

The left hand side of $D_l - IC$ is the expected payoff of D with $c_D = c_l$ in case he reports c_D truthfully and the right hand side is his payoff if he reports $c_D = c_h$. In the latter case, war will break out and he will collect his war payoff $p - c_l$ with probability α_h, and he will be offered t_h with probability $1 - \alpha_h$. If he is offered t_h, he will accept the offer if t_h is at least as good as his war payoff, i.e. $t_h \geq p - c_l$. So, his payoff will be the maximum of t_h and $p - c_l$ in that case.

Note that $D_l - IR$ ensures that $t_l \geq p - c_l > p - c_h$ so that $t_l = \max\{t_l, p - c_h\}$ in $D_h - IC$.

If D turns out to have $c_D = c_\tau$, $\tau \in \{h, l\}$, then the two states will go to war with probability α_τ, and they will reach the peaceful settlement t_τ with probability $1-\alpha_\tau$ where S will receive $1-t_\tau$. Here, S will achieve an expected payoff of

$$\alpha_\tau(1-p-c_S)+(1-\alpha_\tau)(1-t_\tau)$$

S does not know D's type when she offers the contract $\{\alpha_h, t_h, \alpha_l, t_l\}$ but she knows that D is a high-resolve type with probability π, so S's expected payoff from offering $\{\alpha_h, t_h, \alpha_l, t_l\}$ is given by

$$\begin{aligned} V(\{\alpha_h, t_h, \alpha_l, t_l\}) = \pi\left[\alpha_l(1-p-c_S)+(1-\alpha_l)(1-t_l)\right] \\ + (1-\pi)\left[\alpha_h(1-p-c_S)+(1-\alpha_h)(1-t_h)\right] \end{aligned}$$

Then S chooses a contract that solves the following maximization problem

$$\max_{\{\alpha_h, t_h, \alpha_l, t_l\}} V(\{\alpha_h, t_h, \alpha_l, t_l\}) \tag{P1}$$

subject to

$D_l - IR,\ D_h - IR,\ D_l - IC,\ D_h - IC$

and $\alpha_h, \alpha_l, t_h, t_l \in [0,1]$

where $\alpha_h, \alpha_l, t_h, t_l \in [0,1]$ are the usual feasibility constraints on probabilities and shares.

A straightforward analysis of this problem yields the optimal solution.

Let $\pi^* = \frac{c_h - c_l}{c_h + c_S} \in (0,1)$. Then optimal contract is given by

$$\text{if } \pi \geq \pi^* \text{ then } \alpha_l = \alpha_h = 0 \text{ and } t_h = t_l = p - c_l$$

$$\text{if } \pi < \pi^* \text{ then } \alpha_l = 1, \alpha_h = 0,\ t_l = p - c_l \text{ and } t_h = p - c_h$$

This is effectively equivalent to Fearon's optimal take-it-or-leave-it offer. If S is sufficiently confident that she is likely to face a high-resolve type D, $\pi \geq \pi^*$, she solves her risk-return trade-off by offering $p - c_l$ to D and avoids war. Otherwise, she takes the risk of war against the high-resolve type by making a low offer.

A.2 Private Values with Informed Principal

S's cost of war may also be her private information. In this case, S's private information does not affect D's payoff, and this constitutes a private values case.

Maskin and Tirole (1990) develop a method to solve the problem of an informed principal in the case of private values. They assume that some types of principals may violate the individual rationality constraint for the agent, however the individual rationality constraint has to hold in expectations (see their problem F_*^i on page 392). In contrast, a fundamental assumption in crisis bargaining is that no player can be forced to accept any deal that is worse than its war payoff. That is, the individual rationality constraint cannot be violated by any type of principal.[4] Therefore, their approach reduces to the standard principal-agent framework in my crisis bargaining model.

Since S's private information does not affect D's payoff, it does not affect the set of individually rational and incentive compatible contracts, either. Thus, the analysis remains the same and the identity of the player that holds private information does not matter when the informational problem is that of private values.

B Common Values: Distribution of Power

In this section, I study the informational problem that concerns distribution of power between players. Assume that c_D and c_S are common knowledge, but p is equal to p_h with probability π and to $p_l < p_h$ with probability $1-\pi$. Although the distribution of p is common knowledge, only one of the players knows the true value of p. The identity of the informed player matters in this case. First, I consider the scenario that D holds private information.

B.1 Common Values with Uninformed Principal

Assume that D privately knows the true value of p. I will refer to D with $p = p_h$ as the high-resolve type and with $p = p_l$ as the low-resolve type. This case is similar to the previous one because S's choice of contract does not transmit

[4] Technically, this means that the only feasible value for r^i is zero in their problems F_*^i and V_I^i on page 392.

information from S to D, and so S's problem is set up as in (P1). S chooses an individually rational and incentive compatible contract $\{\alpha_h, t_h, \alpha_l, t_l\}$ that maximizes her expected payoff among all individually rational and incentive compatible contracts. Her problem is formulated as follows:

$$\max_{\alpha_l, t_l, \alpha_h, t_h} V(\alpha_l, t_l, \alpha_h, t_h) = \pi \left[\alpha_h(1 - p_h - c_S) + (1 - \alpha_h)(1 - t_h)\right] \qquad \text{(P2)}$$
$$+ (1 - \pi) \left[\alpha_l(1 - p_l - c_S) + (1 - \alpha_l)(1 - t_l)\right]$$

subject to

$$D_l - IR : t_l \geq p_l - c_D,$$
$$D_h - IR : t_h \geq p_h - c_D,$$
$$D_l - IC : \alpha_l(p_l - c_D) + (1 - \alpha_l)t_l \geq \alpha_h(p_l - c_D) + (1 - \alpha_h) \max\{t_h, p_l - c_D\}$$
$$D_h - IC : \alpha_h(p_h - c_D) + (1 - \alpha_h)t_h \geq \alpha_l(p_h - c_D) + (1 - \alpha_l) \max\{t_l, p_h - c_D\}$$

and the usual feasibility constraints $\alpha_i \in [0, 1]$ and $t_i \in [0, 1]$.

The optimal contract is given as follows: Let $\pi^{**} = \frac{p_h - p_l}{(p_h - p_l) + (c_S + c_D)} \in (0, 1)$. Then

if $\pi \geq \pi^{**}$ then $\alpha_l = \alpha_h = 0$ and $t_h = t_l = p_h - c_D$

if $\pi < \pi^{**}$ then $\alpha_l = 0, \alpha_h = 1$, $t_l = p_l - c_D$ and $t_h = p_h - c_D$

This contract is effectively equivalent to a take-it-or-leave-it offer. If S is sufficiently confident that she is likely to face a high-resolve type D, $\pi \geq \pi^{**}$, she solves her risk-return trade-off by offering $p_l - c_D$ to D and thereby avoids war. Otherwise, she takes the risk of war against the high-resolve type D by making a low offer. Fearon's prediction of risk-return trade-off arises in this case as well.

B.2 Common Values with Informed Principal

I only provide the solution of the optimal separating equilibrium here in the appendix. The solution of the optimal pooling equilibrium and the optimal equilibrium are in the main text.

Lemma 6 $\alpha_l^{sep} \geq \alpha_h^{sep}$

Proof. Summing up $S_l - IC$ and $S_h - IC$ yields $\alpha_l^{sep}(p_h - p_l) \geq \alpha_h^{sep}(p_h - p_l)$. Then $p_h > p_l$ implies $\alpha_l^{sep} \geq \alpha_h^{sep}$. ■

Lemma 7 $t_h^{sep} \geq t_l^{sep}$

Proof. By individual rationality for S, it must be the case that $1 - t_l^{sep} \geq 1 - p_l - c_S$. Then $\alpha_l^{sep} \geq \alpha_h^{sep}$ and $S_l - IC$ imply that $1 - t_l^{sep} \geq 1 - t_h^{sep}$, which is equivalent to $t_h^{sep} \geq t_l^{sep}$. ■

Lemma 8 $\alpha_h^{sep} = 0$

Proof. Suppose that $\alpha_h^{sep} > 0$. Then $\alpha_l^{sep} > 0$. Decrease α_h^{sep} by some small $\epsilon > 0$ and α_l^{sep} by some small $\delta > 0$ such that

$$\epsilon(p_h + c_S - t_h) = \delta(p_h + c_S - t_l)$$

Individual rationality for S implies $1-t_h^{sep} \geq 1-p_h-c_S$, equivalently $p_h+c_S \geq t_h^{sep}$. Also $t_h^{sep} \geq t_l^{sep}$ from the previous lemma so that the coefficients of ϵ and δ are both nonnegative. Then $S_h - IC$ and all other constraints continue to hold and V_h^{sep} and V_l^{sep} increase. This is a contradiction so $\alpha_h^{sep} = 0$. ■

Lemma 9 *$S_h - IC$ holds with equality.*

Proof. Suppose that $S_h - IC$ is slack. If $\alpha_l^{sep} > 0$ or $t_l^{sep} > p_l - c_D$, then slightly decreasing α_l^{sep} or t_l^{sep} increases V_l^{sep} without violating any of the constraints. So $\alpha_l^{sep} = 0$ and $t_l^{sep} = p_l - c_D$ must hold. Then $S_h - IC$ becomes $1 - t_h^{sep} \geq 1 - p_l + c_D$, equivalently $t_h^{sep} \leq p_l - c_D$. But this is a contradiction because $t_h^{sep} \geq p_h - c_D$ by $D_h - IR$ and $p_h - c_D > p_l - c_D$. Then $S_h - IC$ holds with equality. ■

Lemma 10 *$S_l - IC$ is slack.*

Proof. Substitute $\alpha_h^{sep} = 0$ in $S_h - IC$ and $S_l - IC$. Then

$$\begin{aligned} 1 - t_h^{sep} &= \alpha_l^{sep}(1 - p_h - c_S) + (1 - \alpha_l^{sep})(1 - t_l^{sep}) \\ &< \alpha_l^{sep}(1 - p_l - c_S) + (1 - \alpha_l^{sep})(1 - t_l^{sep}) \end{aligned}$$

where the equality is the $S_h - IC$ constraint and the inequality is implied by $p_l < p_h$. But the inequality is $S_l - IC$ so $S_l - IC$ is slack. ■

The last two Lemmata prove Result 7.

Lemma 11 $t_h^{sep} = p_h - c_D$

Proof. If $t_h^{sep} > p_h - c_D$ then slightly decrease t_h^{sep}. $D_h - IR$ and $S_l - IC$ continue to hold for a small enough decrease, $D_l - IR$ is not affected, $S_h - IC$ becomes slack and V_h^{sep} increases. This is a contradiction so $t_h^{sep} = p_h - c_D$. ■

So the solution to the problem of the low-resolve S is given by $\alpha_h^{sep} = 0$ and $t_h^{sep} = p_h - c_D$.

The problem of the high-resolve type *becomes*

$$(\alpha_l^{sep}, t_l^{sep}) \text{ solves } \max_{(\alpha_l, t_l)} \alpha_l(1 - p_l - c_S) + (1 - \alpha_l)(1 - t_l)$$

$$\text{subject to}$$

$$D_l - IR : t_l \geq p_l - c_D$$

$$S_h - IC : 1 - p_h - c_S = \alpha_l(1 - p_h - c_S) + (1 - \alpha_l)(1 - t_l)$$

Take the total differential of $S_h - IC$ with respect to α_l and t_l:

$$(p_h + c_S - t_l)d\alpha_l + (1 - \alpha_l)dt_l = 0$$

Take the total differential of the objective function with respect to α_l and t_l and replace the above equality:

$$\begin{aligned} d(\text{objective}) &= (t_l - (p_l + c_S))d\alpha_l - (1 - \alpha_l)dt_l \\ &= (p_h - p_l)d\alpha_l \end{aligned}$$

So increasing α_l and decreasing t_l increases the objective. Then the solution to the problem of the high-resolve type S is given by $t_l^{sep} = p_l - c_D$, and substitute that in $S_h - IC$ to obtain

$$\alpha_l^{sep} = \frac{p_h - p_l}{(p_h - p_l) + (c_D + c_S)}$$

C Take-it-or-leave-it offer

Signalling Equilibrium

First, I will show that D's strategy is optimal given his beliefs.

Rejecting any offer $t < t_l$ is optimal for D since he can guarantee t_l by rejecting it.

Suppose that D receives an offer of $t \geq t_h$. Then he believes that $p = p_h$ with probability 1. Accepting the offer is optimal for him because rejecting it provides him with the same payoff $t_h = p_h - c_D$.

If D receives an offer of $t \in (t_l, t_h)$, then he believes that $p = p_H$ and his expected payoff from rejecting t is $t_h = p_h - c_D$, which is greater than t. So rejecting t is optimal.

If D receives an offer of t_l, then he believes that $p = p_l$ with probability ϕ. Then he is indifferent between accepting and rejecting the offer, so any mixed strategy is optimal for him.

By the definition of α, S_h is indifferent between offering t_l and t_h:

$$t_h = \alpha(1 - t_h) + (1 - \alpha)(1 - t_l)$$

So it is optimal for S_h to offer t_h. Since $t_l < t_h$, this equality implies

$$t_h < \alpha(1 - t_l) + (1 - \alpha)(1 - t_l)$$

so that it is optimal for S_l to offer t_l. D's beliefs are consistent with S's strategy on the equilibrium path.

Bluffing Equilibrium

First, I will show that D's strategy is optimal given his beliefs.

Suppose that D receives an offer of $t \geq t_h$. Then he believes that $p = p_h$ with probability 1. Accepting the offer is optimal for him because rejecting it provides him with the same payoff $t_h = p_h - c_D$.

If D receives an offer of $t \in [0, p_h - c_D] \backslash \{t_l, t_h\}$, then he believes that $p = p_H$ and his expected payoff from rejecting t is $t_h = p_h - c_D$, which is greater than t. So rejecting t is optimal.

If D receives an offer of t_l, then he believes that $p = p_h$ with probability ϕ. To accept this offer with probability $\alpha \in (0, 1)$, he must be indifferent between accepting and rejecting the offer:

$$t_l = \phi(p_h - c_D) + (1 - \phi)(p_l - c_D)$$

where the left hand side is D's payoff from accepting the offer and the right hand side is his expected payoff from rejecting it given his beliefs. This equality holds by definition of t_l, and it is optimal for D to mix between accepting and rejecting the offer.

Next I will show that S's strategy is optimal given D's strategy and beliefs.

Suppose $p = p_h$. If S_h offers $t \in [0, p_h - c_D] \backslash \{t_l, t_h\}$ then D fights with probability 1 so S_h's payoff from offering t is $1 - p_h - c_S$. Her payoff from offering $t_h = p_h - c_D$ is $1 - p_h + c_D$, since D accepts it with probability 1. So offering $t \in [p_l - c_D, p_h - c_D] \backslash \{t_l, t_h\}$ with positive probability cannot be optimal for S_h and she either offers t_l or t_h. For S_h to bluff by offering t_l with positive probability of $\beta \in (0, 1)$, she must be indifferent between offering t_h and t_l:

$$1 - p_h + c_D = \alpha(1 - t_l) + (1 - \alpha)(1 - p_h - c_S)$$

where the left hand side is S's payoff from offering t_h, which D accepts with probability 1, and the right hand side is her expected payoff from offering t_l, which D accepts with probability α. This equality holds by definition of α so it is optimal for S_h to bluff with positive probability.

Suppose that $p = p_l$. If S_l offers t_l, her payoff is

$$\alpha(1 - t_l) + (1 - \alpha)(1 - p_l - c_S)$$

If S offers $1 \geq t_h$, her payoff is less than or equal to $1 - p_h + c_D$ since D accepts $t \geq t_h$ with probability 1. By definition of α, this payoff is less than her payoff from offering t_l. So S_l does not offer $t \geq t_h$. If she offers $t \in [p_l - c_D, p_h - c_D] \backslash \{t_l, t_h\}$ or $t < t_l$, then her payoff is $1 - p_l - c_S$ because then D fights with probability 1. This payoff is less than her payoff from offering t_l if and only if $c_S + c_D \geq \pi(p_h - p_l)$, which I have assumed. So it is optimal for S_l to offer t_l with probability 1.

Finally, I confirm the consistency of D's beliefs on the equilibrium path. S offers t_h only when $p = p_h$, so it must be the case that $p = p_h$ after observing t_h. Since both types of S offer $t_{l,}$ D's belief after receiving t_l must follow Bayes' rule:

$$\Pr(p = p_h | t = t_l) = \phi = \frac{\pi\beta}{\pi\beta + (1 - \pi)} \in [0, \pi]$$

where, given the equilibrium strategies, the numerator is the ex ante probability that $t = t_l$ will be offered by S_h and the denominators is the ex ante probability that $t = t_l$ will be offered. Thus, D's beliefs are consistent with the Bayes' rule on the equilibrium path. Since no offer of $t \in [o, 1]\{t_l, t_h\}$ will be made on the equilibrium path, D's beliefs after receiving such offer can be arbitrary.