Essays in Mechanism Design

by

Eric Gonzalez Sanchez

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2020 by the
Graduate Supervisory Committee:

Alejandro Manelli, Chair
Hector Chade
Edward Schlee

ARIZONA STATE UNIVERSITY

May 2020

ABSTRACT

I study split-pie bargaining problems between two agents. In chapter two, the types of both agents determine the value of outside options – I refer to these as interdependent outside options. Since a direct mechanism stipulates outcomes as functions of agents' types, a player can update beliefs about another player's type upon receiving a recommended outcome. I term this phenomenon as information leakage. I discuss binding arbitration, where players must stay with a recommended outcome, and non-binding arbitration, where players are not obliged to stay with an allocation. The total pie is reduced if the outcome is an outside option. With respect to efficiency, I derive a necessary and sufficient condition for first best mechanisms. These are mechanisms that assign zero probability to outside options for every report received. The condition describes balanced forces in conflict (outside options) and is the same in the cases of binding and non-binding arbitration. I also show a strong link between conflict and information: when conflict exists, information leakage occurs. Hence, non-binding arbitration may seem more restrictive than binding arbitration. To analyze why this is the case, I solve for second best mechanisms with binding arbitration and find a condition under which they can be implemented under non-binding arbitration. Thus, I show that non-binding arbitration can be as effective as binding arbitration in terms of efficiency. I also examine whether the equivalence between binding and non-binding arbitration can cease to hold, and provide analysis of why this happens. In chapter three, the bargaining problem entails no uncertainty but rather envy. Players can feel envy about the allocation of the other player. The Nash Bargaining solution is obtained in this context and some comparative statics are shown. The introduction of envy makes the more envious party a tougher negotiator.

DEDICATION

Gracias a mi familia, por todo.

Toda virtud en este trabajo se debe completamente a mi familia.

Todo error es completamente mío.

I thank my family, for everything.

Every virtue in this work is due entirely to my family.

Every mistake is due entirely to me.

ACKNOWLEDGEMENTS

I, first, thank my family for everything. I also thank my entire committee for staying with me until this moment.

TABLE OF CONTENTS

Page

LIST OF FIGURES v

CHAPTER

1 INTRODUCTION 1

2 BARGAINING WITH INTERDEPENDENT OUTSIDE OPTIONS 5

2.1 Introduction 5

2.2 The Model 10

2.3 The First Best 20

2.4 Beyond The First Best 30

2.5 Conclusions 46

2.6 Appendix 52

3 NASH BARGAINING SOLUTION UNDER EXTERNALITIES, WITH GINO LOYOLA 61

3.1 Introduction 61

3.2 Model 63

3.3 Results 65

3.4 Concluding Remarks 67

3.5 Appendix 68

4 CONCLUSIONS 72

REFERENCES 75

APPENDIX

A STATEMENT 79

LIST OF FIGURES

Figure Page

2.1 Graphical Depiction of Theorem 2. 28

2.2 The Solution to the Binding Arbitration Problem. 37

3.1 Nash Bargaining Solution under Externalities. 64

Chapter 1

INTRODUCTION

In this work I consider two different bargaining problems in which two agents want to split a fixed and known "pie." In chapter two, agents have private information about their type but types do not affect the size of the pie; both agents' types simultaneously determine the outside option for each agent. I refer to this situation as interdependent outside options. This model is motivated by contexts in which the outside option, the failure to reach agreement, is the outcome of war[1]. In such context, it is natural to think that outside options depend on both players military capacity, and the precise military capacity of each player is the player's type. It is also natural, to think that players can always choose to go to war, even after they have accepted an allocation, peacefully.

I adopt a mechanism design perspective to analyze the mentioned problem. I call non-binding arbitration, the case where the designer cannot force players to follow the recommendation. I call binding arbitration, the case where the designer can force the players to follow the recommendation. Both cases are characterized by distinct incentive compatibility and individual rationality conditions. Appealing to the version of the revelation principle in Myerson (1982), there is no loss of generality in focusing on direct revelation mechanisms in binding and non-binding arbitration. In these mechanisms, agents satisfy incentive compatibility and obedience constraints. A mechanism consists of a function that, given both agents' reports, recommends a lottery of allocations of the pie with certain probability, and the outside option with

[1]I thank Carlos Rodriguez-Sickert, Paul Seabright, Miguel Fuentes and Ricardo Guzman, for a joint project and conversations that helped me motivate the problem.

some (complementary) probability. Since the mechanism stipulates outcomes as a function of agents' types, players may update their beliefs about the other player's type upon receiving a recommendation. I refer to this, as "information leakage." Hence, the presence of interdependence in outside options, brings about relevance to the enforcement power of the mechanism designer: if the designer cannot force players to follow a recommendation, the outside option (after updating its value with the updated beliefs) may be claimed by a player.

I am interested in ex ante efficient mechanisms. First, I identify a condition on outside options and distributions of types that is necessary and sufficient for the existence of first best mechanisms. That is to say, mechanism that assign zero probability to the outside option and allocate the pie completely for all reports. The efficiency condition, essentially requires that payoffs from outside options are not "too far apart" in expected terms. I interpret this as balanced forces in conflict. This condition applies equally to binding and non-binding arbitration. In addition, the information leakage phenomenon identified earlier is nonexistent when the efficiency condition holds, because direct mechanisms can be chosen so that no information is revealed by the recommended allocations. Conversely, if first best mechanisms are not implementable, some information is revealed by the allocation proposed by the mechanism. This illustrates a strong link between conflict and information: when there must be conflict there must be information leakage. Second, I identify the optimal mechanisms when the efficiency condition described earlier does not hold. I find the class of second best mechanisms with binding arbitration and identify a sufficient condition under which those mechanisms can be implemented under non-binding arbitration. Thus, I show that non-binding arbitration can be as effective as binding arbitration in some environments. Those environments, are characterized to have very inefficient outside options. In the context of the example of war, the threat of very destructive war, can help the designer's objective.

Chapter three corresponds to a joint work with Gino Loyola, Associate Professor at Department of Management Control and Information Systems, University of Chile [2]. The work is submitted to International Journal of Game Theory. In chapter three, a split-pie bargaining problem is analyzed where now there is no uncertainty, but players can feel envy about the allocation received by the other player. The Nash Bargaining Solution (NBS) is obtained and some properties are discussed. My contribution to such work was in the formal development of the proofs and the exposition as well as discussion of results and possible extensions. The original idea of the paper comes from previous published work of Gino Loyola and discussions with me. Alternative approaches were taken by my coauthor to contrast the result of the proofs I carried out. The paper seeks to analyze, from a theoretical perspective, the role played by feelings and emotions. This has been increasingly discussed in the literature of experimental economics. See, for instance, Güth and Kocher(2014), Fischbacher et al. (2013); Kagel and Wolfe (2001); Pfister and Böhm (2012); Pillutla and Murnighan (1996); Sanfey et al. (2003); van 't Wout et al. (2006). The theoretical literature has mainly focused on possibly not included factors in certain particular negotiation protocols. The paper contributes to the literature by analyzing the impact of envy in a theoretical model where a decentralized solution is proposed. This has the virtue of the analysis to be robust to particular game forms, while satisfying certain properties desired as in the NBS.

The Nash Bargaining solution in this context is characterized and some properties are analyzed. Envy plays a role in making the possibility to reach an agreement difficult. However, conditional on reaching an agreement, it makes a more envious

[2]Department of Management Control and Information Systems, University of Chile, Diagonal Paraguay 257, Of. 1903, 8330015, Santiago, Chile. E-mail: gloyola@fen.uchile.cl, Tel: +56 (2) 22 978 37 68.

player to be a tougher bargaining, giving her/him more bargaining power.

Chapter 2

BARGAINING WITH INTERDEPENDENT OUTSIDE OPTIONS

2.1 Introduction

I study bargaining between two agents over a fixed value, "the pie". The environment exhibits the following characteristics. First, the types of the agents determine the value of their respective outside options – I refer to these as interdependent outside options. The type of an agent is private information for that agent. Second, outside options are inefficient in the sense that the total pie is reduced if the outcome is an outside option, that is, if "negotiation breaks down." I study this problem from a mechanism design perspective. The mechanism designer's objective is efficiency. In particular, the designer seeks to maximize the ex ante expected sum of utilities that agents obtain from participating in the mechanism.

The features described introduce the possibility that a proposed interim outcome may not be enforceable ex post. Note that a mechanism stipulates outcomes as a function of agents' reports. Once reports are submitted, a proposed outcome is announced. Thus the announcement potentially reveals information to each player about the type of the other player, and hence, about the value of the outside option. I refer to this phenomenon as information leakage. In the context of new information, a player may choose not to respect a recommended interim outcome. Thus I consider environments where a mechanism designer can force players to follow recommended outcomes (I call this binding arbitration), and environments where mechanisms must be self-enforcing, that is, the mechanisms may only propose outcomes that both players accept after updating their information (I call this non-binding arbitration).

Examples of the environment described above abound and have been considered extensively in the literature. See, for instance, war applications as in Brito and Intriligator (1985), Hörner et al. (2015), Polacheck and Xiang (2010), and Reed (2003). Industrial organization applications occurin Das and Sengupta (2001) and Goltsman and Pavlov (2014). Political campaigns applications are explored in Skarpedas and Grofman (1995), among others.

As a motivating example, consider two countries negotiating over how to split a disputed territory. The outside option is an armed conflict, potentially rendering part of the land less usable (i.e., the pie is diminished). Each country has private information about its war capabilities, but the value of the outside option (i.e., the outcome of the war) depends on both countries' capabilities. It is not uncommon to seek non-binding arbitration in cases such as this. The case described represents non-enforceability of a recommendation. In the literature, non-enforceability is sometimes called veto rights (see, for instance, Forges (1999) and Compte and Jehiel (2009)).

The natural designer's objective in this case is ex ante efficiency. The designer will seek to minimize armed conflict in general but may favor some scenarios over others. For instance, if two strong military forces enter armed conflict, an observer can expected that a large portion of land will be destroyed. Ideally, the designer would like to avoid conflict completely. In this sense, the "first best" option consists of mechanisms that assign zero probability to outside options and allocate the pie completely, for every report received. I identify a necessary and sufficient condition on the environment for the existence of a first best mechanism. The condition can be interpreted as balanced forces in conflict. The condition is the same in binding and non-binding arbitration. I show that an important feature of first best mechanisms is that they can be taken as withholding information through a proposed outcome.

In contrast, if first best mechanisms are not implementable, information is revealed by the proposed outcomes. This suggests that in the scenario described, the designer may face more restrictions under non-binding arbitration than under binding arbitration, from the standpoint of efficiency. When first best mechanisms are not implementable, I study the mechanisms that maximize the ex ante expected sum of the utilities from participating in the mechanism, considering only those outcomes that are implementable. I call these second best mechanisms. I identify second best mechanisms for the binding arbitration problem, in order to see if non-binding arbitration is more restrictive than binding arbitration. I provide a sufficient condition, under which a designer conducting non-binding arbitration achieves as much as a designer under binding arbitration in terms of the expected ex ante sum of utilities of the agents participating in the mechanism. The condition resembles that of first best mechanisms, suggesting that balanced forces in conflict can help alleviate the information leakage problem. I analyze properties of optimal mechanisms, in particular, leakage of information, and examine whether equivalence between binding and non-binding arbitration holds in general.

For ease of highlighting the contributions of this paper relative to the literature, I use two concepts as my main ingredients: interdependence of outside options and non-binding arbitration. Some of the literature has focused on interdependence, but mainly in the case of binding arbitration. See, for instance, Harsanyi and Selten (1972) and Myerson (1984), who are interested in ex ante efficiency where interdependence plays a role in valuations. In the same vein, McLean and Postlewaite (2015) find conditions to extend VCG mechanisms to interdependent valuations, retaining the ex post incentive compatibility property. In such papers, there is no problem of enforcement: it is assumed that players must stay with proposed outcomes regardless of whether they receive information that updates their beliefs.

A second strain of literature focuses on non-binding arbitration with independence, with players enjoying veto rights. See, for instance, Forges (1999), Shimer and Werning (2015), and Compte and Jehiel (2009). The designer cannot enforce an outcome ex post, because players can veto potential recommendations even after they have agreed to participate in the mechanism. The value of the object or outside option to a player depends only on that player's type. This leaves space for a third school of thought in the literature, one that analyzes interdependent outside options and non-binding arbitration. Two papers in this school are Cramton and Palfrey(1995) and Hörner et al. (2015). Cramton and Plafrey(1995) analyze the problem from a game theoretical perspective in which there are two stages. The first stage is a ratification stage, where every player can veto the proposed mechanism. This is followed by an implementation stage, in which either the proposed mechanism or the status quo is adopted, depending on the first stage result. The authors propose the concept of ratifiable mechanisms, defined as mechanisms for which unanimous ratification is a sequential equilibrium of a two-stage game. In works such as these, there is no objective role for the designer. The authors' concern is with the existence of ratifiable mechanisms, and how to describe those that are objective versus those that are not.

The approach of Hörner et al. (2015) is the closest to that used in the current paper. The first difference is that the 2015 model considers only two types and assumes constant inefficiency from outside options (that is, a fixed portion of the pie is lost in conflict). I assume a continuum of types and no restriction on the inefficiency of outside options. Given the assumption of constancy of inefficiency in the 2015 paper, the authors look for mechanisms that minimize the ex ante probability of conflict. In the present paper, I examine the set of ex ante efficiency mechanisms that would coincide only when inefficiency from outside options is constant. In Hörner et al. (2015), the main interest is in comparing binding and non-binding arbitration in terms of ex ante probability of conflict. They show that in their model, both types of arbitration can produce the same ex ante probability of conflict. In this paper, I examine ex ante efficiency. I demonstrate the necessary and sufficient condition for first best mechanisms to be implementable. I show their properties of concealing information. Then I study second best mechanisms, solve the binding arbitration problem, and show a condition making second best mechanisms implementable with non-binding arbitration. The condition resembles that for first best mechanisms. From the contribution of Hö et al. (2015), the question of “equivalence” between binding and non-binding arbitration is open, to the extent that we do not know whether it is a function of the two types model or if it can hold more generally. I conclude my paper with a discussion of why and how “equivalence” can cease to hold.

The problem analyzed is important for two reasons. First, from a theoretical point of view, the interaction of non-binding arbitration and interdependent outside options brings a natural information component to the design of the mechanism. The information that players obtain is endogenously determined by the mechanism. This has not been explored in the literature (see Hörner et al. (2015)). An “equivalence” between binding and non-binding arbitration is also important from a technical point of view. Binding arbitration, although still complicated by interdependence, is a more tractable problem than non-binding arbitration, so restricting attention to binding arbitration is valuable at the technical level.

Second, the problem analyzed is of importance from the point of view of applications. Take the war example for instance. Avoidance of war is crucial for healthy society. In the context of international conflict, arbitration is naturally thought to be non-binding, since countries cannot be forced to accept given recommendations. My result showing that non-binding arbitration can be as “effective” in terms of efficiency as binding arbitration implies that in these types of problems, there is no efficiency lost if one pursues only non-binding arbitration. Non-binding arbitration may be appealing to actors facing potential conflict, as they could resent the enforcement power of a mediator, or find it restrictive. More generally, any problem in which conflict avoidance is desirable could benefit from my analysis.

In section 2.2, I describe the model and introduce non-binding as well as binding arbitration problems. In section 2.3, I characterize first best mechanisms in both binding and non-binding arbitration. I show a necessary and sufficient condition for mechanisms to be implementable, called the efficiency condition. The condition is the same for both types of arbitration. I also show a link between information leakage and conflict. In section 2.4, I depart from the efficiency condition and solve for second best binding arbitration mechanisms. I describe a condition under which second best mechanisms from binding arbitration can be implemented in non-binding arbitration. Thus, I show that non-binding arbitration can be as "effective" in terms of efficiency as binding arbitration. I discuss the reasons for which the equivalence could fail to hold. Section 2.5 reports conclusions and charts several paths for future work.

2.2 The Model

I consider a bargaining problem where there is a pie of size 1, to be divided among two players. Each player $i \in \{1, 2\}$ has a type $\theta_i \in \Theta_i$, which is private information for that player. I assume $\Theta_i = [\underline{\theta}, \overline{\theta}]$ for all $i = 1, 2$, however, I keep notation the Θ_i, in order to make exposition clear, when necessary. I assume that types distribute independently. For all $i \in \{1, 2\}$, $f_i : \Theta_i \to \mathbb{R}$ denotes the density function for the types of player i, and F_i denotes the respective cumulative distribution function. I assume f_i is continuous and $f(\theta_i) > 0$ for all $\theta_i \in \Theta_i$. I refer to $\Theta := \Theta_1 \times \Theta_2$ as the type space and an element of Θ to be $\theta = (\theta_1, \theta_2)$. That is, I assume the first coordinate of a vector of types represents player 1 and the second one, player 2, unless I write specific subindices. The outside option of player $i \in \{1, 2\}$ is given by the function $u_i : \Theta \to \mathbb{R}$. That is, the outside option of player i depends on the types of both players. I call this, interdependence of outside options. I further assume that u_1 is strictly increasing in the first argument and strictly decreasing in the second. Similarly, u_2 is strictly decreasing in the first argument and strictly increasing in the second argument. In words, the outside option of a player is strictly increasing in his own type and strictly decreasing in the other player's type.

This assumption is in line with the examples mentioned in the introduction. In the war example, the higher the military capacity of a country (higher type), the better the outcome of war for it. And the stronger one's opponent in a war, the worse the outcome for oneself. In an example of collusion of firms, a disagreement to split monopoly profits leads to market competition (outside option). The ability of a firm to succeed in market competition obviously improves its profit, and at the same time lowers the payoffs of the other firm. Similarly, the strength of the political base of a party increases the probability of its passing a bill. However, in the face of stronger political opposition, the probability of passing such abill decreases.

I assume that for any (θ_1, θ_2), $u_1(\theta_1, \theta_2) + u_2(\theta_1, \theta_2) < 1$. I call this the inefficient outside options assumption: for any realization of types, the outside options never split the whole pie (size 1). This corresponds with the inefficiency in the examples. During a war, some land may become barren, or devastation may make reconstruction costly. When parties compete to pass a bill, there are costs of campaigning. Likewise, when firms compete in the market (for instance competing on innovation), the cost of innovation could be reduced if only one firm produces (there would not be a need for differentiation). Finally, I assume the utility for any player of receiving a share of the pie $x \in [0,1]$ is simply x. This is consistent with our motivating examples, the types of players, effectively, play a role only if the players choose their outside options. In the example of war, military capacity does not make a difference at the moment of valuing certain shares of land whose status is agreed upon peacefully. In the case of collusive arrangements for firms, the cost of innovation is not present when splitting monopoly profits. For competing bills, the strength of the political base does not play a role once a bill has been passed.

Let d^*, represent the outside option outcome and let $A = \{(x_1, x_2) : x_1 + x_2 \leq 1$ and $(x_1, x_2) \in [0,1]^2\}$. The set of all possible outcomes is $X = d^* \cup A$. While the value of any element in A is known a priori, the value to the players of d^* depends on both players' types. I study the bargaining problem from a mechanism design perspective. A direct mechanism specifies, for each report $(\theta_1, \theta_2) \in \Theta$, a distribution over the set of outcomes, X. I represent such a mechanism with a pair of functions (π, g), $\pi : \Theta \to [0,1]$ and $g : A \times \Theta \to [0,1]$. The value $\pi(\theta_1, \theta_2)$ indicates the probability assigned to the outside options outcome (d^*), when the report is (θ_1, θ_2). The value $g((x_1, x_2)|\theta_1, \theta_2)$ is the probability (or density, depending on how it is designed) of the outcome $(x_1, x_2) \in A$. The restriction for the direct mechanism to be a distribution over the set of outcomes X is: $\int_A g((x_1, x_2)|(\theta_1, \theta_2))dx_1 dx_2 = 1 - \pi(\theta_1, \theta_2)$, for all $(\theta_1, \theta_2) \in \Theta$. For ease of exposition, for each $i = 1, 2$, $j \neq i$, I define the function $g_i : [0,1] \times \Theta \to [0,1]$, by $g_i(x_i|(\theta_1, \theta_2)) := \int_{[0,1]} g((x_i, x_j)|(\theta_1, \theta_2))dx_j$. Also, whenever I write i and j, it shall be understood as $i \neq j$.

In some scenarios, such as international conflict, it is natural to think that countries always have the option of going to war, even after they have agreed to participate in arbitration and received a recommended outcome. I call **non-binding arbitra-**

tion, the case where the designer cannot force the players to follow the recommendation. I call **binding arbitration** the case where the designer can force players to follow the recommendation. In Hörner et al. (2015), a distinction is made between mediation and arbitration. In their paper, arbitration is what I call binding arbitration and mediation corresponds to what I refer to as non-binding arbitration. The enforcement power of the designer determines distinct incentive compatibility and individual rationality constraints, as we shall see later. Given a direct mechanism (π, g), it is useful to understand the case of non-binding arbitration by thinking in the following way.

First, each player $i \in \{1, 2\}$ chooses to report $\theta_i \in \Theta_i$ or not to participate in the mechanism. In the latter case, the outside option outcome occurs. Once a report is made, it is only observed by the mechanism designer.

Second, if $(\theta_1, \theta_2) \in \Theta$ has been reported, the mechanism recommends either the outside option, with probability $\pi(\theta_1, \theta_2)$ or an allocation $(x_1, x_2) \in A$, drawn from $g(\cdot|(\theta_1, \theta_2))$.

Third, if the allocation (x_1, x_2) has been recommended, each player can accept the recommendation or claim his outside option. It suffices that one of them claims his/her outside option for the outside option outcome to occur.

Some features of the described timing are noteworthy. First, a player's report is only observed by the designer. This is relevant for non-binding arbitration with interdependent outside options. If reports were publicly announced, truthful reports would imply that all agents know their outside option precisely. Since players can claim their outside option, even after participating in the mechanism, they would only accept recommendations that are ex post individually rational. This restriction is significant. It is worth noting that this would be the case with independent outside options (as in Compte and Jehiel (2009)). When outside options depend only on a player's own type, non-binding arbitration forces the designer to satisfy restrictive ex post individually rational allocations. This need not be the case with interdependent outside options. In particular, when first best mechanisms are implementable, the designer can conceal all information about the players' types. Thus, individual rationality constraints are simply satisfied at the interim level.

Second, the designer is able to propose a randomization over the set of allocations for a given pair of types reported. This is relevant when comparing binding and non-binding arbitration. Indeed, as illustrated later, for binding arbitration it is without loss of generality to focus on deterministic allocations for a given pair of types reported. For non-binding arbitration, however, randomizing allocations can make a significant difference. For instance, the designer could implement an optimal binding arbitration mechanism, with non-binding arbitration, by randomizing allocations (see Hörner et al. (2015)). In this case, if the deterministic allocation from the binding arbitration is proposed, it would not be accepted by the highest type. This, because the allocation prescribed is less than the updated outside option. Randomization of allocations helps the designer to induce a "less precise" posterior to the highest type, such that he/she accepts each of the values of the lottery of allocations that can be prescribed for the highest type.

Third, the last part of the timing captures lack of enforcement power, what I call non-binding arbitration. The incentive compatibility (IC) and individual rationality (IR) constraints must take into account the fact that after any received recommendation, the players can still claim their outside options. Note that this applies not only to allocations received when reporting true type (IR constraints), but also to allocations obtained from misreported types (IC constraints).

For non-binding arbitration I use the version of the revelation principle in Myerson (1982). Hence, it is without loss of generality to focus on direct mechanisms that satisfy incentive compatibility and obedience constraints. I call **non-binding arbitration** IC and **non-binding arbitration** IR the corresponding constraints. For binding arbitration, using the revelation principle in Myerson (1979), we can focus on "standard" direct mechanisms, satisfying truthful reports. Thus these direct mechanisms need to satisfy what I call **binding arbitration** IC and **binding arbitration** IR.

A direct mechanism (π, g) satisfies **non-binding arbitration** IC, if for each $i = 1, 2$ and each $\theta_i^* \in \Theta_i$

$$\mathbb{E}_{\theta_j^*}\left[\pi(\theta_i^*, \theta_j^*)u_i(\theta_i^*, \theta_j^*) + \int_{[0,1]} x g_i(x|(\theta_i^*, \theta_j^*))dx\right] \geq$$
$$\mathbb{E}_{\theta_j^*}\left[\pi(\theta_i, \theta_j^*)u_i(\theta_i^*, \theta_j^*) + \int_{[0,1]} \max\{x, \mathbb{E}_z[u_i(\theta_i^*, z)|x, \theta_i]\} g_i(x|(\theta_i, \theta_j^*))dx\right],$$
$$\text{for all } \theta_i \in \Theta_i$$

Where, $\mathbb{E}_z[u_i(\theta_i^*, z)|x, \theta_i] = \int_{\Theta_j} u_i(\theta_i^*, z)\dfrac{f(z)g_i(x|(\theta_i, z))}{\int_{\Theta_j} f(\theta_j)g_i(x|(\theta_i, \theta_j))d\theta_j}dz$, if $\int_{\Theta_j} f(\theta_j)g_i(x|(\theta_i, \theta_j))d\theta_j > 0$ and $\mathbb{E}_z[u_i(\theta_i^*, z)|x, \theta_i] = 0$, otherwise.

Notice that there are two choices for players in the non-binding arbitration IC constraint. The first is the choice of what type to report. The second is whether to accept or not the allocation that is being proposed. The latter choice is done by comparing the allocation proposed with the conditional expectation of the outside option of the player. This conditional expectation is computed using Bayes' rule whenever possible. Thus it takes into account the report chosen, the allocation recommended by the mechanism and the mechanism itself. This is the role of information embedded in the design of the mechanism. However, the information that the mechanism designer can reveal or conceal is not at his/her entire disposal. Later I show that when the condition for first best mechanisms is not satisfied, any mechanism satisfying non-binding arbitration IC and IR must leak some information. That is, regardless of the objective function of the designer, any implementable mechanism is bound to leak some information in cases where conflict (outside options) occurs with strictly positive probability on the type space.

Given the role of information, a deviation to misreport a type may not only be motivated by what the mechanism prescribes for some other type. It could also be motivated to obtain a more informative lottery of allocations prescribed for the other type. This may allow a player to adjust to choose his/her outside option in favorable situations and keep the allocation in other situation. This "strategic" acquisition of information is not present in binding arbitration.

With binding arbitration, whenever an allocation is recommended, the players have to adhere to it. This happens regardless of truth in reporting. Hence, a player that misreports his/her type does not have the mentioned access to strategic information. Misreporting could imply a change in the beliefs of a player, but false information cannot be used to choose which allocations to accept and which to reject. That is why the incentive compatibility constraints are described as they are below.

A direct mechanism (π, g) satisfies **binding arbitration** IC, if for each $i = 1, 2$ and each $\theta_i^* \in \Theta_i$

$$\mathbb{E}_{\theta_j^*}\left[\pi(\theta_i^*, \theta_j^*)u_i(\theta_i^*, \theta_j^*) + \int_{[0,1]} x g_i(x|(\theta_i^*, \theta_j^*))dx\right] \geq$$
$$\mathbb{E}_{\theta_j^*}\left[\pi(\theta_i, \theta_j^*)u_i(\theta_i^*, \theta_j^*) + \int_{[0,1]} x g_i(x|(\theta_i, \theta_j^*))dx\right]$$

for all $\theta_i \in \Theta_i$

A player cannot potentially "benefit" from the information he/she could obtain from reporting a type different from his/her own, as opposed to the case of non-binding arbitration. The flow of information is now shut down by the enforcement power of the mechanism designer.

Non-binding arbitration IR has a role for information. However, the role only exists for those allocations that are prescribed for truthful reports.

A direct mechanism (π, g) satisfies **non-binding arbitration** IR if for each $i = 1, 2$, for every $\theta_i^* \in \Theta_i$, for all $x \in [0, 1]$

$$x \geq \mathbb{E}_z[u_i(\theta_i^*, z)|x, \theta_i^*]$$

Any mechanism induces a posterior distribution over outside options through the allocations that it recommends. With non-binding arbitration, the designer needs to prescribe allocations that are accepted given the updated beliefs of the players, when they report truthfully. This is captured in non-binding IR constraints. Notice that if the outside option of a player depends only on his/her own type, the updated beliefs about the type of the other player would be irrelevant. With interdependent outside options, any information "leaked" by the mechanism becomes relevant. However, when binding arbitration is conducted, players cannot make use of information leaked from the mechanism, in spite of its potential value. Thus the interaction of interdependent outside options with non-binding arbitration makes relevant the information problem embedded in the design of the mechanism. Individual rationality constraints for binding arbitration do not need to take into account the potential information that the players obtain from the mechanism.

A direct mechanism (π, g) satisfies **binding arbitration** IR if for each $i = 1, 2$, for every $\theta_i^* \in \Theta_i$,

$$\mathbb{E}_{\theta_j^*}\left[\pi(\theta_i^*,\theta_j^*)u_i(\theta_i^*,\theta_j^*)+\int_{[0,1]} xg_i(x|(\theta_i^*,\theta_j^*))dx\right] \geq \mathbb{E}_{\theta_j^*}[u_i(\theta_i^*,\theta_j^*)]$$

As information acquisition does not play role in binding arbitration, the IR constraints can be taken at the interim level. The designer's enforcement power allows him/her to provide to the players just the interim expected outside option regardless of their updated beliefs. Therefore, the mechanism designer needs only to ensure that each player decides to participate in the mechanism before any recommendation from it has been made.

To simplify notation and for explanation purposes, I introduce the following notation. Given a direct mechanism (π, g), for each $i \in \{1,2\}$, let $x_i : \Theta_i \times \Theta_j \to [0,1]$ be such that:

$$x_i(\theta_i,\theta_j) = \begin{cases} \dfrac{1}{1-\pi(\theta_i,\theta_j)} \int_{[0,1]} xg_i(x|(\theta_i,\theta_j))dx, & \text{if } \pi(\theta_i,\theta_j) \neq 1 \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

I call $x_i(\theta_i,\theta_j)$ the expected allocation for player i according to (π, g), if the reports were (θ_i,θ_j).

Using this, the utility for player i, of type θ_i, from participating in the mechanism when reporting truthfully and accepting all recommended allocations, is given by $\mathbb{E}_{\theta_j}\left[\pi(\theta_i,\theta_j)u_i(\theta_i,\theta_j) + (1-\pi(\theta_i,\theta_j))\, x_i(\theta_i,\theta_j)\right]$.

The following two lemmas illustrate a useful relation between binding arbitration and non-binding arbitration IC and IR.

Lemma 1. *If a direct mechanism satisfies non-binding arbitration IR, then it satisfies binding arbitration IR.*

Lemma 2. *If a direct mechanism satisfies non-binding arbitration IC, then it satisfies binding arbitration IC.*

The proofs of Lemmas 1 and 2 are straightforward and relegated to the Appendix for brevity. Together Lemmas 1 and 2 say that if a mechanism (π, g) satisfies non-binding arbitration IC and IR then it also satisfies binding arbitration IC and IR. This is important for the objective of the designer.

I am interested in ex ante efficiency. Thus the objective of the designer is to maximize the expected ex ante sum of utilities of the players from participating in the mechanism. Using the notation introduced before, the objective of the designer is,

$$\max_{(\pi,g)} \mathbb{E}_{(\theta_1,\theta_2)} \left[\pi(\theta_1,\theta_2)\left(u_1(\theta_1,\theta_2)+u_2(\theta_1,\theta_2)\right)+\left(1-\pi(\theta_1,\theta_2)\right)\left(x_1(\theta_1,\theta_2)+x_2(\theta_1,\theta_2)\right)\right] \tag{2.2}$$

I call **non-binding arbitration problem** the optimization problem of (2.2) subject to non-binding arbitration IC and IR. I call **binding arbitration problem** the optimization problem of (2.2) subject to binding arbitration IC and IR.

Given Lemmas 1 and 2, it is clear that a designer conducting non-binding arbitration cannot attain a strictly higher value for (2.2) than that obtained when conducting binding arbitration.

Notice that given this objective function and the constraints for the binding arbitration problem, it is without loss of generality for the binding arbitration problem to focus on mechanisms such as (π, x), where $x = (x_1, x_2)$ with each x_i being the expected allocation for player i, as defined in (2.1). I use this for exposition purpose whenever it reduces notation.

2.3 The First Best

I introduce the first best mechanisms for my first result. First best mechanisms assign zero probability to the outside option outcome for any possible report, and the allocation prescribed splits the pie completely, that is $x_1 + x_2 = 1$ for all recommendations. Since outside options are inefficient, assigning zero probability to them for all reports shuts down the inefficiency channel from outside options, and the value of the objective function as in (2.2) is maximized by always allocating the pie completely. The highest possible value of (2.2) is one (regardless of the constraints). Notice that first best mechanisms are, then, ex post efficient. This seems to be very restrictive a priori, but they are nevertheless possible.

Definition 1. *A direct mechanism (π, g) is called **first best** if $\pi(\theta_1, \theta_2) = 0$ for any $(\theta_1, \theta_2) \in \Theta$. In addition, for any (x_1, x_2) such that $g((x_1, x_2)|(\theta_1, \theta_2)) > 0$, for some (θ_1, θ_2), it satisfies $x_1 + x_2 = 1$.*

My first result states that it is possible to implement first best mechanisms, that is, ex-post efficiency. I establish a necessary and sufficient condition for this to hold, which I call the efficiency condition. The environment satisfies the **efficiency condition** if

$$\mathbb{E}_{\theta_2}[u_1(\overline{\theta}, \theta_2)] + \mathbb{E}_{\theta_1}[u_2(\theta_1, \overline{\theta})] \leq 1$$

The efficiency condition entails two things. On the one hand, it requires the expected outside option of the highest types not to be particularly high. That is, the highest types do not obtain too much of the pie. On the other hand, since the outside options are decreasing in the other player's type, lower types do not get extremely small slices of the pie.

Theorem 1. *There is a first best mechanism (π, g) that satisfies non-binding arbitration IC and IR if and only if*

$$\mathbb{E}_{\theta_2}[u_1(\overline{\theta}, \theta_2)] + \mathbb{E}_{\theta_1}[u_2(\theta_1, \overline{\theta})] \leq 1$$

Proof. Suppose $\mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2) + \mathbb{E}_{\theta_1} u_2(\theta_1, \overline{\theta}) \leq 1$. Fix a mechanism (π, g), for which $x_i(\theta_i, \theta_j)$ is the expected allocation that player i would be recommended, according to (π, g), if the reports were (θ_i, θ_j). Let $x_1(\theta_1^*, \theta_2^*) = \mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2)$ for all $(\theta_1^*, \theta_2^*) \in \Theta$ and $x_2(\theta_1^*, \theta_2^*) = 1 - x_1(\theta_1^*, \theta_2^*)$. Let $\pi(\theta_1^*, \theta_2^*) = 0$ for all $(\theta_1^*, \theta_2^*) \in \Theta$. Now, verify that this proposed first best mechanism satisfies non-binding arbitration IC and IR. To see that non-binding arbitration IR holds, note that since the allocation is constant for all reports and π is equal to zero for all reports, we have for all $i, j = 1, 2$, $i \neq j$, $\Theta_j^*(\theta_i^*, \theta_j^*) = [\underline{\theta}, \overline{\theta}]$ for all $(\theta_i^*, \theta_j^*) \in \Theta$. Since, for all $(\theta_1^*, \theta_2^*) \in \Theta$, $x_1(\theta_1^*, \theta_2^*) = \mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2)$ and $x_2(\theta_1^*, \theta_2^*) = 1 - x_1(\theta_1^*, \theta_2^*) \geq \mathbb{E}_{\theta_1} u_2(\theta_1, \overline{\theta})$, using the assumption of outside options increasing in player's own type. Thus we have guaranteed non-binding arbitration IR.

To see that non-binding arbitration IC constraints also hold, note that the allocation recommended is constant for all reports and always better than the highest expected (unconditional) outside option for each player. Also, π is equal to zero everywhere on Θ. Then, for player 1, the expected payoffs for type θ_1^* of sending any report $\theta_1^{**} \in [\underline{\theta}, \overline{\theta}]$ are $\mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2)$. For player 2, the expected payoffs for type θ_2^* of sending any report $\theta_2^{**} \in [\underline{\theta}, \overline{\theta}]$ are $\mathbb{E}_{\theta_2} \left(1 - u_1(\overline{\theta}, \theta_2)\right)$. So non-binding arbitration IC holds.

To see the necessity of the efficiency condition, suppose we have a first best mechanism (π, x), so π is equal to zero everywhere in Θ, and suppose it satisfies non-binding arbitration IC and IR. Using non-binding arbitration IR, and the fact that the mechanism is first best, then for each $(\theta_i^*, \theta_j^*) \in \Theta$ and each i, $x_i(\theta_i^*, \theta_j^*) \geq \mathbb{E}_{\theta_j}[u_i(\theta_i^*, \theta_j) \mid \theta_j \in \Theta_j^*(\theta_i^*, \theta_j^*)]$. Taking expectation in both sides with respect to all θ_j^*, we obtain $x_i^e(\theta_i^*) := \mathbb{E}_{\theta_j} x_i(\theta_i^*, \theta_j) \geq \mathbb{E}_{\theta_j} u_i(\theta_i^*, \theta_j)$. Using this notation, since non-binding arbitration IR holds and the mechanism is first best, for each i, the expected payoffs for type θ_i^* from reporting truthfully are $x_i^e(\theta_i^*)$, and because non-binding arbitration IC holds, for any θ_i^{**},

$$x_i^e(\theta_i^*) \geq \mathbb{E}_{\theta_j}\left[\pi_i\left(\theta_i^*, \theta_i^{**}, \theta_j\right) x_i(\theta_i^{**}, \theta_j) + \left(1 - \pi_i\left(\theta_i^*, \theta_i^{**}, \theta_j\right)\right) u_i(\theta_i^*, \theta_j)\right].$$

Denote $\tilde{\Theta}_j(\theta_i^*, \theta_i^{**}) := \{\tilde{\theta}_j | x_i(\theta_i^{**}, \tilde{\theta}_j) < \mathbb{E}_{\theta_j}[u_i(\theta_i^*, \theta_j) | \theta_j \in \Theta_j^*(\theta_i^{**}, \tilde{\theta}_j)]\}$ and $\tilde{\Theta}_j^c(\theta_i^{**}, \theta_i^*)$ to be its complement relative to $[\underline{\theta}, \overline{\theta}]$, then

$$\begin{aligned} x_i^e(\theta_i^*) \geq & \int_{\tilde{\Theta}_j^c(\theta_i^{**}, \theta_i^*)} x_i(\theta_i^{**}, \theta_j) dF(\theta_j) + \int_{\tilde{\Theta}_j(\theta_i^{**}, \theta_i^*)} u_i(\theta_i^*, \theta_j) dF(\theta_j) \geq \\ & \int_{\tilde{\Theta}_j^c(\theta_i^*, \theta_i^{**})} x_i(\theta_i^{**}, \theta_j) dF(\theta_j) + \int_{\tilde{\Theta}_j(\theta_i^*, \theta_i^{**})} x_i(\theta_i^{**}, \theta_j) dF(\theta_j) \end{aligned}$$

. To see the latter inequality, note that for all $\tilde{\theta}_j \in \tilde{\Theta}_j(\theta_i^*, \theta_i^{**})$, $\mathbb{E}_{\theta_j}[u_i(\theta_i^*, \theta_j) | \theta_j \in \Theta_j^*(\theta_i^{**}, \tilde{\theta}_j)] > x_i(\theta_i^{**}, \tilde{\theta}_j)$. Integrating both sides with respect to all $\tilde{\theta}_j \in \tilde{\Theta}_j(\theta_i^*, \theta_i^{**})$, gives

$$\int_{\tilde{\Theta}_j(\theta_i^*, \theta_i^{**})} u_i(\theta_i^*, \theta_j) dF(\theta_j) \geq \int_{\tilde{\Theta}_j(\theta_i^*, \theta_i^{**})} x_i(\theta_i^{**}, \theta_j) dF(\theta_j)$$

and hence, the latter inequality. Similarly, for type θ_i^{**}, we have $x_i^e(\theta_i^{**}) \geq x_i^e(\theta_i^*)$. Hence, x_i^e is constant for all i and we denote its value simply by x_i^e. Since $x_2(\theta_1, \theta_2) = 1 - x_1(\theta_1, \theta_2)$, taking expectations with respect to θ_1 and θ_2, we have $x_1^e + x_2^e = 1$. Taking expectation from non-binding arbitration IR with respect to all θ_j, we have $x_i^e \geq \mathbb{E}_{\theta_j} u_i(\theta_i, \theta_j)$ and since outside options are increasing on own type, $x_1^e \geq \mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2)$ and $1 - x_1^e \geq \mathbb{E}_{\theta_1} u_2(\theta_1, \overline{\theta})$. Hence, $\mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2) + \mathbb{E}_{\theta_1} u_2(\theta_1, \overline{\theta}) \leq 1$ □

The condition for Theorem 1 suggests a very interesting insight within the context of the literature on conflict. When the efficiency condition holds, the expected outside option of the highest types is not very large, but because of our assumption on the outside options (decreasing in the other player's type), the lower types do not obtain too little of the pie from their outside options. The condition can be interpreted as balanced forces at conflict. That is, types are not "too far apart" in expectations in terms of outside options. Since the efficiency condition is necessary and sufficient, it suggests a clear connection between conflict and uncertainty. When lower types are "too far apart" from higher types in terms of payoffs of outside options, there is a big difference in facing lower rather than higher types. By Theorem 1 conflict cannot be avoided in this case.

To illustrate the efficiency condition more specifically, I show two examples. In both cases, for simplicity of exposition I assume both players have the same utility of outside options. The first example considers uniform distribution and outside options as contest success functions (Skarpedas (1996)). That is, for all $i = \{1, 2\}$, $u_i(\theta_i, \theta_j) = \dfrac{\theta_i}{\theta_i + \theta_j}(1 - \ell)$, where ℓ is the part of the pie that is lost due to inefficiency of the outside options. The efficiency condition then becomes,

$$(1-\ell)\frac{2\overline{\theta}}{\overline{\theta}-\underline{\theta}}\left(log\left(\frac{2\overline{\theta}}{\overline{\theta}+\underline{\theta}}\right)\right) \leq 1$$

Notice that if ℓ is sufficiently close to one, the efficiency condition would be satisfied. From the perspective of the designer, high inefficiency from outside options helps him/her achieve efficiency. This is intuitively sound in the case of war: when armed conflict is highly destructive, both players prefer to settle according to a "peaceful" alternative.

The efficiency condition is a condition, jointly, on utilities (from outside options) and distributions of types. The second example I show seeks to illustrates this. Suppose the utility of outside options is such that for each $i = \{1, 2\}$, $u_i(\theta_i, \theta_j) = a\theta_i - b\theta_j$, with $a, b > 0$. I refer to this case as separable utility. The efficiency condition becomes

$$\frac{(a\overline{\theta} - 1/2)}{b} \leq \mathbb{E}(\theta)$$

In this case, the efficiency condition requires the first moment of the distribution of types not to be too low. In other words, when players go to the outside option outcome, they expect to face a relatively high type. This means that the expected payoffs obtained from outside options are not high. Hence, outside options are not attractive for players, a priori, and therefore, they prefer to settle to avoid conflict.

The following Corollary to Theorem 1 allows me to discuss the role of information in first best mechanisms.

Corollary. *Given a first best mechanism (π, g), for each $i = 1, 2$, $x_i(\theta_i, \theta_j)$ can be taken to be constant.*

In a first best mechanism, the player's expected allocations can be taken to be constant. This implies that the designer can avoid leaking any information whatsoever through the mechanism. This observation comes directly from the efficiency condition. When it is met, the designer (for instance) can give $\mathbb{E}_{\theta_2} u_1(\overline{\theta}_1, \theta_2)$ to player 1 and $\mathbb{E}_{\theta_1} u_2(\theta_1, \overline{\theta}_2)$ for any report. Since the allocation is not conveying any information, the non-binding individual rationality constraints are taken just at the interim level. The allocation satisfies non-binding arbitration IR for the highest type. Therefore it satisfies it also for all the lower types. This is because outside options are increasing in the type of the player. Lastly, for non-binding arbitration IC, notice that any reported type gives the same allocation, which is bigger than the interim expected outside options for all types. Therefore, accepting the recommended allocation obtained from misreporting is optimal. Thus the utility of reporting truthfully is the same as that of misreporting.

As noted above, the maximum possible value for the objective function in (2.2) is the size of the pie (assumed to be one). This is an upper bound regardless of the constraints. Given Theorem 1, this value is obtained in the non-binding arbitration problem with first best mechanisms, when the efficiency condition holds. By Lemmas 1 and 2, the first best mechanisms of non-binding arbitration can be implemented in binding arbitration. Hence, both binding and non-binding arbitration problems attain the same value for the objective function. I state this more formally in the following Remark.

Remark 1: Suppose $\mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2) + \mathbb{E}_{\theta_1} u_2(\theta_1, \overline{\theta}) \leq 1$. Fix a mechanism (π^*, g^*) that is a solution to the non-binding binding arbitration problem. Then (π^*, g^*) is also a solution to the binding arbitration problem.

This result follows from Theorem 1 and Lemmas 1 and 2. Whenever $\mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2) + \mathbb{E}_{\theta_1} u_2(\theta_1, \overline{\theta}) \leq 1$, non-binding and binding arbitration problems can assign zero probability to the outside options. Thus when the efficiency condition holds, the non-binding arbitration problem does not lead to a worsening of efficiency as compared with the binding arbitration problem. The efficiency condition allows the mechanism designer to implement first best mechanisms.

Furthermore, if there is a first best mechanism that satisfies binding arbitration IC and IR, the efficiency condition must hold. This follow as a direct observation from the proof of Theorem 1. I formally state it as follows.

Remark 2: There is a first best mechanism (π, g) that satisfies binding arbitration

IC and IR if and only if $\mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2) + \mathbb{E}_{\theta_1} u_2(\theta_1, \overline{\theta}) \leq 1$.

The efficiency condition is necessary and sufficient for implementing binding and non-binding arbitration first best mechanisms. When the efficiency condition holds, the first best mechanism can be taken to conceal all information, therefore bypassing the information leakage problem. This translates into the same IR constraints for both problems at the interim level. On incentive compatibility constraints, this means that the allocation from misreporting is the same as the one obtained with truthful reports accepted at the interim level. Since there is no information leakage, the allocation is still accepted once a player misreports his/her type. Hence, the IC constraints are also regarded as the same in both binding and non-binding arbitration.

To see the necessity of the efficiency condition, note that if a first best mechanism is implementable with binding arbitration, the expected allocation for each player is constant as a function of his type by binding arbitration IC. By Lemma 2, a first best mechanism that is implementable with non-binding arbitration also has to satisfy this. Then, by binding arbitration IR and Lemma 1, first best mechanisms for binding and non-binding arbitration need to satisfy the condition that the expected allocation for each player is greater than the expected outside option for the highest type. Since allocations come from splitting a pie of size 1, the expected value of the allocations cannot add up to a value more than 1. This means that the expected outside options of the highest type for each player cannot add up to a value greater than one.

The question remains as to the solution to both the binding arbitration and non-binding arbitration problems without the efficiency condition. The efficiency condition can be "generalized" in the sense that the designer can try to find the largest region for zero probability of conflict. The following proposition shows that it is always possible to have a "peaceful" region of the type space, that is, a region where whenever two types of this region meet, they never resort to their outside option.

Theorem 2. *Suppose each u_i is continuous. There is always a profile (θ_1^*, θ_2^*) with $\underline{\theta} < \theta_1^*$ and $\underline{\theta} < \theta_2^*$ so that, for some mechanism (π, g), the following holds:*

1. $\pi(\theta_1, \theta_2) = 0$ *for all* $(\theta_1, \theta_2) \in [\underline{\theta}, \theta_1^*] \times [\underline{\theta}, \theta_2^*)$
2. (π, g) *satisfies non-binding arbitration IR and IC.*

Proof. Define $N : \Theta \to \mathbb{R}$ as follows,

$N(\theta_1, \theta_2) = \frac{1}{F_2(\theta_2)} \int_{[\underline{\theta}, \theta_2]} u_1(\theta_1, z) dF_2(z) + \frac{1}{F_1(\theta_1)} \int_{[\underline{\theta}, \theta_1]} u_2(z, \theta_2) dF_1(z).$

Note that if $\mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2) + \mathbb{E}_{\theta_1} u_2(\theta_1, \overline{\theta}) \leq 1$, we are in the context of Theorem 1 and, hence, it can be taken $\theta_1^* = \theta_2^* = \overline{\theta}$. If $\mathbb{E}_{\theta_2} u_1(\overline{\theta}, \theta_2) + \mathbb{E}_{\theta_1} u_2(\theta_1, \overline{\theta}) > 1$, then $N(\overline{\theta}, \overline{\theta}) > 1$. By the inefficiency assumption, $u_1(\underline{\theta}, \underline{\theta}) + u_2(\underline{\theta}, \underline{\theta}) < 1$. Let $Z = \{(\theta_1, \theta_2) \in \Theta : N(\theta_1, \theta_2) = 1\}$. Then, by continuity of u_i and f_i, Z is non-empty. Pick an element $(\theta_1^*, \theta_2^*) \in Z$, then $\theta_1^* > \underline{\theta}$ and $\theta_2^* > \underline{\theta}$.

Let $[(\theta_1^*, \theta_2^*)] = \{(\theta_1, \theta_2) \in \Theta \mid \theta_1 \leq \theta_1^* \text{ and } \theta_2 \leq \theta_2^*\}$. Let $\pi(\theta_1, \theta_2) = 0$ for all $(\theta_1, \theta_2) \in [(\theta_1^*, \theta_2^*)]$. For all $i = 1, 2$, $i \neq j$, let $x_i(\theta_1, \theta_2) = \mathbb{E}_{\theta_j'} u_i(\theta_i^*, \theta_j')$ for all $(\theta_1, \theta_2) \in [(\theta_1^*, \theta_2^*)]$ and $\pi(\theta_1, \theta_2) = 1$ for all $(\theta_1, \theta_2) \notin [(\theta_1^*, \theta_2^*)]$. It is clear that whenever the recommended allocation is reached, all players will accept it. So non-binding arbitration IR holds. To see that non-binding arbitration IC holds, note that if $\theta_i > \theta_i^*$ misreports to $\theta_i' > \theta_i^*$ his expected payoff remains the same and equal to $\mathbb{E}_{\theta_j} u_i(\theta_i, \theta_j)$. If he misreports to $\theta_i'' \leq \theta_i^*$ his expected payoff is $F_j(\theta_j^*)\mathbb{E}_{\theta_j}\left[u_i(\theta_i^*, \theta_j) | \theta_j \leq \theta_j^*\right] + (1 - F_j(\theta_j^*))\mathbb{E}_{\theta_j}\left[u_i(\theta_i, \theta_j) | \theta_j \geq \theta_j^*\right]$. Since outside options are strictly increasing in own type and $\theta_i > \theta_i^*$, misreporting to θ_i'' has strictly less expected payoffs.

If $\theta_i \leq \theta_i^*$ misreports to $\theta_i' \leq \theta_i^*$ his expected payoff remains the same and equal to

$F_j(\theta_j^*)\mathbb{E}_{\theta_j}\left[u_i(\theta_i^*, \theta_j)|\theta_j \le \theta_j^*\right] + (1 - F_j(\theta_j^*))\mathbb{E}_{\theta_j}\left[u_i(\theta_i, \theta_j)|\theta_j \ge \theta_j^*\right] \mathbb{E}_{\theta_j} u_i(\theta_i, \theta_j)$. If he misreports to $\theta_i'' > \theta_i^*$ his expected payoff is $\mathbb{E}_{\theta_j} u_i(\theta_i, \theta_j)$. Since outside options are strictly increasing in own type and $\theta_i < \theta_i^*$, misreporting to θ_i'' has strictly less expected payoffs Thus, non-binding arbitration IC holds. □

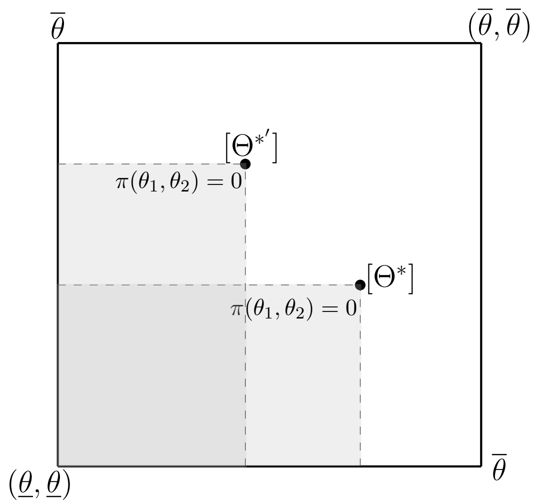

Figure 2.1: Graphical depiction of Theorem 2.

Figure 2.1 illustrates Theorem 2. The designer can choose among many rectangles of the type space. I have shown two: Θ^* and $\Theta^{*'}$. For each of these sets, $\pi(\theta_1, \theta_2) = 0$ for any report of types within that set. Also, as in Corollary 1, the designer is able to withhold information within a given set by recommending a fixed expected allocation for any report within such a set. The intuition here is important. The mechanisms in Theorem 2 are not necessarily the solutions to the arbitration problem. However, they are desirable from the perspective of a designer with "lexicographic preferences for peace." In the context of international conflict, this seems reasonable when war is likely to be highly destructive. The importance of the mechanisms is that they provide intuition and some features of the solution to the binding arbitration problem

(Theorem 3), and they help explain how that solution can be implemented with non-binding arbitration (Theorem 4).

Observe that the problem without the efficiency condition can be interpreted as a problem of resource constraints. In particular, the condition on Theorem 1 not being satisfied can be interpreted as a budget constraint breaking efficiency. Without the efficiency condition, it would not be a problem if the designer could use resources from elsewhere. Note that absent any constraints in possible beliefs, the designer would like to convince the high types that they are facing high types "to the extent possible", in order to relax the budget constraint of splitting 1. On the other hand, lower types do not have the problem of constrained resources when facing other low types. Another way of saying this is to observe that the efficiency condition holds at the bottom of the type space. Remember that the intuition of the efficiency condition is that types are similar or balanced, and thus, the payoffs of the conflict outcome do not vary much. Note that this is the case if we group low types. Within this group, types are balanced, and hence, the efficiency condition holds. This is indeed the logic for the solution of the binding arbitration problem in Theorem 3. In Theorem 4, I show that a solution of these characteristics is implementable with non-binding arbitration and is actually optimal when the efficiency condition is not satisfied. The designer conceals some information at the lower region of the type space. Within this group, players do not know exactly who they are facing, but at the same time they know they are in the low group. The higher region of the type space is where budget constraints plays a major role. In these regions high types may face lower types, and the difference between the payoffs of both is big. Some information must be revealed when seeking efficiency. It is possible to pool higher and lower types in the way described above, which illustrates the intuition behind Theorem 3.

2.4 Beyond The First Best

For this section I depart from the efficiency condition, and hence from first best mechanisms, to continue to study efficiency. I am interested in the impact on efficiency of the enforcement power of the mechanism designer. The answer to the question of when the efficiency condition holds was given in the previous section. Both binding and non-binding arbitration problems attain the first best. The question remains how to proceed when the first best cannot be implementable. To answer this, I study the binding arbitration problem when the efficiency condition does not hold. I show the solution to the binding arbitration problem in Theorem 3, under a condition in the primitives that allows π to be increasing. In Theorem 4, I show a sufficient condition for binding and non-binding arbitration to be "equivalent" in terms of efficiency when the first best is not implementable.

To motivate the discussion beyond the efficiency condition, I introduce the following definition.

Definition 2. *We say a mechanism (π, g) fully conceals information if for all i and each θ_i,*
$\mathbb{E}_{\theta_j}[u_i(\theta_i, \theta_j)|x_i, \theta_i] = \mathbb{E}_{\theta_j}[u_i(\theta_i, \theta_j)]$, *for all x_i such that $g((x_i, x_j)|(\theta_i, \theta_j)) > 0$ for some θ_j and some x_j.*

The following result allows me to discuss the efficiency condition in the language of information design.

Proposition 1. *Suppose the environment does not satisfy the* ***efficiency condition****. Let (π, g) satisfy non-binding arbitration IC and IR. Then (π, g) does not fully conceal information.*

Proof. Suppose by contradiction that (π, g) fully conceals information and that the model does not satisfy the efficiency condition. By fully concealing for $(\overline{\theta}, \overline{\theta})$, for each (x_1, x_2) such that $g((x_1, x_2)|(\overline{\theta}, \overline{\theta})) > 0$, we have $x_1 \geq \mathbb{E}_{\theta_2}[u_1(\overline{\theta}, \theta_2)]$ and $x_2 \geq \mathbb{E}_{\theta_1}[u_2(\theta_1, \overline{\theta})]$, by non-binding arbitration IR. Thus, $1 \geq x_1 + x_2 \geq \mathbb{E}_{\theta_2}[u_1(\overline{\theta}, \theta_2)] + \mathbb{E}_{\theta_1}[u_2(\theta_1, \overline{\theta})]$. But then, this implies the efficiency condition, leading to a contradiction.

□

Proposition 1 states that whenever there is some probability of conflict, the designer must leak some information. When higher types have higher outside options (i.e., we do not have the efficiency condition), the designer needs to satisfy very demanding IR constraints. This cannot be done by concealing information when two high types face each other (they both have high outside options). The designer, then, may prefer to reveal some information to higher type players to alleviate the stringent IR. For instance, revealing to a high type that he is facing a high type makes both players have a less demanding IR, since both know they would not get much from the outside option. However, information sharing needs to be done in a way that is incentive compatible. Balancing these two forces, the designer is bound to reveal some information.

Proposition 1 shows a strong connection between conflict and information leakage. Without the efficiency condition, the designer cannot induce any random beliefs he/she chooses. The designer is restricted as to what information he/she can convey. This represents a difference with respect to the traditional literature on information design. As observed, fully concealing information is possible only when the model satisfies the efficiency condition. This suggests that the designer who seeks efficiency "weakly prefers" concealing information.

I introduce the following definition to explain the next results of this section.

Definition 3. *Let* $\ell : [\underline{\theta}, \overline{\theta}]^2 \to [0,1]$*, be defined as follows. For each* (θ_1, θ_2)

$$\ell(\theta_1, \theta_2) = 1 - u_1(\theta_1, \theta_2) + u_2(\theta_1, \theta_2)$$

I call ℓ *the efficiency loss function.*

For each pair of types (θ_1, θ_2), $\ell(\theta_1, \theta_2)$ indicates the part of the pie that is lost when type θ_1 and type θ_2 go to their outside options. For this reason, the situation is called efficiency loss. As discussed in the description of the model, efficiency loss is assumed to be always strictly positive. That is, for each pair (θ_1, θ_2), $\ell(\theta_1, \theta_2) > 0$. In the example of international conflict, this indicates how destructive war can be. In cartel formation, it indicates the decrease in profits from monopoly to the sum of the profits of competing firms.

Additionally, I define the following function, let $J : [\underline{\theta}, \overline{\theta}] \times [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$ be such that,

$$J(\theta_i, \theta_j) = \frac{F_i(\theta_i)}{f_i(\theta_i)} \left(\frac{\partial u_i(\theta_i, \theta_j)/\partial \theta_i}{\ell(\theta_i, \theta_j)} \right) + \frac{F_j(\theta_j)}{f_j(\theta_j)} \left(\frac{\partial u_j(\theta_i, \theta_j)/\partial \theta_j}{\ell(\theta_i, \theta_j)} \right)$$

.

The function J is the analogue of the virtual value function in Myerson (1981). Given the binding arbitration problem, when J is increasing, the solution to the binding arbitration problem involves π increasing. The following result illustrates the solution to the binding arbitration problem using this condition on J. In what follows, for simplicity, I focus on symmetric environments. By a symmetric environment, I mean, $\Theta_i = \Theta_j = [\underline{\theta}, \overline{\theta}]$, $u_i(\theta_i, \theta_j) = u_j(\theta_j, \theta_i)$, for each $(\theta_i, \theta_j) \in [\underline{\theta}, \overline{\theta}]$ and $f_i = f_j$, where f_i is the density of types for player i. I assume each u_i is C^1.

Theorem 3. *Fix a symmetric environment. Suppose that for each i, u_i is C^1 and that J is increasing. Then, any (π, g) solution to the binding arbitration problem partitions the type space into two regions: Θ^* and Θ^{*c}, such that $\pi = 1$ almost everywhere in Θ^{*c} and $\pi = 0$ almost everywhere in Θ^*.*

The proof of Theorem 3 can be found in the Appendix. I provide an outline and explanation of the proof in this section. I first introduce some notation for ease of exposition. For the binding arbitration problem, given a mechanism (π, g), I use (π, x), where x is the expected allocation as defined earlier in equation (2.1). For a given mechanism (π, x), I define for each i the interim utility of participating in the mechanism, by $V_i : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$ such that

$$V_i(\theta_i) = \mathbb{E}_{\theta_j}\left[\pi(\theta_i, \theta_j)u_i(\theta_i, \theta_j) + (1 - \pi(\theta_i, \theta_j))x_i(\theta_i, \theta_j)\right].$$

I write V to mean $V = (V_1, V_2)$

The following two lemmas are used to prove Theorem 3.

Lemma 3. *If a mechanism (π, x) satisfies binding arbitration IC and binding arbitration IR for the highest type, then it satisfies binding arbitration IC and IR.*

The proof of Lemma 3 can be found in the Appendix. Lemma 3 allows me to reduce the binding arbitration problem's IR constraints. In particular, it states that is sufficient to satisfy binding arbitration IR for the highest type in the binding arbitration problem. The following lemma states that the traditional envelope condition is necessary for binding arbitration IC. Notice, however, that in general it is not possible to obtain an IC characterization. In spite of that, Lemma 4 is still useful for solving the binding arbitration problem.

Lemma 4. *Suppose, for each i, u_i is a C^1 function. If a mechanism (π, x) satisfies binding arbitration IC, then for all $i = 1, 2$, and all θ_i,*

$$V_i(\theta_i) = V_i(\underline{\theta}) + \mathbb{E}_{\theta_j}\left[\int_{\underline{\theta}}^{\theta_i} \pi(z, \theta_j)\frac{\partial u_i(z, \theta_j)}{\partial z}dz\right]$$

Proof. By assuming for all $i = 1, 2$, u_i being C^1, all the assumptions from Theorem 2 in Milgrom and Seagal (2002) follow.

Denote $V_i(\theta|\theta_i) := \mathbb{E}_{\theta_j}[\pi(\theta, \theta_j)u_i(\theta_i, \theta_j) + (1 - \pi(\theta, \theta_j))x_i(\theta, \theta_j)]$. Since, u_i is C^1, for any given θ, $V_i(\theta|\cdot)$ is absolutely continuous. By the same assumption, and since π is a probability and the set of types is contained in a compact space, there is an integrable function $b : [\underline{\theta}, \overline{\theta}] \to \Re$ such that $\left|\frac{\partial V_i(\theta, \theta_i)}{\partial \theta_i}\right| \leq b(\theta_i)$ and $V_i(\theta|\cdot)$ is differentiable for all θ. Finally by binding arbitration IC, $V_i(\theta, \theta_i)$ is maximized at $\theta = \theta_i$. Therefore, the envelope condition follows from Milgrom and Segal (2002), Theorem 2. □

Lemma 4 gives a necessary condition for incentive compatibility when transforming the binding arbitration problem from choosing (π, g) to choosing (π, V). The role of the assumption on J is of relevance in this case. If J is increasing, the solution (π, V) to a relaxed binding arbitration problem has the property that π is increasing. The envelope condition from Lemma 4 together with the fact that π increasing gives binding arbitration IC. Thus we can solve a relaxed problem and obtain the solution to the binding arbitration problem.

I introduce the following relaxed problem to best explain the proof of Theorem 3.

$$\max_{(\pi,g)} \mathbb{E}_{(\theta_1,\theta_2)} \left[\pi(\theta_1,\theta_2)\left(u_1(\theta_1,\theta_2)+u_2(\theta_1,\theta_2)\right)+(1-\pi(\theta_1,\theta_2))\left(x_1(\theta_1,\theta_2)+x_2(\theta_1,\theta_2)\right)\right]$$

s.t.

$$\text{For all } i \text{ and all } \theta_i,\ V_i(\theta_i) = V_i(\underline{\theta}) + \mathbb{E}_{\theta_j}\left[\int_{\underline{\theta}}^{\theta_i} \pi(z,\theta_j)\frac{\partial u_i(z,\theta_j)}{\partial z}dz\right]$$

$$\text{For all } i,\ V_i(\overline{\theta}) \geq \mathbb{E}_{\theta_j} u_i(\overline{\theta},\theta_j)$$

$$\mathbb{E}_{\theta_i,\theta_j}\left[V_i(\theta_i)+V_j(\theta_j)\right] \leq \mathbb{E}_{\theta_i,\theta_j}\left[\pi(\theta_i,\theta_j)\left(u_i(\theta_i,\theta_j)+u_j(\theta_i,\theta_j)\right)+(1-\pi(\theta_i,\theta_j))\right]$$

The first constraint is called the envelope condition. The second is highest type binding arbitration IR. The third is called expected feasibility. The first constraint is necessary for binding arbitration IC, by Lemma 4. The second constraint suffices for binding arbitration IR, by Lemma 3. Notice that this problem is a relaxation of the binding arbitration problem, because the binding arbitration IC constraints may not hold. That is why I call it a Binding Arbitration Relaxed Problem (BARP). Once the binding arbitration problem is transformed from choosing (π, g) to choosing (π, V), there is one question remaining about whether or not the chosen (π, V) is actually feasible, in the sense that can be obtained from some lottery of allocations. That is the role of what I call expected feasibility in the BARP.

The outline of the proof of Theorem 3 is as follows. First I solve the BARP. Denote (π, V) as a solution to it. I show that (π, V) satisfies the notion that π is increasing, because J is increasing. By Lemma 4 and π increasing, (π, V) satisfies binding arbitration IC. By Lemma 3 it satisfies binding arbitration IR. Thus, (π, V) is a solution to the binding arbitration problem.

The following problem helps to explain the intuition behind Theorem 3. As shown in the proof of Theorem 3, the BARP can be written as a problem with one constraint linear in π and an objective function also linear in π, as follows.

$$\min_{\pi} \int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\overline{\theta}} \ell(\theta_i, \theta_j)\pi(\theta_i, \theta_j) f(\theta_i) f(\theta_j) d\theta_i d\theta_j$$

subject to

$$\int_{\underline{\theta}}^{\overline{\theta}} \int_{\underline{\theta}}^{\overline{\theta}} \left[\pi(\theta_i, \theta_j) \left(F(\theta_i) \frac{\partial u_i(\theta_i, \theta_j)}{\partial \theta_i} f(\theta_j) - (\ell(\theta_i, \theta_j)/2) f(\theta_i) f(\theta_j)\right) d\theta_i d\theta_j\right]$$
$$= \mathbb{E}_{\theta_j}[u_i(\overline{\theta}, \theta_j)] - 1/2$$

The construction of J, then, reflects the trade-off between satisfying the highest type IR and worsening the objective function (which is now trying to be minimized). The trade-off exists because of the need to maintain incentive compatibility. The mechanism designer has to satisfy the highest type individual rationality. This can be done by giving him/her high allocation or high probability of conflict. Given the envelope condition, increasing π for a given type increases his/her expected utility from participating in the mechanism. Notice, however, that giving a high allocation to the highest type would be the best way to satisfy his/her IR constraint, given the objective function. But high allocation for the highest type would mean that lower types would like to report the highest type. Therefore, to maintain incentive compatibility, the mechanism designer has to balance the incentives that π provides, with the impact they have on the objective function. Increasing π for the highest type provides higher utility to satisfy his/her IR constraint in an incentive compatible way, but it worsens the value of the objective function. The ratio between the "weights" of

π in the objective function with those from the constraint reflects such a trade-off. If that ratio is decreasing, the optimal mechanism starts assigning probability one from top to bottom of the type space as in Figure 2.2, until the highest type IR is satisfied. The ratio is decreasing if J is increasing. Thus the designer "prefers" to put higher probability of conflict in the higher region of the type space.

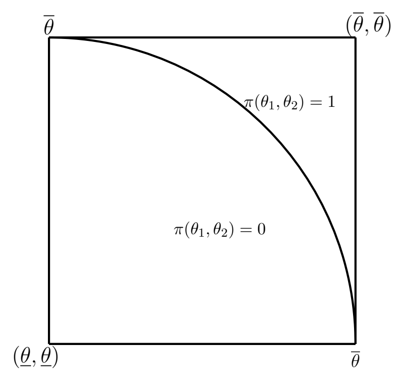

Figure 2.2: The solution to the binding arbitration problem.

The solution to the binding arbitration problem when the efficiency condition does not hold can be regarded as the second best. The optimal mechanism "pools" low types in a "peaceful" region and higher types into conflict. Notice that this is in line with the argument in Theorem 2. To see why the second best mechanism pools high types and low types to the extremes of conflict and zero conflict, respectively, it is important to see the forces from IR and IC of the binding arbitration problem. For that, the BARP is instrumental. The J function illustrates the trade-off of satisfying the highest type while maintaining incentives for truth telling. Since the BARP can be written as one linear constraint problem with linear objective, the ratio between the weights of π in the objective function with those in the constraint yields the trade-off described above. In a pair (θ_1, θ_2) in which such ratio is high, increasing π at (θ_1, θ_2) worsens the objective function more than it helps to satisfy the highest type individual rationality. When this ratio is the lowest (not negative), the designer prefers to start increasing the probability of conflict, because the low ratio indicates that the utility of the highest type is increased by a large amount compared to the worsening of the objective function. When J is increasing, the ratio is decreasing. Hence, the designer "starts" assigning $\pi = 1$ at the highest region of the type space (lowest value of the mentioned ratio) and "stops" when the highest type IR is satisfied with equality.

I turn to the question of whether we can implement a mechanism such as the one described in Theorem 3. By Lemmas 1 and 2, if that were possible, the binding arbitration problem would be equivalent to non-binding arbitration in terms of ex ante efficiency. The discussions of Theorem 1 and Theorem 2 are useful to motivate the proof of Theorem 4. Notice that we are in a scenario similar to that of Theorem 2: a peaceful region and a full conflict region. The only difference is that the solution in Theorem 3 does not necessarily have a rectangle of the type space as a peaceful region. The idea of how to implement such a solution is derived from the idea of the implementation in Theorem 2: concealing information within the peaceful region. The zero conflict region here however may not be easily implementable in terms of concealing information. The role of inefficiency is crucial to be able to implement such a mechanism.

The following simplifies the explanation of the proof of Theorem 4. Given (π, V), we say (π, x) generates (π, V), if

$$V_i(\theta_i) = \mathbb{E}_{\theta_j}\left[\pi(\overline{\theta}_i, \theta_j)u_i(\overline{\theta}_i, \theta_j) + (1 - \pi(\overline{\theta}_i, \theta_j))x_i(\overline{\theta}_i, \theta_j)\right].$$

Theorem 4. *Assume a symmetric model J is increasing, and each u_i is C^1. There is $\ell^* \in (0, 1)$ such that if $\ell \geq \ell^*$, a solution to the binding arbitration is a solution to the non-binding arbitration problem.*

The objective function for both binding and non-binding arbitration problems is the same. Hence, Lemma 1 and Lemma 2 imply that a solution to the non-binding arbitration problem gives a value for the objective function that is at most that attained by a solution to the binding arbitration problem. Therefore, to answer positively to the question of equivalence, it suffices that there is a solution to the binding arbitration problem that satisfies non-binding arbitration IC and IR.

The outline of the proof of Theorem 4 is as follows. I start with a solution (π, V) to the BARP. Then, I show that one can construct an allocation function x so that (π, x) generates V and that there is a (π, g), for which (π, x) satisfies non-binding arbitration IC and IR.

Proof of Theorem 4.

Proof. I use $x_i(\theta_i) = \mathbb{E}_{\theta_j}[x_i(\theta_i, \theta_j)]$, for ease of exposition. Let (π, V) be a solution to the BARP. I construct an allocation function $x : [\underline{\theta}, \overline{\theta}]^2 \to [0, 1]^2$ such that:

i) for any (θ_i, θ_j) such that $\pi(\theta_i, \theta_j) = 0$, $x_i(\theta_i) + x_j(\theta_j) \leq 1$,

ii) (π, x) generates (π, V) and

iii) (π, x) is implementable with non-binding arbitration.

Take (π, V) to be a solution to (BARP).

Denote $\Theta_j^*(\theta_i) = \{\theta_j : \pi(\theta_i, \theta_j) = 0\}$ and $\Theta_j^{*c}(\theta_i)$, its complement relative to $[\underline{\theta}, \overline{\theta}]$. Let $m(\Theta_j^*(\theta_i)) = \int_{\Theta_j^*(\theta_i)} f(\theta_j) d\theta_j$. Define $\tilde{\theta}_i(\theta_j) = \inf\{\theta_i : \pi(\theta_i, \theta_j) = 1\}$ and if $\inf\{\theta_i : \pi(\theta_i, \theta_j) = 1\}$ is empty. Let $\tilde{\theta}_i(\theta_j) = \overline{\theta}$. Denote by $\underline{V}$ the value $V_i(\underline{\theta})$

Using this notation and the fact that the solution can be described as in Theorem 3, we can write

$$\underline{V} = \int_{\Theta_j} u_i(\tilde{\theta}_i(\theta_j), \theta_j) f(\theta_j) d\theta_j$$

Therefore, for any i and any θ_i

$$V_i(\theta_i) = \int_{\Theta_j} u_i(\tilde{\theta}_i(\theta_j), \theta_j) f(\theta_j) d\theta_j + \int_{\Theta^{*c}(\theta_i)} u_i(\theta_i, \theta_j) f(\theta_j) d\theta_j - \int_{\Theta^{*c}(\theta_i)} u_i(\tilde{\theta}_i(\theta_j), \theta_j) f(\theta_j) d\theta_j$$

Thus,

$$V_i(\theta_i) = \int_{\Theta_j^*(\theta_i)} u_i(\tilde{\theta}_i(\theta_j), \theta_j) f(\theta_j) d\theta_j + \int_{\Theta^{*c}(\theta_i)} u_i(\theta_i, \theta_j) f(\theta_j) d\theta_j$$

Hence, I impose the following condition to construct x:

$$\int_{\Theta_j^{*c}(\theta_i)} u_i(\theta_i, \theta_j) f(\theta_j) d\theta_j + \int_{\Theta_j^*(\theta_i)} x_i(\theta_i, \theta_j) f(\theta_j) d\theta_j = \int_{\Theta_j^*} u_i(\tilde{\theta}_i(\theta_j), \theta_j) f(\theta_j) d\theta_j + \int_{\Theta^{*c}(\theta_i)} u_i(\theta_i, \theta_j) f(\theta_j) d\theta_j$$

So, I let:

$$x_i(\theta_i) = \left(1/m\left(\Theta_j^*\left(\theta_i\right)\right)\right) \int_{\Theta_j^*(\theta_i)} u_i(\tilde{\theta}_i(\theta_j), \theta_j) f(\theta_j)$$

Now, take (θ, θ') such that $\pi(\theta, \theta') = 0$. Then, we can write:

$$x_i(\theta) + x_j(\theta') = \frac{1}{m(\Theta_i^*(\theta'))} \int_{\Theta_i^*(\theta')} [(1 - \ell(\theta_i, \theta_j)) - u_i(\theta_i, \tilde{\theta}_j(\theta_i))] f(\theta_i) + \frac{1}{m(\Theta_j^*(\theta))} \int_{\Theta_j^*(\theta)} u_i(\tilde{\theta}_i(\theta_j), \theta_j) f(\theta_j)$$

We want to show that for any (θ, θ') such that $\pi(\theta, \theta') = 0$, $x_i(\theta) + x_j(\theta') \leq 1$.

If we do that, we observe that if (π, V) is a solution to the BARP, then (π, x) with $x(\theta_i, \theta_j) = x_i(\theta_i)$ for all θ_j is a solution to the binding arbitration problem.

I start with

$$x_i(\theta) + x_j(\theta') = (1/m(\Theta_i^*(\theta'))) \int_{\Theta_i^*(\theta')} [(1 - \ell(\theta_i, \theta_j)) - u_i(\theta_i, \tilde{\theta}_j(\theta_i))] f(\theta_i)$$
$$+ (1/m(\Theta_j^*(\theta))) \int_{\Theta_j^*(\theta)} u_i(\tilde{\theta}_i(\theta_j), \theta_j) f(\theta_j)$$

So,

$$x_i(\theta) + x_j(\theta') = (1 - \ell(\theta_i, \theta_j))(1/m(\Theta_i^*(\theta'))) \int_{\Theta_i^*(\theta')} [-u_i(\theta_i, \tilde{\theta}_j(\theta_i))] f(\theta_i)$$
$$+ (1/m(\Theta_j^*(\theta))) \int_{\Theta_j^*(\theta)} u_i(\tilde{\theta}_i(\theta_j), \theta_j) f(\theta_j)$$

Let $q_i(\theta_i, \theta_j)$ be such that $(1 - \ell(\theta_i, \theta_j)) q_i(\theta_i, \theta_j) = u_i(\theta_i, \theta_j)$, then $x_i(\theta) + x_j(\theta')$, equals

$$(1 - \ell(\theta_i, \theta_j)) + (1 - \ell(\theta_i, \theta_j)) \left(1/m(\Theta_j^*(\theta))\right) \int_{\Theta_j^*(\theta)} q_i(\tilde{\theta}_i(\theta_j), \theta_j) f(\theta_j)$$
$$- (1 - \ell(\theta_i, \theta_j)) \left(1/m(\Theta_i^*(\theta'))\right) \int_{\Theta_i^*(\theta')} q_i(\theta_i, \tilde{\theta}_j(\theta_i))] f(\theta_i)$$

Thus,

$$x_i(\theta) + x_j(\theta') < (1 - \ell(\theta_i, \theta_j)) \left[1 - (1/m(\Theta_i^*(\theta'))) \int_{\Theta_i^*(\theta')} q_i(\theta_i, \tilde{\theta}_j(\theta_i))] f(\theta_i)\right]$$
$$+ (1 - \ell(\theta_i, \theta_j))$$

And since $u_i > 0$, we obtain $\int_{\Theta_i^*(\theta')} q_i(\theta_i, \tilde{\theta}_j(\theta_i))]f(\theta_i) > 0$. Hence

$$x_i(\theta) + x_j(\theta') < 2(1 - \ell(\theta_i, \theta_j))$$

So, when $\ell = 1/2$, $x_i(\theta) + x_j(\theta') < 1$.

Then we conclude that there is $\ell^* \in (0, 1/2)$ such that when $\ell \geq \ell^*$, $x_i(\theta) + x_j(\theta') \leq 1$.

Take (π, V) that is a solution to the BARP, and let x be an allocation function as constructed.

By binding arbitration IR, for each θ_i, we obtain

$$\int_{\Theta_j^*(\theta_i)} x_i(\theta_i, \theta_j) f(\theta_j) d\theta_j + \int_{\Theta_j^{*c}(\theta_i)} u_i(\theta_i, \theta_j) f(\theta_j) d\theta_j \geq$$
$$\int_{\Theta_j^*(\theta_i)} u_i(\theta_i, \theta_j) f(\theta_j) d\theta_j + \int_{\Theta_j^{*c}(\theta_i)} u_i(\theta_i, \theta_j) f(\theta_j) d\theta_j$$

thus,

$$\int_{\Theta_j^*(\theta_i)} x_i(\theta_i, \theta_j) f(\theta_j) d\theta_j \geq \int_{\Theta_j^*(\theta_i)} u_i(\theta_i, \theta_j) f(\theta_j) d\theta_j$$

I define $x_i^* : [\underline{\theta}, \overline{\theta}] \to [0,1]$ by $x_i^*(\theta_i) \int_{\Theta_j^*(\theta_i)} f(\theta_j) d\theta_j = \int_{\Theta_j^*(\theta_i)} x_i(\theta_i, \theta_j) f(\theta_j) d\theta_j$. Now I define g^* to be such that, for each θ_i, $g_i^*(x_i(\theta_i)|(\theta_i, \theta_j)) = 1 - \pi(\theta_i, \theta_j)$, for each θ_j. It is clear that (π, g^*), by the construction, satisfies non-binding arbitration IR. Whenever θ_i is so that $\Theta_j^*(\theta_i)$ is not empty (otherwise non-binding arbitration IR is irrelevant), the following holds

$$x_i^*(\theta_i) \geq \frac{\int_{\Theta_j^*(\theta_i)} u_i(\theta_i, \theta_j) f(\theta_j) d\theta_j}{\int_{\Theta_j^*(\theta_i)} f(\theta_j) d\theta_j} = \mathbb{E}_z[u_i(\theta_i, z)|x_i(\theta_i), \theta_i]$$

. Hence, non-binding arbitration IR is satisfied. I now show that (π, g^*) satisfies non-binding arbitration IC. Given any type θ_i and $\theta_i' \neq \theta_i$, we need to verify

$$\begin{aligned}
&\int_{\Theta_j^*(\theta_i)} x_i(\theta_i,\theta_j) f(\theta_j) d\theta_j + \int_{\Theta_j^{*c}(\theta_i)} u_i(\theta_i,\theta_j) f(\theta_j) d\theta_j \geq \\
&\int_{\Theta_j^{*c}(\theta_i')} u_i(\theta_i,\theta_j) f(\theta_j) d\theta_j + \int_{\Theta_j^*(\theta_i')} \max\{x_i(\theta_i'), \mathbb{E}_z[u_i(\theta_i,z)|x_i(\theta_i'),\theta_i']\} f(\theta_j) d\theta_j
\end{aligned} \tag{2.3}$$

Since $x_i(\theta_i')$ is constant across θ_j, we can have only two cases.

First $x_i(\theta_i') \geq \mathbb{E}_z[u_i(\theta_i,z)|x_i(\theta_i'),\theta_i']$ in which case, (2.3) becomes

$$\begin{aligned}
\int_{\Theta_j^*(\theta_i)} x_i(\theta_i,\theta_j) f(\theta_j) d\theta_j + \int_{\Theta_j^{*c}(\theta_i)} u_i(\theta_i,\theta_j) f(\theta_j) d\theta_j \geq \\
\int_{\Theta_j^{*c}(\theta_i')} u_i(\theta_i,\theta_j) f(\theta_j) d\theta_j + \int_{\Theta_j^*(\theta_i')} x_i(\theta_i') f(\theta_j) d\theta_j
\end{aligned}$$

. That is, binding arbitration IC holds for θ_i when he reports θ_i'. Since (π, x) satisfies binding arbitration IC, (π, g^*) satisfies (2.3).

The second case is when $x_i(\theta_i') < \mathbb{E}_z[u_i(\theta_i,z)|x_i(\theta_i'),\theta_i']$. Then, (2.3) becomes

$$\begin{aligned}
&\int_{\Theta_j^*(\theta_i)} x_i(\theta_i,\theta_j) f(\theta_j) d\theta_j + \int_{\Theta_j^{*c}(\theta_i)} u_i(\theta_i,\theta_j) f(\theta_j) d\theta_j \geq \\
&\int_{\Theta_j^{*c}(\theta_i')} u_i(\theta_i,\theta_j) f(\theta_j) d\theta_j + \int_{\Theta_j^*(\theta_i')} u_i(\theta_i,\theta_j) f(\theta_j) d\theta_j = \mathbb{E}_{\theta_j}[u_i(\theta_i,\theta_j)]
\end{aligned}$$

Which holds, since (π, g^*) satisfies binding arbitration IR, because (π, x) does so. Therefore, (π, g^*) is a solution to the non-binding arbitration problem since it maintains π from a solution (π, x) to the binding arbitration and satisfies non-binding arbitration IC and IR. □

Theorem 4 states that under the assumptions described above, a designer who is constrained to conduct non-binding arbitration mechanisms can be as effective, in terms of ex ante efficiency, as a designer who conducts binding arbitration. In particular, the proof shows that this can be done by partitioning the type space into a peaceful region and a conflict region. The recommended allocation can be taken to reveal only that a player is in the peaceful region but not who he is facing. Since the probability of conflict is increasing under J, the optimal mechanism pools lower types into peaceful regions and higher types into conflict regions. Thus whenever an allocation is received, players know they face a lower type but not exactly what type the other player may be.

Theorem 4 is important for at least two reasons. First, if the equivalence between the problems hold, then solving the binding arbitration problem requires only studying the expected probability of conflict and the utilities of the players participating. This is a much easier task than solving a non-binding arbitration problem. Second, Theorum 4 is important in many applications. If equivalence holds, then the power of the institution (the mechanism designer) could be taken as unimportant for the desired objective. This would explain why countries accept binding arbitration in conflict resolution.

For completeness I mention a class of examples where the assumption of J increasing holds.

Remark 3: Suppose for each i, that F_i is a log-concave and absolutely continuous function, and each u_i is C^1, and $\frac{\partial u_i(\theta_i,\theta_j)/\partial\theta_i}{\delta(\theta_i,\theta_j)}$ is increasing in both arguments. Then J is increasing.

Log concave distributions and increasing partial derivatives for the two players' outside options satisfy this assumption with the case of constant ℓ.

In looking for means of breaking the equivalence, my paper suggests that analyzing decreases in J would be productive. Notice that the assumption on J is an assumption jointly on the utilities of the outside option and the distribution of types. It is important to note that without the condition of J increasing, the complexity of solving the binding arbitration problem grows significantly. It is not possible to obtain an IC characterization for binding arbitration IC. Binding arbitration IC in general does not imply π increasing. Neither does it imply that $\mathbb{E}_{\theta_j}\pi(\theta_i, \theta_j)$. The interdependence of outside options implies some expected "co-monotonicity" of π and the outside option. The ironing required to solve this problem is then difficult, since it becomes multidimensional. Restricting attention to $\mathbb{E}_{\theta_j}\pi(\theta_i, \theta_j)$ is not sufficient for incentives.

In spite of all this, some comments are worth mentioning. Motivated by Theorem 4, analyzing the case of J decreasing and low ℓ conveys some intuition for how to break the equivalence between binding and non-binding arbitration. The idea is as follows: with J decreasing, some kind of simplifying needs to be done for binding arbitration IC. Suppose this is done "point wise". Then at the ex post level, some IC constraint must bind. At the ex post level, it is possible to obtain an IC characterization. The conjecture is that some upward misreport must be binding for at least some type (one could actually expect that many IC constraints would bind). With at least one IC constraint binding (for misreporting a higher type), ℓ low means that there is not enough inefficiency to "reallocate" by the mechanism. That is, the expected allocation of misreporting cannot be much bigger in expected terms than that from truth telling. This means that misreporting in non-binding arbitration can strictly improve upon misreporting in binding arbitration. But since the binding arbitration IC was binding, misreporting in non-binding arbitration is strictly better than truth telling. Then the solution to the binding arbitration problem is not implementable with non-binding arbitration, because of the non-binding arbitration IC.

2.5 Conclusions

I consider the question of how to split a pie between two players. The utility of outside options depends on both players' types. I refer to this dependence as the interdependency of outside options. I study the problem from the perspective of designing mechanisms that seek ex ante efficiency. The presence of interdependent outside options brings relevance to the enforcement power of the mechanism designer. I study two cases. First, I consider a case when players cannot be forced to stay with the recommendation of a designer. I call this case non-binding arbitration. I also consider a case where players have to stay with each recommendation of the designer. I call this case binding arbitration. For the case of non-binding arbitration, I consider direct mechanisms that are incentive compatible and satisfy obedience, appealing to the version of the revelation principle in Myerson (1982). In the case of binding arbitration, I restrict attention to incentive compatible mechanisms, appealing to the revelation principle in Myerson (1979). Since direct mechanisms recommend outcomes as functions of the types reported by the players, any recommendation potentially conveys information about the type of the other player. Since outside options are interdependent, this information is useful for the players. With non-binding arbitration, a player can use this information and claim an outside option when it is convenient given his/her updated beliefs. Therefore, the designer who conducts non-binding arbitration has to take into account potential "information leakage."

Outside options are inefficient, in the sense that some part of the pie is lost if the outside option outcome occurs. Since I am interested in efficiency, the first best consists of mechanisms that assign probability zero to the outside option outcome and allocate the pie completely. In Theorem 1, I provide a necessary and sufficient condition (the efficiency condition) for first best mechanisms to be implementable. The efficiency condition is the same for both binding and non-binding arbitration. It can be interpreted as describing balanced forces in conflict. I show that it allows the designer to fully conceal information so that no player knows what type he is facing upon receiving a recommended allocation. This allows the designer to bypass the information leakage problem. I also show that whenever the efficiency condition does not hold, the designer is bound to reveal some information.

I depart from the efficiency condition to analyze efficiency beyond the first best. I study the mechanisms that maximize the ex ante expected sum of utilities of the players of participating in the mechanism. I solve this problem with binding arbitration, in Theorem 3. I also provide a sufficient condition for the solution to the binding arbitration problem to be implementable with non-binding arbitration, in Theorem 4. The condition resembles the efficiency condition, in the sense that it allows the designer to conceal information for the lower types. Theorem 4 means that a designer who conducts non-binding arbitration can be as effective as one who conducts binding arbitration in terms of efficiency. This may explain why countries accept binding arbitration when facing some types of international discord.

The abstract features of the model can accommodate many applications, including those related to international conflict, collusion or cartel formation among firms, and political party rivalry, to mention a few. These applications are significant because of the impact they have on society at large. Determining whether institutions can be effective in arbitrating problems is highly relevant. My work contributes to resolving the efficiency problem across many fields.

The importance of the problem analyzed is also relevant at the technical level. The conjunction of interdependent outside options and non-binding arbitration implies that players can potentially learn information from a given mechanism and use it to veto a recommended allocation. This restricts the designer. The mechanism designer needs to be careful with information potentially revealed. The implications for mechanism design are not adequately explored in the literature.

Some avenues for future research are implied in this paper. Theorem 4 shows a condition for the "equivalence" (in terms of efficiency) between binding and non-binding arbitration. When the condition does not hold, the question of this equivalence remains open. Some insights in this paper may suggest ways of answering the question. Additionally, the information design problem embedded into matter of the mechanism itself is worth exploring . Allowing the designer to send signals to the players could advance the objective of efficiency. Such signals could shape players' beliefs, as noted in Balzer and Schneider (2019). Applying tools from information design literature could also shed light on the equivalence issue. Constrained information design has recently begun to be explored in Doval and Skreta (2018).

REFERENCES

[1] Ausubel, L. M., Cramton, P., & Deneckere, R. J. (2002). Bargaining with incomplete information. In R.J. Aumann & S. Hart (Eds.), *Handbook of game theory with economic applications* (1st ed., vol. 3, pp. 1897-1945). Elsevier. https://EconPapers.repec.org/RePEc:eee:gamchp:3-50.

[2] Balzer, B., & Schneider, J. (2019). Belief management and optimal arbitration. *mimeo*. https://josndr.github.io/Duality.pdf.

[3] Bester, H., & Warneryd, K. (2006). Conflict and the social contract. *The Scandinavian Journal of Economics*, *108*(2), 231-249. https://doi.org/10.1111/j.1467-9442.2006.00452.x

[4] Brito, D. L., & Intriligator, M. D. (1985). Conflict, war and re-distribution. *The American Political Science Review*, *79*(4), 943-957. https://doi.org/10.2307/1956242

[5] Chatterjee, K., & Lee, C. (1998). Bargaining and search with incomplete information about outside options. *Games and Economic Behavior*, *22*(2), 203-237. https://doi.org/10.1006/game.1997.0586

[6] Chatterjee, K., & Samuelson, W. (1983). Bargaining under incomplete information. *Operations Research*, *31*(5), 835-851. https://doi.org/10.1287/opre.31.5.835

[7] Chatterjee, K. & Samuelson, W. (1987). Bargaining with two-sided incomplete information: An infinite horizon model with alternating offers. *The Review of Economic Studies*, *54*(2), 175-192. https://doi.org/10.2307/2297510

[8] Chatterjee, K. & Samuelson, W. (1988). Bargaining under two-sided incomplete information: The unrestricted offers case. *Operations Research*, *36*(4), 605-618. https://doi.org/10.1287/opre.36.4.605

[9] Chikte, S. & Deshmukh, S. (1987). The role of external search in bilateral bargaining. *Operations Research*, *35*(2), 198-205. https://doi.org/10.1287/opre.36.4.605

[10] Compte, O. & Jehiel, P. (2009). Veto constraint in mechanism design: Inefficiency with correlated types. *American Economic Journal: Microeconomics*, *1*(1), 182-206. https://doi.org/10.1257/mic.1.1.182

[11] Cramton, C. (1984). Bargaining with incomplete information: An infinite-horizon model with two-sided uncertainty. *The Review of Economic Studies*, *51*(4), 579-593. https://doi.org/10.2307/2297780

[12] Cramton, P. & Palfrey, T. (1995). Ratifiable mechanisms: Learning from disagreement. *Games and Economic Behavior*, *10*(2), 255–283. https://doi.org/10.1006/game.1995.1032

[13] Das, S. & Sengupta, S. (2001). Asymmetric information bargaining and international mergers. *Journal of Economics and Management Strategy*, *10*(4), 565-590. https://doi.org/10.1111/j.1430-9134.2001.00565.x

[14] Doval, L. & Skreta, V. (2018). Constrained information design: Toolkit. *Working paper*. https://arxiv.org/pdf/1811.03588.pdf

[15] Evans, R. (1989). Sequential bargaining with correlated values. *The Review of Economic Studies*, *56*(4), 499-510. https://doi.org/10.2307/2297497

[16] Forges, F. (1999). Ex post individually rational trading mechanisms. In A. Alkan, C.D. Aliprantis & N.C. Yannelis (Eds.), *Current Trends in Economics. Studies in Economic Theory* (1st ed., vol 8, pp. 157-175). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-03750-8_10

[17] Fudenberg, D. & Tirole, J. (1983). Sequential bargaining with incomplete information. *The Review of Economic Studies*, *50*(2), 221-247. https://doi.org/10.2307/2297414

[18] Fudenberg, D., Levine, D. & Tirole, J. (1987). Incomplete information bargaining with outside opportunities. *Quarterly Journal of Economics*, *102*(1), 37-50. https://doi.org/10.2307/1884679

[19] Goltsman, M. & Pavlov, G. (2014). Communication in Cournot oligopoly. *Journal of Economic Theory*, *153*, 152–176. https://doi.org/10.1016/j.jet.2014.06.008

[20] Grossman, S. & Perry, M. (1986). Sequential bargaining under asymmetric information. *Journal of Economic Theory*, *39*(1), 120-154. https://doi.org/10.1016/0022-0531(86)90023-2

[21] Harsanyi, J. & Selten, R. (1972). Generalized Nash solution for two-person bargaining games with incomplete information. *Management Science*, *18*(5), 80-106. https://doi.org/10.1287/mnsc.18.5.80

[22] Hörner, J., Morelli, M. & Squintani, F. (2015). Mediation and peace. *Review of Economic Studies*, *82*(4), 1483–1501. https://doi.org/10.1093/restud/rdv022

[23] Leininger, W., Linhart, P & Radner, R. (1989). Equilibria of the sealed-bid mechanism for bargaining with incomplete information. *Journal of Economic Theory*, *48*(1), 63-106. https://doi.org/10.1016/0022-0531(89)90120-8

[24] McLean, R. & Postlewaite, A. (2015). Implementation with interdependent valuations. *Theoretical Economics*, *10*(3), 923–952. https://doi.org/10.3982/te1440

[25] Milgrom, P. & Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, *70*(2), 583-601. https://doi.org/10.1111/1468-0262.00296

[26] Myerson, R. (1979). Incentive compatibility and the bargaining problem. *Econometrica*, *47*(1), 61-74. https://doi.org/10.2307/1912346

[27] Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research*, *6*(1), 58-73. https://doi.org/10.1287/moor.6.1.58

[28] Myerson, R. (1982). Optimal coordination mechanisms in generalized principal-agent problems. *Journal of Mathematical Economics*, *10*(1), 67-81. https://doi.org/10.1016/0304-4068(82)90006-4

[29] Myerson, R. (1984). Two-person bargaining problems with incomplete information. *Econometrica*, *52*(2), 461-487. https://doi.org/10.2307/1911499

[30] Polachek, S. & Xiang, J. (2010). How opportunity costs decrease the probability of war in an incomplete information game. *International Organization*, *64*(1), 133-144. https://doi.org/10.1017/S002081830999018X

[31] Reed, W. (2003). Information power and war. *The American Political Science Review*, *97*(4), 633-641. https://doi.org/10.1017/S0003055403000923

[32] Rubinstein, A. (1982). Perfect equilibrium in bargaining model. *Econometrica*, *50*(1), 97-109. https://doi.org/10.2307/1912531

[33] Samuelson, W. (1984). Bargaining under asymmetric information. *Econometrica*, *52*(4), 995-1006. https://doi.org/10.2307/1911195

[34] Shimer, R & Werning, I. (2019). Efficiency and information transmission in bilateral trading. *Review of Economic Dynamics*, *33*, 154-176. https://doi.org/10.1016/j.red.2019.02.004

[35] Skarpedas, S. & Grofman, B. (1995). Modeling negative campaigning. *The American Political Science Review*, *89*(1), 49-61. https://doi.org/10.2307/2083074

2.6 Appendix

Proof of Lemma 1.

Proof. Fix a mechanism (π, g) that satisfies non-binding arbitration IR, then for each i, and each $\theta_i \in \Theta_i$

$x \int_{\Theta_i} f(\theta_j) g_i(x|(\theta_i, \theta_j)) d\theta_j \geq \int_{\Theta_i} u_i(\theta_i, z) f(z) f(x|(\theta_i, z)) dz$ for all $x \in [0,1]$. Integrating over all x, gives

$\int_{[0,1]} x \left[\int_{\Theta_i} f(z) g(x|\theta_i, z) dz \right] dx \geq \int_{[0,1]} \left[\int_{\Theta_i} u_i(\theta_i, z) f(\theta_j) g_i(x|(\theta_i, \theta_j)) d\theta_j \right] dx$. Rearranging gives

$\mathbb{E}_{\theta_j^*} \left[\int_{[0,1]} x g_i(x|(\theta_i, \theta_j^*)) dx \right] \geq \mathbb{E}_{\theta_j^*}[(1-\pi(\theta_i, \theta_j^*)) u_i(\theta_i, \theta_j^*)]$. Adding $\mathbb{E}_{\theta_j^*}[\pi(\theta_i, \theta_j^*) u_i(\theta_i, \theta_j^*)]$ on both sides gives $\mathbb{E}_{\theta_j^*} \left[\pi(\theta_i, \theta_j^*) u_i(\theta_i, \theta_j^*) + \int_{[0,1]} x g_i(x|(\overline{\theta}_i, \theta_j^*)) dx \right] \geq \mathbb{E}_{\theta_j^*}[u_i(\theta_i, \theta_j^*)]$, which implies that (π, g) satisfies binding arbitration IR. □

Proof of Lemma 2

Proof. Fix a mechanism (π, g) that satisfies non-binding arbitration IC, then, for each $i, j = 1, 2$, $i \neq j$, for any θ_i,

$\mathbb{E}_{\theta_j^*} \left[\pi(\theta_i, \theta_j^*) u_i(\theta_i, \theta_j^*) + \int_{[0,1]} x g_i(x|(\theta_i, \theta_j^*)) dx \right] \geq$

$\mathbb{E}_{\theta_j^*} \left[\pi(\theta_i, \theta_j^*) u_i(\theta_i, \theta_j^*) + \int_{[0,1]} \max\{x, \mathbb{E}_z[u_i(\theta_i, z)|x, \theta_i]\} g_i(x|(\theta_i, \theta_j^*)) dx \right] \geq$

$\mathbb{E}_{\theta_j^*} \left[\pi(\theta_i, \theta_j^*) u_i(\theta_i, \theta_j^*) + \int_{[0,1]} x g_i(x|(\theta_i, \theta_j^*)) dx \right]$, where the latter inequality follows by the use of the max operator. Hence, (π, g) satisfies binding arbitration IC. □

Proof of Lemma 3

Proof. Take (π, x) that satisfies binding arbitration IC and binding arbitration IR, for the highest type. Then for all $i = 1, 2$, for all θ_i we have, by binding arbitration IC,

$$\mathbb{E}_{\theta_j}[\pi(\theta_i, \theta_j) u_i(\theta_i, \theta_j) + (1 - \pi(\theta_i, \theta_j)) x_i(\theta_i, \theta_j)] \geq$$
$$\mathbb{E}_{\theta_j}[\pi(\overline{\theta}, \theta_j) u_i(\theta_i, \theta_j) + (1 - \pi(\overline{\theta}, \theta_j)) x_i(\overline{\theta}, \theta_j)]$$

On the other hand, by binding arbitration IR for the highest type, then for all $i = 1, 2$,

$\mathbb{E}_{\theta_j}[(1-\pi(\overline{\theta},\theta_j))x_i(\overline{\theta},\theta_j)] \geq \mathbb{E}_{\theta_j}[(1-\pi(\overline{\theta},\theta_j))u_i(\overline{\theta},\theta_j)$. Using this in the above inequality from binding arbitration IC, we get, for all $i=1,2$, $j\neq i$, all θ_i

$$\mathbb{E}_{\theta_j}[\pi(\theta_i,\theta_j)u_i(\theta_i,\theta_j)+(1-\pi(\theta_i,\theta_j))x_i(\theta_i,\theta_j)] \geq$$
$$\mathbb{E}_{\theta_j}[\pi(\overline{\theta},\theta_j)u_i(\theta_i,\theta_j)+(1-\pi(\overline{\theta},\theta_j))u_i(\overline{\theta},\theta_j)] \geq$$
$$\mathbb{E}_{\theta_j}[\pi(\overline{\theta},\theta_j)u_i(\theta_i,\theta_j)+(1-\pi(\overline{\theta},\theta_j))u_i(\theta_i,\theta_j)],$$

by the assumption of increasing outside options. Thus, we obtain binding arbitration IR. □

Proof of Theorem 3.

Proof. I call the following equation, as before, the expected feasibility constraint and I express it

$$\mathbb{E}_{\theta_i,\theta_j}\left[V_i(\theta_i)+V_j(\theta_j)\right] \leq \mathbb{E}_{\theta_i,\theta_j}\left[\pi(\theta_i,\theta_j)\left(u_i(\theta_i,\theta_j)+u_j(\theta_i,\theta_j)\right)+(1-\pi(\theta_i,\theta_j))\right] \quad (2.4)$$

Note that is without loss of generality to assume $x_1(\theta_1,\theta_2)+x_2(\theta_1,\theta_2)=1$ for all (θ_1,θ_2). To see this, note that, if $x_1(\theta_1,\theta_2)+x_2(\theta_1,\theta_2)<1$ on a strictly positive measure (according to f), one can increase the objective function and relaxing (2.4) at the same time. This means, that the objective function of the binding arbitration problem can be rewritten as:

$$\max_{(\pi,g)} \mathbb{E}_{(\theta_1,\theta_2)}\left[-\ell(\theta_i,\theta_j)\pi(\theta_1,\theta_2)\right]$$

Using the envelope condition of Lemma 4 and symmetry we can rewrite the previous constraint as

$$V_i(\underline{\theta}) + \mathbb{E}_{\theta_i,\theta_j}\left[\int_{\underline{\theta}}^{\theta_i} \pi(\theta_i,\theta_j)\frac{\partial u_i(z,\theta_j)}{\partial z}dz\right] = 1/2 - \mathbb{E}_{\theta_i,\theta_j}[\pi(\theta_i,\theta_j)\ell(\theta_i,\theta_j)/2] \tag{2.5}$$

Notice it is written as equality given the rewritten objective function. To see this, suppose the expected feasibility constraint did not hold with equality, we could increase $V_i(\underline{\theta})$ and decrease π on a strictly positive measure (under f), improving the objective function. Ultimately, integrating by parts (2.5) gives

$$V_i(\underline{\theta}) + \mathbb{E}_{\theta_j}\left[\int_{\underline{\theta}}^{\overline{\theta}} \pi(\theta_i,\theta_j)(1-F(\theta_i))\frac{\partial u_i(\theta_i,\theta_j)}{\partial \theta_i}d\theta_i\right] = 1/2 - \mathbb{E}_{\theta_i,\theta_j}[\pi(\theta_i,\theta_j)\ell(\theta_i,\theta_j)/2] \tag{2.6}$$

On the other hand, as stated before, by Lemma 3, I focus on binding arbitration IR for the highest type only, which using Lemma 4 is written as follows,

$$V_i(\underline{\theta}) + \mathbb{E}_{\theta_j}\left[\int_{\underline{\theta}}^{\overline{\theta}} \pi(\theta_i,\theta_j)\frac{\partial u_i(\theta_i,\theta_j)}{\partial \theta_i}d\theta_i\right] \geq \mathbb{E}_{\theta_j}[u_i(\overline{\theta},\theta_j)]$$

As well as before, it is without loss of generality to write this constraint with equality. Note that if it were not binding, we could decrease π on a strictly positive measure (under f) and hence improve the objective function. So we write

$$V_i(\underline{\theta}) + \mathbb{E}_{\theta_j}\left[\int_{\underline{\theta}}^{\overline{\theta}} \pi(\theta_i,\theta_j)\frac{\partial u_i(\theta_i,\theta_j)}{\partial \theta_i}d\theta_i\right] = \mathbb{E}_{\theta_j}[u_i(\overline{\theta},\theta_j)]$$

Combining this two equations, we obtain only one constraint,

$$\mathbb{E}_{\theta_j}\left[\int_{\underline{\theta}}^{\overline{\theta}} \pi(\theta_i,\theta_j)F(\theta_i)\frac{\partial u_i(\theta_i,\theta_j)}{\partial \theta_i}d\theta_i\right] - \mathbb{E}_{\theta_i,\theta_j}[\pi(\theta_i,\theta_j)\ell(\theta_i,\theta_j)/2] = \mathbb{E}_{\theta_j}[u_i(\overline{\theta},\theta_j)] - 1/2$$

That is,

$$\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\theta}}^{\overline{\theta}}\left[\pi(\theta_i,\theta_j)F(\theta_i)\frac{\partial u_i(\theta_i,\theta_j)}{\partial\theta_i}f(\theta_j)d\theta_i d\theta_j\right]$$
$$-\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\theta}}^{\overline{\theta}}\left[\pi(\theta_i,\theta_j)\left(\ell(\theta_i,\theta_j)/2\right)f(\theta_i)f(\theta_j)d\theta_i d\theta_j\right]=$$

$$\mathbb{E}_{\theta_j}[u_i(\overline{\theta},\theta_j)]-1/2$$

Thus, solving the following problem, gives the solution to the BARP

$$\min_{\pi}\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\theta}}^{\overline{\theta}}\ell(\theta_i,\theta_j)\pi(\theta_i,\theta_j)f(\theta_i)f(\theta_j)d\theta_i d\theta_j$$

subject to

$$\int_{\underline{\theta}}^{\overline{\theta}}\int_{\underline{\theta}}^{\overline{\theta}}\left[\pi(\theta_i,\theta_j)\left(F(\theta_i)\frac{\partial u_i(\theta_i,\theta_j)}{\partial\theta_i}f(\theta_j)-(\ell(\theta_i,\theta_j)/2)f(\theta_i)f(\theta_j)\right)d\theta_i d\theta_j\right]$$
$$=\mathbb{E}_{\theta_j}[u_i(\overline{\theta},\theta_j)]-1/2$$

Note that this problem has an objective function and one constraint that are both linear in π. We then, can optimize pointwise using the following method, the formal discussion is in Lemma 5 and Lemma 6 in the Appendix. For each (θ_i,θ_j) we construct a ratio for $\pi(\theta_i,\theta_j)$ dividing the number multiplying it in the objective function with the one in the constraint. Since each $\pi(\theta_i,\theta_j)\in[0,1]$ we set each of them equal to one, until the constraint holds, starting with those for which the mentioned ratio is minimum. We first point out that by symmetry $\pi(\theta_i,\theta_j)=\pi(\theta_i,\theta_j)$, so this is taken into account to construct the mentioned ratio. Define the function $\beta:[\underline{\theta},\overline{\theta}]\times[\underline{\theta},\overline{\theta}]\to\Re$ as such ratio, thus,

$$\beta(\theta_i,\theta_j)=\frac{2f(\theta_i)f(\theta_j)\ell(\theta_i,\theta_j)}{F(\theta_i)\frac{\partial u_i(\theta_i,\theta_j)}{\partial\theta_i}f(\theta_j)+F(\theta_j)\frac{\partial u_j(\theta_i,\theta_j)}{\partial\theta_j}f(\theta_i)-(\ell(\theta_i,\theta_j))f(\theta_i)f(\theta_j)}$$

or, simplifying, using $f > 0$,

$$\beta(\theta_i, \theta_j) = \frac{2}{J(\theta_i, \theta_j) - 1}$$

Let $Z^* \subset [\underline{\theta}, \overline{\theta}]^2$ to satisfy the following properties. First, set $\pi^*(\theta_i, \theta_j) = 1$ for all $(\theta_i, \theta_j) \in Z^*$ and $\pi^*(\theta_i', \theta_j') = 0$ for all $(\theta_i', \theta_j') \notin Z^*$, and such that

$$\mathbb{E}_{\theta_j}\left[\int_{\underline{\theta}}^{\overline{\theta}} \pi^*(\theta_i, \theta_j)\frac{\partial u_i(\theta_i, \theta_j)}{\partial \theta_i} F(\theta_i) d\theta_i\right] - \mathbb{E}_{\theta_i, \theta_j}[\pi^*(\theta_i, \theta_j)\ell(\theta_i, \theta_j)/2]$$
$$= \mathbb{E}_{\theta_j}[u_i(\overline{\theta}, \theta_j)] - 1/2$$

Second, if $(\theta_i', \theta_j') \notin Z^*$ then $\beta(\theta_i', \theta_j') \geq \beta(\theta_i, \theta_j)$ for any $(\theta_i, \theta_j) \in Z^*$. By the assumptions maintained, since J is increasing, β is decreasing. Thus, (π^*, V^*), defined as above, is increasing and by Lemma 6 in the Appendix, represents the solution to the BARP. By Lemma 3 it satisfies binding arbitration IR. By π^* increasing and Lemma 4, it satisfies binding arbitration IC. To see this, let (π, V) be a solution to the BARP. Let $\theta_i \neq \theta_i'$. The following needs to hold for binding arbitration IC,

$$V_i(\theta_i) \geq \mathbb{E}_{\theta_j}[\pi(\theta_i', \theta_j)u_i(\theta_i, \theta_j) + (1 - \pi(\theta_i', \theta_j))\, x_i(\theta_i', \theta_j)],$$

or

$$V_i(\theta_i) \geq V_i(\theta_i') + \mathbb{E}_{\theta_j}[\pi(\theta_i', \theta_j)\,(u_i(\theta_i, \theta_j) - u_i(\theta_i', \theta_j))].$$

Let $\theta_i \geq \theta_i'$, therefore, the following must be verified,

$$V_i(\theta_i) = V_i(\underline{\theta}) + \mathbb{E}_{\theta_j}\left[\int_{\underline{\theta}}^{\theta_i} \pi(z, \theta_j)\frac{\partial u_i(z, \theta_j)}{\partial z} dz\right] \geq$$

$$Vi(\underline{\theta}) + \mathbb{E}_{\theta_j}\left[\int_{\underline{\theta}}^{\theta_i'} \pi(z, \theta_j)\frac{\partial u_i(z, \theta_j)}{\partial z} dz\right] + \mathbb{E}_{\theta_j}\pi(\theta_i', \theta_j)[u_i(\theta_i, \theta_j) - u_i(\theta_i', \theta_j)].$$

Or,

$$\mathbb{E}_{\theta_j}\left[\int_{\theta_i'}^{\theta_i}\pi(z,\theta_j)\frac{\partial u_i(z,\theta_j)}{\partial z}dz\right]\geq \mathbb{E}_{\theta_j}\pi(\theta_i',\theta_j)[u_i(\theta_i,\theta_j)-u_i(\theta_i',\theta_j)].$$

By monotonicity of π constructed,

$$\begin{aligned}\mathbb{E}_{\theta_j}\left[\int_{\theta_i'}^{\theta_i}\pi(z,\theta_j)\frac{\partial u_i(z,\theta_j)}{\partial z}dz\right]&\geq \mathbb{E}_{\theta_j}\left[\int_{\theta_i'}^{\theta_i}\pi(\theta_i',\theta_j)\frac{\partial u_i(z,\theta_j)}{\partial z}dz\right]\\&=\mathbb{E}_{\theta_j}\pi(\theta_i',\theta_j)[u_i(\theta_i,\theta_j)-u_i(\theta_i',\theta_j)]\end{aligned}$$

Similarly, for $\theta_i \leq \theta_i'$. So the solution proposed satisfies binding arbitration IC and binding arbitration IR. □

Let $X=[\underline{x},\overline{x}]$ be an interval of the real line and $a(x), b(x)$ are continuous functions strictly positive for all $x \in X$ and let $\int_X b(x)dx > A > 0$. Consider the following optimization problem denoted **(E)**. Let $r(x)=\dfrac{a(x)}{b(x)}$. Assume r is monotone.

Consider the following problem.

$$\textbf{(E)}\qquad \min_{f:X\to[0,1]} B(f)=\int_X a(x)f(x)dx$$

s.t.

$$\int_X b(x)f(x)dx\geq A$$

Observation 1: If $b(x)$ is continuous, then $\int_{\underline{x}}^{x} b(z)dz$ is continuous as function of x

Lemma 5. *Suppose f^* is a solution to **(E)**. Then $\int_X b(x)f^*(x)dx = A$.*

Proof. By contradiction suppose f^* is a solution and $\int_X b(x)f^*(x)dx > A$. Let $\epsilon > 0$ be such that $\int_X b(x)f^*(x)dx = A+\epsilon$. Put $\alpha=\dfrac{A}{A+\epsilon}$, then $\alpha\in(0,1)$. Let $f^\alpha=\alpha f^*$. Then $\int_X b(x)f^\alpha(x)dx = A$ and strictly improves the objective. □

Lemma 6. *The following is a solution to* (***E***). *Let* $\alpha^* \in \Re$ *and* f^* *be such that for all* $x \in X$ *whenever* $r(x) \leq \alpha^*$, $f^*(x) = 1$ *and* $f^*(x) = 0$ *otherwise and* $\int_X b(x)f^*(x)dx = A$

Proof. By Observation 1 and monotonicity of r, there is an $\alpha^* \in \mathbb{R}$ such that f^* constructed as in the Lemma accomplishes $\int_X b(x)f^*(x)dx = A$.

Now, by contradiction, suppose there is an $\tilde{f}$ such that $B(\tilde{f}) < B(f^*)$. We can write $\tilde{f}(x) = f^*(x) + \ell(x)$ for all $x \in X$. That is, $\ell(x)$ denotes the change made to f^* in order to obtain $\tilde{f}$. f^* partitions X into X_0 and X_1. With X_0 denoting the set of all x such that $f^*(x) = 0$ and X_1 the set of all x such that $f^*(x) = 1$. By Lemma 5, we can take $\tilde{f}$ to attain the constraint with equality. On the other hand, by construction f^* holds the constraint with equality. It follows that

$$-\int_{X_1} \ell(x)b(x)dx = \int_{X_0} \ell(x)b(x)dx \tag{2.7}$$

or

$$-\int_{X_1} \ell(x)b(x)dx = \int_{X_0} \ell(x)a(x)(1/r(x))dx \tag{2.8}$$

Since $r(x) \geq \alpha^*$, for all $x \in X_0$, then

$$-\int_{X_1} \ell(x)b(x)dx = \int_{X_0} \ell(x)a(x)1/r(x)dx \leq \int_{X_0} \ell(x)a(x)(1/\alpha^*)dx \tag{2.9}$$

On the other hand, taking difference of the objective functions, we obtain

$$B(\tilde{f}) - B(f^*) = \int_X \ell(x)a(x)dx = \int_{X_0} \ell(x)a(x)dx + \int_{X_1} \ell(x)a(x)dx \tag{2.10}$$

Using (2.10)

$$B(\tilde{f})-B(f^*) \geq -\int_{X_1} \ell(x)b(x)\alpha^* dx+\int_{X_1} \ell(x)a(x)dx = \int_{X_1} \ell(x)b(x)(r(x)-\alpha^*)dx \geq 0 \tag{2.11}$$

Note that $\ell(x) \leq 0$ for all $x \in X_1$, since f^* can only be decreased in X_1. At the same time $r(x) \leq \alpha^*$ for all $x \in X_1$ by construction of the solution. Lastly, since $b(x) > 0$ for all x, we obtain the last inequality.

□

Chapter 3

NASH BARGAINING SOLUTION UNDER EXTERNALITIES, WITH GINO LOYOLA

ABSTRACT

The Nash bargaining solution (NBS) of a negotiation is characterized in which two parties experience a linear envy externality. Our results indicate that such a negative externality plays a dual role in the NBS outcome: (i) it increases the probability of a disagreement, and (ii) conditional on an agreement, it constitutes a source of bargaining power. Despite the second result, we show that an increase in envy produces a negative welfare effect because it reduces both parties' NBS payoffs.

3.1 Introduction

One of the main issues in recent research on experimental economics is the role played by feelings and emotional states in bargaining situations (Güth and Kocher (2014), Fischbacher et al. (2013); Kagel and Wolfe (2001); Pfister and Böhm (2012); Pillutla and Murnighan (1996); Sanfey et al. (2003); van 't Wout et al. (2006)). To model the evidence presented in this research, a growing body of game-theory literature has arisen that explores the role of other-regarding elements in the preferences of negotiators in bargaining processes such as envy (Kirchsteiger (1994)), inequity aversion (Fehr and Schmidt (1999)), fairness (Kahneman et al. (1986); Nowak et al. (2000); Xie et al. (2012)), reciprocity (Falk and Fischbacher (2006)) and trust (Berg et al. (1995)). However, all of these studies have focused on decentralized bargaining procedures such as the ultimatum game, the Nash demand game or variants of the alternating offers model, paying little attention to centralized negotiation processes.

To fill this gap, the present article sets out to characterize the sharing rule that a mediator adopting the Nash bargaining solution (NBS) (Nash (1950)) criterion should propose when the parties to a dispute are subject to an envy-based externality. We assume that each party experiences such a negative externality that is proportional to the surplus stake going to his/her counterpart. Our analysis identifies the role played by envy in the bargaining outcome and the conditions for an agreement or disagreement and the characteristics of any agreement that might be arrived at.

More specifically, we show that envy's role in shaping bargaining game outcomes is a dual one. On the one hand, if envy levels are high then disagreement is more likely because the parties are more resistant to mediation proposals involving a concession-making process. On the other hand, conditional on reaching an agreement, the envy externality constitutes a source of bargaining power since it works like a commitment tactic making the envious party a tougher negotiator. This conclusion follows from the fact that while a party's *share* under the NBS is increasing in his/her own envy level, it is decreasing in that of his/her counterpart.

We demonstrate, however, that the foregoing result is just a second-order effect given that an increase in envy produces a negative first-order effect on payoffs and thus ultimately has a negative effect on welfare. In particular, we show formally that an increase in the envy externality worsens both bargainers' well-being as each one's *utility* under the NBS is decreasing in his/her own envy level as well as in that of his/her counterpart.

This paper proceeds as follows. Section 2.2 presents the bargaining problem a mediator must solve when distributing a surplus between two envious parties under the NBS. Section 2.3 characterizes the negotiated solution of this problem, emphasizing the role played by envy in the bargaining outcome. Section 2.4 discusses the main conclusions. All proofs are given in the Appendix.

3.2 Model

Two players, A and B, are bargaining over the division of a surplus $\pi > 0$. The set of feasible agreements is $X = \{(x_A, x_B) : 0 \leq x_A \leq \pi \text{ and } x_B = \pi - x_A\}$. If $(x_A, x_B) \in X$, then x_A is assigned to player A and x_B to player B. The payoffs to A and B are $U_A(x_A, x_B) = x_A - \theta_A x_B$ and $U_B(x_A, x_B) = x_B - \theta_B x_A$, respectively, where $\theta_i > 0$ for each $i \in \{A, B\}$ represents the degree of marginal negative externality experienced by player i due to the portion of the cake that goes to the other player. Disagreement payoffs are given by $d_A \geq 0$ and $d_B \geq 0$.

Now define $\Omega \equiv \{(u_A, u_B) : \text{there is an } x \in X \text{ with } U_A(x) = u_A \text{ and } U_B(x) = u_B\}$ as the set of all possible utility pairs attainable through agreement. For an arbitrary utility of player A, $u_A \in [U_A(0, \pi), U_A(\pi, 0)]$, the corresponding utility for player B whenever there is an $x = (x_A, x_B) \in X$ with $u_A = U_A(x)$ can be obtained as follows. First, rewrite $U_A(x)$ in terms only of x_A as $U_A(x_A) = (1 + \theta_A)x_A - \theta_A \pi$. Next, find the value of x_A that satisfies $U_A(x_A) = u_A$, which turns out to be $x_A(u_A) = \frac{u_A + \theta_A \pi}{1 + \theta_A}$. Then the utility of B given the utility assigned to A is

$$\begin{aligned} g(u_A) &\equiv U_B(x_A(u_A), \pi - x_A(u_A)) \\ &= \frac{1 - \theta_A \theta_B}{1 + \theta_A}\pi - \frac{1 + \theta_B}{1 + \theta_A}u_A. \end{aligned}$$

This in turn means that $\Omega = \{(u_A, u_B) : U_A(0, \pi) \leq u_A \leq U_A(\pi, 0) \text{ and } u_B = g(u_A)\}$. Thus, Ω represents the graph of the function $g : [-\theta_A \pi, \pi] \to \mathbb{R}$, which is represented as a straight red line in Figure 3.1.

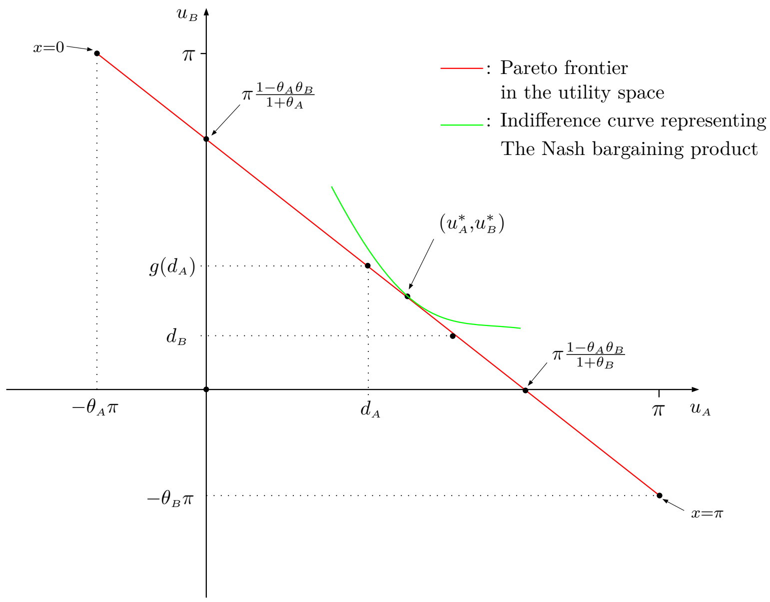

Figure 3.1: Nash bargaining solution under externalities.

We are interested in the Nash Bargaining Solution (NBS) for this setting. Using the previous notation, and denoting $\Theta \equiv \{(u_A, u_B) \in \Omega : u_A \geq d_A \text{ and } u_B \geq d_B\}$, the NBS is the unique pair of utilities (u_A^*, u_B^*) that solves the following problem:

$$\max_{(u_A, u_B) \in \Theta} (u_A - d_A)(u_B - d_B),$$

where the objective function is known as the Nash bargaining product and is represented as a green indifference curve in Figure 3.1. Using g we can find the utility the NBS assigns to player A by solving the following problem:

$$\max_{u_A} (u_A - d_A)\left(\frac{1-\theta_A\theta_B}{1+\theta_A}\pi - \frac{1+\theta_B}{1+\theta_A}u_A - d_B\right)$$

$$\text{subject to}$$

$$u_A \geq d_A$$
$$u_A \leq \frac{1-\theta_A\theta_B}{1+\theta_B}\pi - \frac{1+\theta_A}{1+\theta_B}d_B.$$

As can be seen, the two constraints imply that a necessary condition for the existence of a solution of this problem is given by

$$d_A \leq \frac{1-\theta_A\theta_B}{1+\theta_B}\pi - \frac{1+\theta_A}{1+\theta_B}d_B, \tag{3.1}$$

otherwise the problem would be infeasible.

Now define $\underline{x_i}$ as the minimum share that induces player i to accept an agreement such that $U_i(\underline{x_i}, \pi - \underline{x_i}) = d_i$, from which we obtain

$$\underline{x_i} = \frac{d_i + \theta_i \pi}{1+\theta_i} \tag{3.2}$$

for $i, j \in \{A, B\}$, $i \neq j$. Consider then the following condition

$$\underline{x_A} + \underline{x_B} \leq \pi. \tag{3.3}$$

3.3 Results

We first establish the next auxiliary result.

Lemma 1 *Condition (3.1) is equivalent to condition (3.3).*

Our main result is then stated in the following proposition:

Proposition 1 *If condition (3.3) holds, then the NBS can be described as*

$$\begin{aligned} u_A^* &= \frac{1-\theta_A\theta_B}{2(1+\theta_B)}\pi - \frac{1+\theta_A}{2(1+\theta_B)}d_B + \frac{d_A}{2} \\ u_B^* &= \frac{1-\theta_A\theta_B}{2(1+\theta_A)}\pi - \frac{1+\theta_B}{2(1+\theta_A)}d_A + \frac{d_B}{2}. \end{aligned}$$

Otherwise, there is a disagreement.

Taken together Lemma 1 and Proposition 1 imply that each inequality (3.1) and

(3.3) is both a necessary and a sufficient condition for an agreement. In particular, inequality (3.3) is just the classical condition for an agreement, but adjusted by externalities. There will therefore exist an NBS acceptable for both parties as long as the total surplus to be divided is sufficient to cover at least the sum of their respective minimum aspirations. Note further that[1]

$$\frac{\partial \underline{x_i}}{\partial \theta_i} = -\frac{d_i - \pi}{(\theta_i + 1)^2} > 0.$$

The sign of this partial derivative implies that condition (3.3) is less likely to be satisfied if the envy level θ of either party increases. We conclude then that a negotiation breakdown can occur if the envy levels of parties are sufficiently high.

As noted earlier, we can derive x_A for a given utility u_A as $x_A(u_A) = \frac{u_A + \theta_A \pi}{1+\theta_A}$. With this formula we can also derive the associated share of the cake that is assigned to each player by the NBS. If we denote $x_i^* \equiv x_i^*(u_i^*)$ as the share of the cake assigned to player i, we then obtain the following expressions:

$$\begin{aligned} x_A^* &= \frac{\pi(1+\theta_A(2+\theta_B)) + (1+\theta_B)d_A - (1+\theta_A)d_B}{2(1+\theta_A)(1+\theta_B)} \\ x_B^* &= \frac{\pi(1+\theta_B(2+\theta_A)) + (1+\theta_A)d_B - (1+\theta_B)d_A}{2(1+\theta_A)(1+\theta_B)}. \end{aligned}$$

These shares can also be written to resemble the split-the-difference rule (Muthoo (1999), Chapter 2) as follows

$$x_i^* = \underline{x_i} + \frac{1}{2}(\pi - \underline{x_i} - \underline{x_j})$$

for $i, j \in \{A, B\}$, $i \neq j$. Expressed in this way, the NBS can be seen as delivering

[1]The sign of this derivative is true because condition (3.3) implies that $\underline{x_i} < \pi$, which in turn is equivalent to the inequality $d_i < \pi$.

the minimum aspiration to each player and then just splitting the remaining surplus $\pi - \underline{x_A} - \underline{x_B}$ equally between them.

Corollary 1 *(Source of bargaining power). The shares under the NBS are such that x_i^* is increasing in θ_i and decreasing in θ_j for $i, j \in \{A, B\}$ $i \neq j$.*

Therefore, although envy could be seen at first glance as a sentiment that weakens a negotiator, it is in fact a source of bargaining power. Conditional on reaching an agreement, this externality works as a kind of commitment tactic (Muthoo (1996)) because it constrains a negotiator's capacity to make concessions to the other negotiator, allowing him/her to obtain a larger share of the surplus to be divided and hence reduce the share of his/her counterpart.

Despite the foregoing result, it should be noted that a marginal increase in a party's envy level has a dual impact, affecting positively his/her own share and negatively that of his/her rival (Corollary 1) but also negatively his/her utility with assigned shares held constant. Indeed, as we show formally, the magnitude of the latter impact on utility dominates that of the former on shares. This result is stated in the following corollary:

Corollary 2 *Utilities under the NBS are such that u_i^* is decreasing in θ_i and θ_j for $i, j \in \{A, B\}$ $i \neq j$.*

We therefore conclude that an increase in any player's envy level reduces the utility of both players, and thus, in welfare terms, has ultimately a negative effect.

3.4 Concluding Remarks

It was shown that envy externalities play a dual role in the bargaining outcome of a negotiation conducted under the NBS criterion. Although higher envy levels increase the probability of a disagreement, conditional on reaching an agreement such an externality is a source of bargaining power for the player who experiences it.

This duality opens the door to an extension of our model in which, at a stage previous to applying the NBS, negotiators may strategically over-report the true level of their envy at the risk of provoking a negotiation breakdown. In this new context, designing an efficient and truth-telling bargaining procedure would be an interesting exercise that has yet to be explored.

3.5 Appendix

Proof of Lemma 1. Condition (3.1) is equivalent to

$$(1+\theta_B)d_A + (1+\theta_A)d_B \leq (1-\theta_A\theta_B)\pi.$$

Adding $(\theta_A + \theta_B + \theta_A\theta_B)\pi$ to both sides of the inequality and then rearranging, we obtain

$$\frac{d_A + \theta_A\pi}{1+\theta_A} + \frac{d_B + \theta_B\pi}{1+\theta_B} \leq \pi,$$

which, given the definition of $\underline{x_i}$ in (3.2), completes the proof. □

Proof of Proposition 1. First, we prove the case where condition (3.3) is satisfied. When the problem expressed only in terms of u_A is solved unconstrained, it can readily be seen from the first-order condition that the equation for u_A^* is as shown. The strict concavity of the objective function ensures that the first-order condition gives a unique global maximum. It remains to check the two constraints. We begin by rearranging the NBS:

$$u_A^* = \frac{1}{2}\left(\frac{1-\theta_A\theta_B}{1+\theta_B}\pi - \frac{1+\theta_A}{1+\theta_B}d_B\right) + \frac{1}{2}d_A.$$

It follows that $u_A^* \geq d_A$ since the term in brackets is greater than or equal to d_A because by Lemma 1 if condition (3.3) holds then condition (3.1) also does so. The same naturally applies for $u_B^* \geq d_B$, and thus the second constraint is also satisfied as this is just a rewritten version in terms of u_A. The equation for u_B^* is easily checked by applying g to u_A^*.

Now turning to the case where condition (3.3) is not satisfied, then by Lemma 1, neither is condition (3.1). This implies that the NBS problem is infeasible and thus there exists no pair of utilities that is an acceptable agreement to both parties. □

Proof of Corollary 1. It is easily shown that

$$\frac{\partial x_i^*}{\partial \theta_i} = -\frac{d_i - \pi}{2\left(\theta_i + 1\right)^2}$$

and

$$\frac{\partial x_i^*}{\partial \theta_j} = \frac{d_j - \pi}{2\left(\theta_j + 1\right)^2},$$

for all $i \in \{A, B\}$; $i \neq j$. Condition (3.3) ensures that $\underline{x_i} < \pi$, and therefore that $d_i < \pi$ for each $i \in \{A, B\}$. This implies that the first derivative takes a positive sign and the second a negative one. □

Proof of Corollary 2. It is clear by inspection of the functional forms of (u_A^*, u_B^*) given in Proposition 1 that the utility assigned to a player by the NBS is decreasing in its own degree of externality. To see that the solution is also decreasing in the degree of externality of the other player, we take the partial derivative

$$\frac{\partial u_i^*}{\partial \theta_j} = \frac{\left(\theta_i + 1\right)\left(d_j - \pi\right)}{2\left(\theta_j + 1\right)^2} < 0,$$

for $i, j \in \{A, B\}$ $i \neq j$, where the negative sign holds because condition (3.3) implies $\underline{x_j} < \pi$ and therefore also the inequality $d_j < \pi$. □

REFERENCES

[1] Berg, J., Dickhaut, J., McCabe, K.: Trust, reciprocity, and social history. Games and Economic Behavior, **10**(1), 122-142 (1995)

[2] Falk, A., Fischbacher, U.: A theory of reciprocity. Games and Economic Behavior, **54**(2), 293-315 (2006)

[3] Fehr, E., Schmidt, K.: A theory of fairness, competition, and cooperation. The Quarterly Journal of Economics, **114**(3), 817-868 (1999)

[4] Fischbacher, U., Hertwig, R., Bruhin, A.: How to model heterogeneity in costly punishment: Insights from responders' response times. Journal of Behavioral Decision Making, **26**(5), 462-476 (2013)

[5] Güth, W., Kocher, M.: More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. Journal of Economic Behavior & Organization, **108**, 396-409 (2014)

[6] Kagel, J., Wolfe, K.: Tests of fairness models based on equity considerations in a three-person ultimatum game. Experimental Economics, **4**(3), 203-219 (2001)

[7] Kahneman, D., Knetsch, J., Thaler, R.: Fairness and the assumptions of economics. The Journal of Business, **59**(4), 285-300 (1986)

[8] Kirchsteiger, G.: The role of envy in ultimatum games. Journal of Economic Behavior & Organization, **25**(3), 373-389 (1994)

[9] Nash, J.: The bargaining problem, Econometrica, **18**: 155-162 (1950)

[10] Muthoo, A.: A Bargaining model based on the commitment tactic. Journal of Economic Theory, **69**, 134-152 (1996)

[11] Muthoo, A.: Bargaining Theory with Applications. Cambridge University Press, Cambridge (1999)

[12] Nowak, M., Page, K., Sigmund, K.: Fairness versus reason in the ultimatum game. Science, **289**(5485), 1773-1775 (2000)

[13] Pfister, H.-R., Böhm, G.: Responder feelings in a three-player three-option ultimatum game: affective determinants of rejection behavior. Games, **3**(1), 1-29 (2012)

[14] Pillutla, M., Murnighan, J.: Unfairness, anger, and spite: Emotional rejections of ultimatum offers. Organizational Behavior and Human Decision Processes, **68**(3), 208-224 (1996)

[15] Sanfey, A., Rilling, J., Aronson, J., Nystrom, L., Cohen, J.: The neural basis of economic decision-making in the ultimatum game. Science, **300**(5626), 1755-1758 (2003)

[16] van 't Wout, M., Kahn, R., Sanfey, A., Aleman, A.: Affective state and decision-making in the ultimatum game. Experimental Brain Research, **169**(4), 564-568 (2006)

[17] Xie, W., Li, Y., Wang, Y., Li, K.: Responders' dissatisfaction may provoke fair offer. Journal of Economic Interaction and Coordination, **7**(2), 197-207 (2012)

Chapter 4

CONCLUSIONS

I have considered two types of bargaining problems, between two players that want to split a fixed size "pie." In chapter two, outside options of the players were interdependent, in the sense that they depended on the types of both players. I analyzed the problem from a mechanism design perspective. I showed a necessary and sufficient condition (the efficiency condition) for first best mechanism to be implementable even with asymmetry of information. The efficiency condition can be roughly interpreted as balanced forces at conflict (the outside options). Then, I depart from this condition to find the second best mechanisms. Given the information leakage problem that arises when the efficiency condition does not hold, second best mechanisms may differ depending on the type of arbitration that the mechanism designer can perform. I solve for the second best mechanisms with binding arbitration and find a condition for them to be implementable with non-binding arbitration. Such condition can be interpreted as highly inefficient outside options. This means that as conflict (outside options) become more destructive of the "pie," the designer with non-binding arbitration can be as effective as one with binding arbitration.

As future work opened by the research exposed here, I can mention some observations worth noting. The problem of bargaining with interdependent outside options analyzed, leads to an interesting information problem embedded in the design of the mechanism that seeks to "solve" to the problem. In particular, one could transform the problem into a constrained information design problem. Transforming the problem in this way is interesting in at least two regards. First, it allows to analyze the problem in a more intuitive language of information and, second, allows to analyze a broader set of problems, the so called ratifiable games as in Cramton and Palfrey (1995).

A more concrete route, looking for future from chapter two is to know whether the equivalence between binding and non-binding arbitration can be broken. In particular given the current work there are strong conjectures and some lemmas that lead to think that the mentioned equivalence is broken.

It is also important to mention that the abstract features of the model can accommodate many applications. For example particular applications as in industrial organization in the sense of collusion or cartel formation among firms are interesting to analyze using this paper. Negative campaign in political science is also another route possible to analyze using the tools developed here. These applications are significant because of the impact they have on society at large. Determining whether institutions can be effective in arbitrating problems is highly relevant. My work contributes to solving the efficiency problem across many practical applications as well.

In chapter three, I included a joint work with Gino Loyola. The bargaining problem consisted of splitting a pie (of fixed size) when the players can feel envy about the other players' allocation. The Nash Bargaining Solution (NBS) was obtained and some properties were analyzed. It was shown that envy introduces a source of bargaining power to the players when solving for the NBS. This interesting feature, opens up the field to study the problem when uncertainty about the degree of envy of the other player is included. Because of the properties showed of the NBS, if the designer tries to implement it, incentives lead the players to claim to be very envious. However, at the same time this could lead to break the negotiation. Balancing these two forces is relevant when designing a mechanism in this context. As future work, the analysis of such problem seems promising. The property of the envy serving as bargaining power would contradict the incentives for truthful reports making allegedly impossible to attain the first best as opposed to what happens in chapter two.

REFERENCES

Ausubel, L. M., Cramton, P., & Deneckere, R. J. (2002). Bargaining with incomplete information. In R.J. Aumann & S. Hart (Eds.), *Handbook of game theory with economic applications* (1st ed., vol. 3, pp. 1897-1945). Elsevier. https://EconPapers.repec.org/RePEc:eee:gamchp:3-50

Balzer, B., & Schneider, J. (2019). Belief management and optimal arbitration. *mimeo*. https://josndr.github.io/Duality.pdf

Berg, J., Dickhaut, J. & McCabe, K. (1995) Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*(1), 122-142. https://doi.org/10.1006/game.1995.1027

Bester, H., & Warneryd, K. (2006). Conflict and the social contract. *The Scandinavian Journal of Economics*, *108*(2), 231-249. https://doi.org/10.1111/j.1467-9442.2006.00452.x

Brito, D. L., & Intriligator, M. D. (1985). Conflict, war and redistribution. *The American Political Science Review*, *79*(4), 943-957. https://doi.org/10.2307/1956242

Chatterjee, K., & Lee, C. (1998). Bargaining and search with incomplete information about outside options. *Games and Economic Behavior*, *22*(2), 203-237. https://doi.org/10.1006/game.1997.0586

Chatterjee, K., & Samuelson, W. (1983). Bargaining under incomplete information. *Operations Research*, *31*(5), 835-851. https://doi.org/10.1287/opre.31.5.835

Chatterjee, K. & Samuelson, W. (1987). Bargaining with two-sided incomplete information: An infinite horizon model with alternating offers. *The Review of Economic Studies*, *54*(2), 175-192. https://doi.org/10.2307/2297510

Chatterjee, K. & Samuelson, W. (1988). Bargaining under two-sided incomplete information: The unrestricted offers case. *Operations Research*, *36*(4), 605-618. https://doi.org/10.1287/opre.36.4.605

Chikte, S. & Deshmukh, S. (1987). The role of external search in bilateral bargaining. *Operations Research*, *35*(2), 198-205. https://doi.org/10.1287/opre.36.4.605

Compte, O. & Jehiel, P. (2009). Veto constraint in mechanism design: Inefficiency with correlated types. *American Economic Journal: Microeconomics*, *1*(1), 182-206. https://doi.org/10.1257/mic.1.1.182

Cramton, C. (1984). Bargaining with incomplete information: An infinite-horizon model with two-sided uncertainty. *The Review of Economic Studies*, *51*(4), 579-593. https://doi.org/10.2307/2297780

Cramton, P. & Palfrey, T. (1995). Ratifiable mechanisms: Learning from disagreement. *Games and Economic Behavior*, *10*(2), 255–283. https://doi.org/10.1006/game.1995.1032

Das, S. & Sengupta, S. (2001). Asymmetric information bargaining and international mergers. *Journal of Economics and Management Strategy*, *10*(4), 565-590. https://doi.org/10.1111/j.1430-9134.2001.00565.x

Doval, L. & Skreta, V. (2018). Constrained information design: Toolkit. *Working paper*. https://arxiv.org/pdf/1811.03588.pdf

Evans, R. (1989). Sequential bargaining with correlated values. *The Review of Economic Studies*, *56*(4), 499-510. https://doi.org/10.2307/2297497

Falk, A. & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, *54*(2), 293-315. https://doi.org/10.1016/j.geb.2005.03.001

Fehr, E. & Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817-868. https://doi.org/10.1162/003355399556151

Fischbacher, U., Hertwig & R., Bruhin, A. (2013). How to model heterogeneity in costly punishment: Insights from responders' response times. *Journal of Behavioral Decision Making*, *26*(5), 462-476. https://doi.org/10.1002/bdm.1779

Forges, F. (1999). Ex post individually rational trading mechanisms. In A. Alkan, C.D. Aliprantis & N.C. Yannelis (Eds.), *Current Trends in Economics. Studies in Economic Theory* (1st ed., vol 8, pp. 157-175). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-03750-8_10

Fudenberg, D. & Tirole, J. (1983). Sequential bargaining with incomplete information. *The Review of Economic Studies*, *50*(2), 221-247. https://doi.org/10.2307/2297414

Fudenberg, D., Levine, D. & Tirole, J. (1987). Incomplete information bargaining with outside opportunities. *Quarterly Journal of Economics*, *102*(1), 37-50. https://doi.org/10.2307/1884679

Goltsman, M. & Pavlov, G. (2014). Communication in Cournot oligopoly. *Journal of Economic Theory*, *153*, 152–176. https://doi.org/10.1016/j.jet.2014.06.008

Grossman, S. & Perry, M. (1986). Sequential bargaining under asymmetric information. *Journal of Economic Theory*, *39*(1), 120-154. https://doi.org/10.1016/0022-0531(86)90023-2

Güth, W. & Kocher, M. (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. *Journal of Economic Behavior & Organization*, *108*, 396-409. https://doi.org/10.1016/j.jebo.2014.06.006

Harsanyi, J. & Selten, R. (1972). Generalized Nash solution for two-person bargaining games with incomplete information. *Management Science*, *18*(5), 80-106. https://doi.org/10.1287/mnsc.18.5.80

Hörner, J., Morelli, M. & Squintani, F. (2015). Mediation and peace. *Review of Economic Studies*, *82*(4), 1483–1501. https://doi.org/10.1093/restud/rdv022

Kagel, J. & Wolfe, K. (2001). Tests of fairness models based on equity considerations in a three-person ultimatum game. *Experimental Economics*, *4*(3), 203-219. https://doi.org/10.1023/A:1013290819565

Kahneman, D., Knetsch, J. & Thaler, R. (1986). Fairness and the assumptions of economics. *The Journal of Business*, *59*(4), 285-300. https://doi.org/10.1086/296367

Kirchsteiger, G. (1994). The role of envy in ultimatum games. *Journal of Economic Behavior & Organization*, *25*(3), 373-389. https://doi.org/10.1016/0167-2681(94)90106-6

Leininger, W., Linhart, P & Radner, R. (1989). Equilibria of the sealed-bid mechanism for bargaining with incomplete information. *Journal of Economic Theory*, *48*(1), 63-106. https://doi.org/10.1016/0022-0531(89)90120-8

McLean, R. & Postlewaite, A. (2015). Implementation with interdependent valuations. *Theoretical Economics*, *10*(3), 923–952. https://doi.org/10.3982/te1440

Milgrom, P. & Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, *70*(2), 583-601. https://doi.org/10.1111/1468-0262.00296

Muthoo, A. (1996). A Bargaining model based on the commitment tactic. *Journal of Economic Theory*, *69*(1), 134-152. https://doi.org/10.1006/jeth.1996.0041

Muthoo, A. (1999). *Bargaining Theory with Applications*. Cambridge University Press. https://doi.org/10.1017/cbo9780511607950

Myerson, R. (1979). Incentive compatibility and the bargaining problem. *Econometrica*, *47*(1), 61-74. https://doi.org/10.2307/1912346

Myerson, R. (1981). Optimal auction design. *Mathematics of Operations Research*, *6*(1), 58-73. https://doi.org/10.1287/moor.6.1.58

Myerson, R. (1982). Optimal coordination mechanisms in generalized principal-agent problems. *Journal of Mathematical Economics*, *10*(1), 67-81. https://doi.org/10.1016/0304-4068(82)90006-4

Myerson, R. (1984). Two-person bargaining problems with incomplete information. *Econometrica*, *52*(2), 461-487. https://doi.org/10.2307/1911499

Nash, J. (1950). The bargaining problem. *Econometrica*, *18*(2), 155-162. https://doi.org/10.2307/1907266

Nowak, M., Page, K. & Sigmund, K. (2000). Fairness versus reason in the ultimatum game. *Science*, *289*(5485), 1773-1775. https://doi.org/10.1126/science.289.5485.1773

Pfister, H.-R. & Böhm, G. (2012). Responder feelings in a three-player three-option ultimatum game: Affective determinants of rejection behavior. *Games*, *3*(1), 1-29. https://doi.org/10.3390/g3010001

Pillutla, M. & Murnighan, J. (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes*, *68*(3), 208-224. https://doi.org/10.1006/obhd.1996.0100

Polachek, S. & Xiang, J. (2010). How opportunity costs decrease the probability of war in an incomplete information game. *International Organization*, *64*(1), 133-144. https://doi.org/10.1017/S002081830999018X

Reed, W. (2003). Information power and war. *The American Political Science Review*, *97*(4), 633-641. https://doi.org/10.1017/S0003055403000923

Rubinstein, A. (1982). Perfect equilibrium in bargaining model. *Econometrica*, *50*(1), 97-109. https://doi.org/10.2307/1912531

Samuelson, W. (1984). Bargaining under asymmetric information. *Econometrica*, *52*(4), 995-1006. https://doi.org/10.2307/1911195

Sanfey, A., Rilling, J., Aronson, J., Nystrom, L. & Cohen, J. (2003). The neural basis of economic decision-making in the ultimatum game. *Science*, *300*(5626), 1755-1758. https://doi.org/10.1126/science.1082976

Shimer, R & Werning, I. (2019). Efficiency and information transmission in bilateral trading. *Review of Economic Dynamics*, *33*, 154-176. https://doi.org/10.1016/j.red.2019.02.004

Skarpedas, S. & Grofman, B. (1995). Modeling negative campaigning. *The American Political Science Review*, *89*(1), 49-61. https://doi.org/10.2307/2083074

van 't Wout, M., Kahn, R., Sanfey, A. & Aleman, A. (2006). Affective state and decision-making in the ultimatum game. *Experimental brain research*, *169*(4), 564-568. https://doi.org/10.1007/s00221-006-0346-5

Xie, W., Li, Y., Wang, Y. & Li, K. (2012). Responders' dissatisfaction may provoke fair offer. *Journal of Economic Interaction and Coordination*, 7(2), 197-207. https://doi.org/10.1007/s11403-012-0092-7

APPENDIX A

STATEMENT

I hereby state that Gino Loyola have granted me permission to use our joint work in my thesis.