FUNDAÇÃO GETULIO VARGAS
ESCOLA DE ECONOMIA DE SÃO PAULO

LUIS GUSTAVO DE AZEVEDO GOMES

CONFLICT ARBITRATION WITH ALLIANCES

São Paulo

2021

LUIS GUSTAVO DE AZEVEDO GOMES

CONFLICT ARBITRATION WITH ALLIANCES

Dissertação apresentada à Escola de Economia de São Paulo como pré-requisito à obtenção de título de mestre em Economia de Empresas.

Orientador: Marcos Yamada Nakaguma.

Coorientador: Daniel Monte.

São Paulo

2021

Gomes, Luis Gustavo de Azevedo.
Conflict arbitration with alliances / Luis Gustavo de Azevedo Gomes. - 2021.
53 f.

Orientador: Marcos Yamada Nakaguma.
Co-orientador: Daniel Monte.
Dissertação (mestrado CMEE) – Fundação Getulio Vargas, Escola de Economia de São Paulo.

1. Conflito - Administração. 2. Negociação. 3. Alianças estratégicas (Negócios). 4. Visualização da informação. I. Nakaguma, Marcos Yamada. II. Monte, Daniel. III. Dissertação (mestrado CMEE) – Escola de Economia de São Paulo. IV. Fundação Getulio Vargas. V. Título.

CDU 33

Ficha Catalográfica elaborada por: Isabele Oliveira dos Santos Garcia CRB SP-010191/O
Biblioteca Karl A. Boedecker da Fundação Getulio Vargas - SP

LUIS GUSTAVO DE AZEVEDO GOMES

CONFLICT ARBITRATION WITH ALLIANCES

Dissertação apresentada à Escola de Economia de São Paulo como pré-requisito à obtenção de título de mestre em Economia de Empresas.

Data de aprovação: 28/06/2021

Banca examinadora:

Prof. Dr. Marcos Yamada Nakaguma
FGV-EESP (Orientador)

Prof. Dr. Daniel Monte
FGV-EESP (Coorientador)

Prof. Dr. Gabriel de Abreu Madeira
FEA-USP

Agradecimentos

Em primeiro lugar, gostaria de agradecer aos meus orientadores, os Professores Marcos Yamada Nakaguma e Daniel Monte, que estimularam meu interesse pela pesquisa e por sempre me axuliarem, mesmo em meio à todas as mudanças que se impuseram ao longo desse último ano.

Aos meus pais Heliana e Luiz Gonzaga e ao meu padrasto, Virgílio, por todo seu amor, dedicação e esforço em minha formação. Por seu apoio incondicional durante todo o meu mestrado e por terem sempre alimentado meu desejo por aprender.

Ao Professor Gabriel Madeira, pela sua participação na banca e pelos valiosos comentários e sugestões oferecidos a este trabalho.

Aos professores do mestrado acadêmico da EESP, por toda a sua dedicação e exigência nos cursos ministrados.

Aos funcionários da EESP, que durante todo o período que pudemos frequentar, sempre mantiveram nossos ambientes e equipamentos de estudo em ordem. Um agradecimento especial ao senhor Ademir, o porteiro, que, sempre com seu bom humor, nos avisava do horário de fechamento da sala de estudos.

A todos os meus colegas e amigos da EESP, de minha turma e dos anos anteriores. Nunca imaginei ter vida social tão intensa e ativa durante uma fase que muitos julgam ser tão solitária absorvida pelos estudos. Aos amigos da *salinha*: Pokefilho, Jô, LH, Daniel (o Rei do R) e Fê. Aos amigos da *Famiglia Saligna*: Doutor (Olavo), Doutor (Caio Castro), Tavinho, Flopinho, Alê, Jota, João, José, Matão, Gabiru, Makoto, Nickão, Ogeda e Ariel.

À Yasmin, o raio que sol que ilumina os meus dias, que, mesmo fisicamente distante, esteve sempre ao meu lado. O seu apoio, carinho, afeto e confiança fizeram a minha vida muito mais serena e feliz nos meus últimos meses do mestrado.

À minha querida avó Josefa, que mesmo diante das inúmeras dificuldades que se abateram sobre sua via, nunca abalaram sua insistência e dedicação na educação de seus filhos e netos. Será para sempre o meu maior exemplo de garra e força de vontade para vencer os obstáculos que a vida impõe.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento

de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Quaisquer falhas e omissões são de minha inteira responsabilidade.

Resumo

Este artigo analiza o possível papel da formação de alianças como uma ferramenta para a prevenção de conflitos. Considera-se um ambiente com tipos binários onde três jogadores disputam um montante de tamanho 1, mas nenhum jogador conhece o tipo do outro. Uma terceira parte não-viesada para qualquer jogador, com poder de comprometimento e de forçar suas decisões desenha um protocolo de arbitragem que minimiza a probabilidade de confronto. Mostra-se que, sob certas circunstâncias, alianças não ajudam o árbitro a diminuir a probabilidade de conflito, mas, por meio de simulações, explicita-se situações onde elas podem ser úteis.

Palavras-chave: conflitos, alianças, design de informação, desenho de mecanismo.

Abstract

This paper analyzes the possible role of the formation of alliances as a tool to conflict prevention. We consider an environment with binary types where three players contest a pie of size 1, but no player knows each other's type. An unbiased third party with commitment and enforcement power designs an arbitration protocol that minimizes the probability of engagement. We show that under some circumstances alliances do not allow the arbitrator to decrease the occurrence of conflict, but, through simulation exercises, we provide examples where they could indeed be useful.

Keywords: conflicts, alliances, information design, mechanism design.

List of Figures

Figure 1 – Visualization of Lemma 1 17
Figure 2 – Alliance Configurations 21
Figure 3 – Visual Representation of Arbitration Strategy with Alliances 23
Figure 4 – Expected war payoff of H is $> 1/3$ 41
Figure 5 – Low type IC does not bind 41
Figure 6 – Simulation results on α for report with two L and a single H 47
Figure 7 – Simulation results on α for report with two H and a single L 48
Figure 8 – Simulation results on α and κ for report with two H and a single L . . 49
Figure 9 – Simulation results on α and κ for report with two H and a single L . . 50
Figure 10 – Simulation results on μ for report with two L and a single H 51
Figure 11 – Simulation results on μ for report with two H and a single L 52

Contents

1 INTRODUCTION 9

2 LITERATURE REVIEW 11

3 BENCHMARK MODEL 13

4 FIXING WARS AND ALLIANCES 20
4.1 Probability of victory and sharing rule 20
4.2 The modified arbitration problem 21
4.3 Equivalence of Programs 25

5 WHEN DO ALLIANCES WORK? 29
5.1 Alliances without synergies 29
5.2 Alliances with synergies 31

6 CONCLUSION 33

Bibliography 34

APPENDIX A – PROOFS 37
A.1 Proof of Proposition 3.1 37
A.2 Proof of Proposition 4.1: 38

APPENDIX B – THE GENERAL PROBLEM WITH ALLIANCES . 45

APPENDIX C – SIMULATION RESULTS – FIGURES 47

1 Introduction

Conflicts and alliances between the involved parties are among the most pervasive features of human history. Since conflicts, even in a broader sense, are costly affairs, it makes sense to examine on what are the best tools to help prevent them. As alliances have the power to redistribute strength among the disputant parties and thus, shape the outcomes of a conflict, it is reasonable to think they could be an instrument to achieve a peaceful resolution. It is indeed an ongoing debate on the international relations literature, whether alliances are useful to deter conflicts or induce further conflicts instead. On one side, works such as Leeds (2003), Johnson and Leeds (2011) and Benson (2012) point that alliances can promote peace by raising the costs of waging war against the allied parties and thus, deterring potential aggressors; on the other side, studies like Snyder (1984), Vasquez (1993) and Senese and Vasquez (2008) argue quite the contrary: rival countries, by taking actions to increase their military security (including the formation of alliances), may induce "hostile spirals" that end up increasing the risk of conflict.

In our paper, we tackle this issue with a different approach. Instead of looking if the individual parties are willing to form alliances and if, as a consequence, fewer conflicts erupt, we take the point of view of a central planner whose sole purpose is to decrease the frequency of war occurrence. Our research question is to investigate the alliance formation issue through a mechanism design approach: under which circumstances an unbiased central planner would find alliance formation a useful tool to reduce conflict frequency?

Our modeling approach follows closely that of Hörner et al. (2015), but with elements from Skaperdas (1998) to account for the alliance structures. The environment of our analysis has three players that are on the verge of entering a costly conflict over some surplus. The status quo is such that each player has 1/3 of the surplus. Furthermore, players can be of two types, high or low but each one does not know the other's type. There is an unbiased third party that tries to impede the escalation of the dispute by stipulating whether the players start the conflict or remain in peace. He is an arbitrator, meaning that he can enforce the outcome of his strategy upon the players. This arbitrator, however, does not possess any other information other than what the players provide. In addition, he is committed to the outcome of his strategy, regardless of the reports sent by the players.

Initially, the arbitrator can only specify if the players should engage in conflict or not, but later on, we allow him to fix binary alliances among the players. Alliances have two direct consequences to the players. First, they change the probability of the member players win the dispute, through a contest success function. Second, it implies players must follow a rule to share the remnants of the surplus after the conflict occurs. We also focus on an environment without the possibility of peaceful divisions of the surplus other than the status quo.

We show that alliances are irrelevant to decrease the frequency of conflict if two conditions are met. First, if the share a high type receives when allied to a low type is below the share her strength represents within the alliance and, second, if alliances entail neither positive nor negative synergies between the allied players. These assumptions shape the possible conflict payoffs in a way that alliances worsen the prospects of the high type player, by yielding a disproportionately high share of the spoils to the low type. As a result, if the arbitrator were to use alliance structures, he would need to increase the frequency with which he spurs conflict to ensure the participation of the high type.

Through simulation exercises, we also explore the limits of this result and find examples where alliances are indeed relevant tools and can decrease the overall probability of war. The main insights are that the arbitrator will prefer to form alliances between low and high types if the strength of the low type is positive, although small, and the share received by the high is large; also, alliances between low types are favored if they induce some coordination advantage or disadvantage between the partners.

The remainder of this paper is organized as follows. Chapter 2 makes a brief review on the related literature. Chapter 3 introduces and solves our baseline model. Chapter 4 describes the model including alliances. Chapter 5 presents simulation results and chapter 6 concludes.

2 Literature Review

This study follows closely the modeling approach adopted by Hörner et al. (2015). In their setting, two players dispute over a pie of size one, and each player type is of private knowledge. The player type defines the relative strength if they decide to solve the dispute through a costly war. Players can be of two types: high or low. The only information that is common knowledge is the prior probability of each type. The authors use this setting to show that there exists a communication protocol such that mediated peace talks - where the third party conducting the talks has no enforcement power over the proposed settlement agreement - lead to the same frequency of conflict as the arbitration – where the third party can, indeed, enforce agreements, such as in common judicial disputes.

In this paper, we take advantage of the setting proposed and expand it to fit a conflict with three conflicting parties. Moreover, the goal of this study is different. Our goal is not to compare the effectiveness of the third party under different enforcement capacities, but to argue whether the formation of alliances can help a third party to even further decrease conflicts in an incomplete information setting.

Another important difference in our framework is that, given a set of messages, the third party only puts the disputing players in the status quo or allows them to go to war, and does not propose a settlement agreement to split the surplus peacefully. Under this optic, our third party behaves more similarly to an information designer capable of commitment, such as in Kamenica and Gentzkow (2011), than as an arbitrator as in Hörner et al. (2015) and Bester and Wärneryd (2006). There is indeed some literature applying information design to conflict resolution. A prime example is the paper from Hennigs (2019), which directly applies the Bayesian persuasion approach from Kamenica and Gentzkow (2011) in a two-player dispute. Aside from the number of players, our paper also differs from Hennigs (2019) because we do not assume the arbitrator has the information on the player types, and thus, we focus our attention only on settings where players report their type truthfully to the arbitrator (such as in Hörner et al. (2015)).

The focus on alliance structures under imperfect information is also a novel aspect of this work. As pointed out in Bloch (2012), most of the current theoretic studies on alliances consider a complete information setting (see, for example, Skaperdas (1998),

Jackson and Nei (2015), König et al. (2017) and Hiller (2017)). Indeed, to this date, very few papers have delved into the issue. One of them is Konrad and Morath (2018). Yet, their setting differs from ours as they are interested in the conditions underlying the decision between two players to ally against a common enemy, whereas we proceed in the direction of mechanism and information design. Another study is Manzini and Mariotti (2009), but their focus is on bargains between an already fixed alliance and another player and do not investigate mechanism design.

3 Benchmark Model

Three players contest a pie of size normalized to 1. War decreases the surplus to $\theta < 1$. Each player (we refer to both types as *she*) can be of low (L) or high (H) type. A player's type is private information, only known by the player herself. A player can be of H or L type with probability q and $(1-q)$, respectively. We set the peace payoffs to be the equal split of the full pie, which means each player gets 1/3 of the total surplus in this case.

Whenever a conflict occurs, all three players engage simultaneously. The winner takes all the surplus that is not destroyed in the conflict (θ). Each player's probability of victory depends on her and the other players' strengths. We denote the L and H types strengths as M_L and M_H, respectively, where $M_L, M_H \in \mathbb{R}_+$. We denote the type space $\{M_L, M_H\}$ by Ω. To translate these strengths into victory probability, we use the following rule. If all players are of the same type, they all have a probability of 1/3 to win the conflict. If there are players of distinct types, then the probability of victory is given by the ratio of a player's strength and the sum of the strengths of all players in the contest. In other words, the probability of victory equals the share of a player's strength in the contest.

Throughout this section, we make the assumption that $M_H = 1$ and $M_L = 0$, implying that the low player type always loses a conflict if any of her opponents are of high type. Moreover, a high type player wins all the spoils when her opponents are both of low type and wins with a probability of 1/2 if one of them is of high type. We present a more general version of the model in the next chapter. The important feature of this assumption is that the low type has a very low military strength in relation to the high type.

We also assume that the spoils, θ, fall within a specific interval: $\theta \in (1/3, 2/3)$. Together with our assumptions on the strengths, this implies that the expected payoff of starting the conflict is higher than 1/3 for the high type player if she is the sole high type contesting the pie, and lower than 1/3 if there is at least another high type player contesting the surplus. This means that a high type player would rather start the conflict instead of remaining in peace if she believes her opponents are weak, but she does not

wish to wage war against a strong adversary.

In the benchmark arbitration program, a third party (*he*) collects information privately from the players and makes binding decisions to solve the problem. It is, indeed, highly unlikely that a third party could have such power over sovereign entities such as countries, but it is useful to have a guide on what is the maximum peace attainable. Another reason is that there are other instances where conflict arbitration is indeed possible, such as international trade. Our arbitrator has the sole purpose of decreasing the probability of conflict and, in the benchmark case, his only instrument is to recommend the players to stay in peace (and split the pie evenly) or to wage war.

We can set up the arbitration game as a mechanism design problem. So, after being informed of its type, each player decides whether or not to participate in the arbitration program. The outside option for participation is the conflict among all players. If all players agree to participate, then each player i must send a message $m_i \in \{L, H\}$ to the arbitrator, which could be truthful or not. The arbitrator, now with the reported types in hand then decides, with probability $\pi(m_1, m_2, m_3)$, to incite the conflict among the players.

We focus on equilibria in which the players decide to participate and report their types truthfully. Also, because the optimal arbitration program is linear, then we can restrict our attention to the symmetric solution without loss of generality. Note that symmetry implies that different message vectors but with the same composition of player types are treated equally by the arbitrator – for instance, because of the symmetry, message vectors (L, L, H) and (H, L, L) go to war with the same probability.

The optimal arbitration program determines the probabilities π so as to minimize the probability of conflict among the players:

$$\min_{\boldsymbol{\pi}} \left\{ (1-q)^3\pi_1 + 3q(1-q)^2\pi_2 + 3q^2(1-q)\pi_3 + q^3\pi_4 \right\} \tag{3.1}$$

where $\pi_1 \coloneqq \pi(L, L, L)$, $\pi_2 \coloneqq \pi(L, L, H)$, $\pi_3 = \pi(L, H, H)$, $\pi_4 = \pi(H, H, H)$. This optimization is subjected to the truthful reporting constraints (the incentive compatibility constraints), and to constraints that guarantee the players are willing to participate in the program (the participation constraints). Specifically, the participation constraint of a high type player is given by:

$$(1-q)^2\left(\theta\pi_2+\frac{1}{3}(1-\pi_2)\right)+2q(1-q)\left(\frac{\theta}{2}\pi_3+\frac{1}{3}(1-\pi_3)\right)+q^2\left(\frac{\theta}{3}\pi_4+\frac{1}{3}(1-\pi_4)\right)\geqslant$$
$$(1-q)^2\theta+2q(1-q)\frac{\theta}{2}+q^2\frac{\theta}{3}$$

The left-hand side is the expected payoff of a high type player upon revealing truthfully his type. With probability $(1-q)^2$ the opponents are of a low type, thus if the players enter in conflict – which happens with probability π_2–, then the high type player can capture the totality of the remaining surplus, θ. In case of peace, she receives the status quo payoff of 1/3. With probability $2q(1-q)$, then one of the opponents is of high type and the other, of low type, and so the relevant probability of conflict is π_3. Now, in case of conflict, the high type player splits the remaining surplus with the other high type player, yielding a payoff of $\theta/2$. With probability q^2, the other players are of high type and then, when they start a war, they split equally the remainder of the pie.

The right-hand side is the *ex-ante* payoff of refusing to arbitrate the conflict and hence triggering war. A high type player earn θ when both of her opponents are weak; $\theta/2$, if one is low and the other, high; and $\theta/3$ if the others are all high.

The participation constraint for the low type is, similarly:

$$(1-q)^2\left(\frac{\theta}{3}\pi_1+\frac{1}{3}(1-\pi_1)\right)+2q(1-q)\frac{1}{3}(1-\pi_2)+q^2\frac{1}{3}(1-\pi_3)\geqslant(1-q)^2\frac{\theta}{3}$$

The only major difference with the high type's is the right-hand side. The low type player only earns something from the war if her opponents are also of low type.

The incentive compatibility constraint for the low type is:

$$(1-q^2)\left(\frac{\theta}{3}\pi_1+\frac{1}{3}(1-\pi_1)\right)+2q(1-q)\frac{1}{3}(1-\pi_2)+q^2\frac{1}{3}(1-\pi_3)\geqslant$$
$$(1-q)^2\left(\frac{\theta}{3}\pi_2+\frac{1}{3}(1-\pi_2)\right)+2q(1-q)\frac{1}{3}(1-\pi_3)+q^2\frac{1}{3}(1-\pi_4)$$

Note that the left-hand side is the same as in the participation constraint, and the right-hand side follows a similar logic, but since it represents the case in which the player

is untruthful about his type to the arbitrator – saying she is of high type, in this case – the conflict probabilities have changed. Now, whenever her opponents are both of low type (which occurs with probability $(1-q)^2$) the arbitrator receives a message vector with two lows and one high, and thus, sends them to war with probability π_2. The same reasoning can be applied to explain the other terms of the right-hand side. In an analogous manner, the incentive-compatibility constraint of the high type player is given by:

$$(1-q^2)\left(\theta\pi_2 + \frac{1}{3}(1-\pi_2)\right) + 2q(1-q)\left(\frac{\theta}{2}\pi_3 + \frac{1}{3}(1-\pi_3)\right) + q^2\left(\frac{\theta}{3}\pi_4 + \frac{1}{3}(1-\pi_4)\right) \geqslant$$
$$(1-q^2)\left(\theta\pi_1 + \frac{1}{3}(1-\pi_1)\right) + 2q(1-q)\left(\frac{\theta}{2}\pi_2 + \frac{1}{3}(1-\pi_2)\right) + q^2\left(\frac{\theta}{3}\pi_3 + \frac{1}{3}(1-\pi_3)\right)$$

The solution to this arbitration program features the following properties. Let $\Pi(q,\theta) := (1-q)^2\theta + 2q(1-q)\frac{\theta}{2} + q^2\frac{\theta}{3}$, the *ex-ante* expected payoff of conflict for a high type without the arbitration.

Proposition 3.1. *The arbitration solution for the case without alliance structures is such that :*

(1) The (L, L, L) triad never enters in conflict for all the considered parametric space ($\pi_1 = 0\ \forall(q,\theta) \in [0,1] \times \left[\frac{1}{3}, \frac{2}{3}\right]$).

(2) No triad ever enters conflict if the high type ex-ante expected conflict payoff is less then 1/3 – the peace payoff ($\pi_1 = \pi_2 = \pi_3 = \pi_4 = 0$ if $\Pi(q,\theta) \leqslant \frac{1}{3}$).

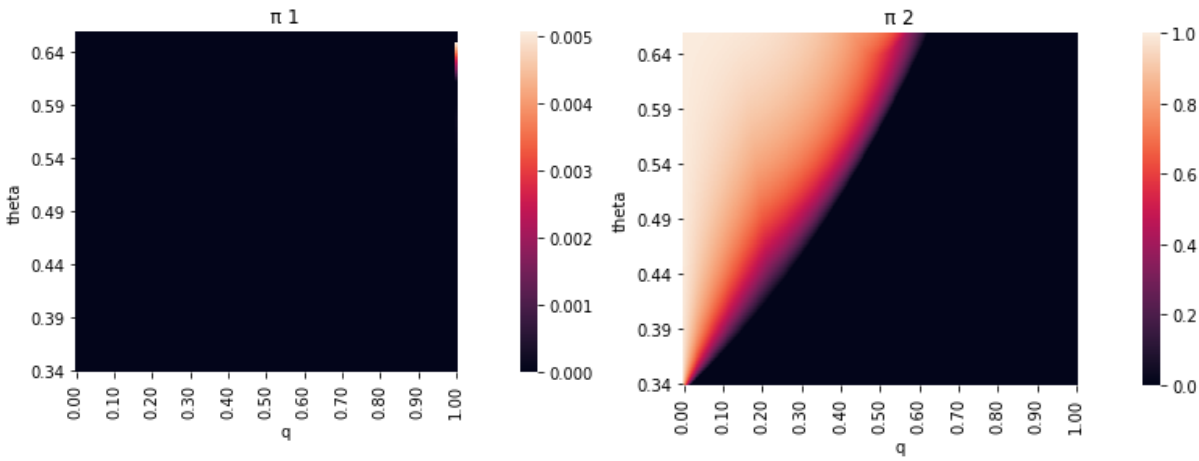
(3) The triads with two low types and a single high type fight with positive probability $\pi_2 > 0$ if ex-ante expected conflict payoff of the high type is larger than 1/3. If, in addition, q and θ are such that $q/(1-q) \leqslant (1-\theta)/2$, then the triads with two high types and a single low type do not fight. If $q/(1-q) > (1-\theta)/2$, then they fight with positive probability.

This lemma can be visualized in the figure 1.

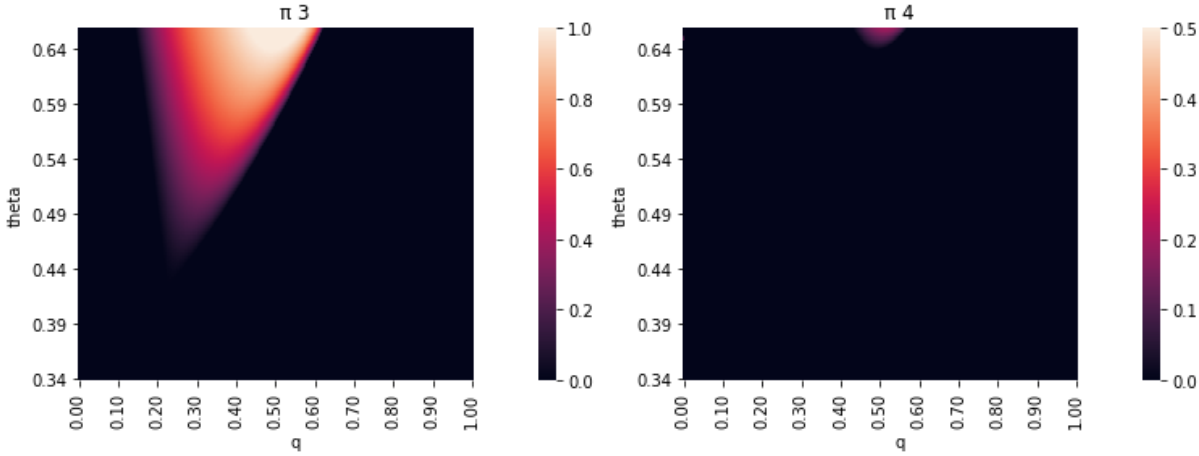
The reasoning behind the arbitration solution is the following. An arbitrator never sends the players to war when he receives the message they are all of low type. This is because low type countries do not benefit from war, and thus increasing the probability

Figure 1 – Visualization of Lemma 1

(a) All L (b) Two L, one H



(c) One L, two H (d) All H



of such conflict reduces the value of participating in the arbitration program. On the other hand, increasing the probability of conflict among low type countries increases the incentive for the high player to lie about her type – because she would then increase the chance of entering a profitable conflict where her opponents are weak.

If a high type player believes there is a high probability that at least one of her opponents is of high type, or in other words, believes her *ex-ante* expected value of conflict to be lower than the peace payoff of 1/3, then the arbitrator can prevent every possible confrontation. The reason being that under this circumstance no player ever wishes to start a war.

Nevertheless, in the opposite case where the high type player has a prior belief

that the conflict will be profitable, the arbitrator must assign a positive war probability whenever she receives the message there are 2 low type countries and a single high player (π_2). The reason is that this is the only composition of player types in which war generates a payoff strictly higher than peace for the high type player. So the arbitrator can at the same time ensure the participation of the high type player and decrease the overall war probability by increasing the likelihood of this conflict, which is the most profitable for a high type player.

Also, while the proportion of high type players is low, this strategy serves as a counter for the low type player to lie about her type: if there are few high type players and a low type player lies, chances are her opponents are of low type, but since the arbitrator receives a message with two low types and a single high type, there is a high probability the arbitrator will incite war. Because of the high threat of conflict, the low type player has the incentive to tell the truth.

Of course, this effect diminishes as the proportion of high types increases, since the chance that at least one of the opponents is of high type becomes large and thus, the low type player has an incentive to lie, so that the reported type composition received by the arbitrator has two high type countries and a single low type. Because inciting this war decreases the expected arbitration payoff of the high type, the arbitrator does not want to increase the chances of it occurring. However, he must do so to prevent the low type from being untruthful about her type. In fact, the incentive for lying for the low type player keeps increasing with the proportion of high type players.

The main idea conveyed by this result is that the biggest challenge the arbitrator has to overcome to reduce the occurrence of conflict is to guarantee the participation of the high type player. If the high type player anticipates the conflict to be profitable, the arbitrator must make him weakly better off accepting the program than going straight into war. In this sense, arbitration is able to reduce the occurrence of wars by increasing the chances that upon entering a conflict, a high type player will face only weak opponents.

Obviously, wars are not profitable to low type countries and so, if there is a mechanism that decreases wars, she will surely join in. The issue is that, in knowing that the arbitrator wishes to diminish the chances a high type player encounters another high type in a war, a low type player faces an incentive to misreport her type. For this reason, the arbitrator must increase, in some cases, the probability of war among high type dominant triads.

One could draw a parallel between our mechanism design setting and that of Bayesian Persuasion. The arbitrator acts similarly to a sender that designs a signal (the war probabilities) conditional on what he observes from nature (the strength of each player). He designs these signals so that the agent with the deviant preference (the high type player, that wishes to begin the confrontation) becomes indifferent between choosing his *ex-ante* preferred action (going straight to war) and joining the arbitration program. The major difference is, notably, that the sender does not directly observe the composition of player types, but has to rely on reports made by the very same player, which imposes a constraint on how much he can decrease the overall war probability.

4 Fixing Wars and Alliances

Now, we introduce to the benchmark model the possibility to change the war payoff by introducing alliances. An alliance is a coalition of players that share both their military resources and the spoils remaining in the aftermath of the conflict.

In addition to fixing conflict or peace to the players, the arbitrator can now specify a military alliance between any two of them. This means that the arbitration strategy also includes a distribution over possible alliance configurations conditional on the reports of types. We assume the outside option of this arbitration program to be the start of a conflict without any alliance structure.

4.1 Probability of victory and sharing rule

Let I denote the set of players. An "alliance structure" is a partition A of the set of players. For instance, with $I = \{1, 2, 3\}$, a possibility for A is $\{\{1, 2\}, \{3\}\}$. The set of all possible partitions is denoted by $\mathcal{A}$. An "alliance" is an element a of partition A. Let a_i denote the alliance in partition A that includes player i – keeping our example, we have that $a_1 = \{1, 2\}$, meaning that 1 and 2 are allied to each other.

An alliance has two implications. First, it means that each player's victory probability now considers the strengths of the other players, and second, that allied players divide the spoils of war via a sharing rule. The probability of victory of a player i is a function $f^i : \Omega^3 \times \mathcal{A} \to [0, 1]$, that we assume to follow the axiomatization proposed by Skaperdas (1998), and to take the specific form, similar to that suggested by Tullock (2001), of

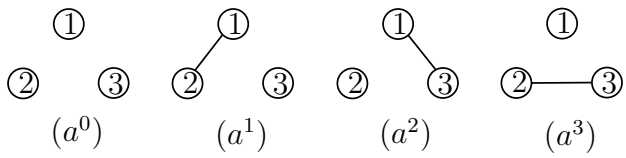
$$f^i((M_i)_{i=1}^3, A) = \begin{cases} \dfrac{\left(\sum_{j \in a_i} M_j\right)^\mu}{\sum_{a \in A}\left(\sum_{j \in a} M_j\right)^\mu}, & \text{if } \exists\ M_i \neq 0 \\ \dfrac{|a_i|^\mu}{\sum_{a \in A} |a|^\mu}, & \text{if } M_i = 0\ \forall\ i\ \in I \end{cases} \tag{4.1}$$

where $|a|$ denotes the number of members in alliance set a. This particular form makes the probability of victory to depend on the ratio between the strengths of the weak and the strong player (so that, from now on, we define $\kappa := M_L/M_H$). μ, which

we assume to be strictly positive, has the property of capturing possible coordination advantages or disadvantages. $\mu > 1$ carries the meaning of "superadditiveness", or, as we call henceforth,"positive synergies", that is, the strength of the alliance is greater than the sum of its individual strengths, whereas $\mu < 1$ means the allies hamper each other's efforts – "negative synergies". When $\mu = 1$ we have the situation where the alliance does not bring any coordination advantage or disadvantages – in other words, where the effort of the group is exactly the sum of efforts of its members.

We consider alliances that either associate the players in a pairwise manner or leave them alone (empty alliance structure). For the case under study, where $I = \{1, 2, 3\}$, this means we exclude the partition $\{\{1, 2, 3\}\}$. The reason for so is that, in case of conflict, the opposing sides are clear. See figure 2 for a graphical representation.

Figure 2 – Alliance Configurations



For now, we remain mostly agnostic about the sharing rule being used by the players, with the remark that: if the allied players are found to be of the same type, they split the spoils evenly, and if they are of a different type, the high type player is allocated an amount $\alpha \in [0, 1]$ of the spoils.

4.2 The modified arbitration problem

As in the benchmark configuration, an unbiased third party privately collects information from the players and then makes binding recommendations to them. The difference is that the arbitrator now has the power to force alliances between two players and then, choose to start or not the conflict, given the chosen alliance structure. The arbitrator is unbiased in the sense he wishes to decrease as much as possible the probability of conflict. He is constrained in an *ex-ante* manner by the participation and incentive compatibility constraints of the low and high type players.

The timing of the mechanism is the following. First, the arbitrator devises his strategy for the given values of the parameters: the high type prior, q, the remaining

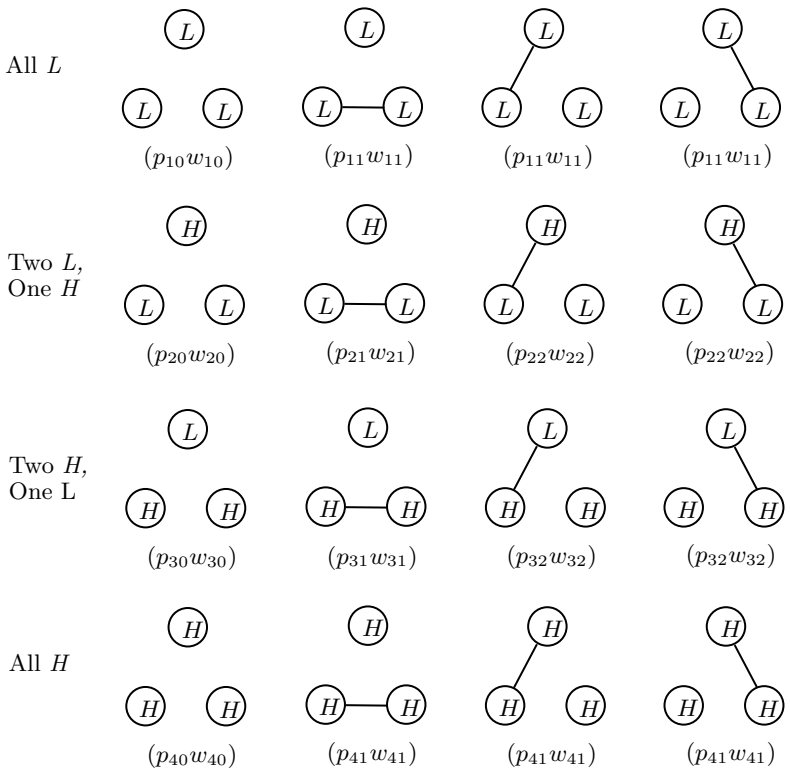
surplus after war, θ, the military strength parameters, κ and μ, and the allies' sharing rule, α. The players start the game without any alliance. Given the arbitration strategy, the players then decide whether or not to accept it. If at least one of them disagrees, there is conflict without any alliances imposed, so the players go to war all by themselves. If all of them agree, then each decides whether or not to send a truthful report about her type. Given the reported types, the arbitrator uses his strategy to prescribe an alliance structure to the players and to decide whether or not to send these players to war.

War is a type reveling phenomenon: even if a player lies to the arbitrator about her type, if war occurs and it turns out she is allied to a player of a different type, she receives the share according to her true type, not her reported type. For instance, suppose a low type player lies and reports she is of high type; in addition, suppose the arbitrator prescribes her to be allied with a player who, after the conflict, turns out to be of high type. Then the untruthful low type player will still receive the fraction $1 - \alpha$ of the spoils. In other words, lying about your own type does not mean you will receive the spoils according to the type you reported.

The strategy of the arbitrator consists of two conditional probability distributions. First a probability distribution $p : \Omega^3 \to \Delta(\mathcal{A})$ stipulating a probability to each possible alliance structure given a vector of reported types and second, $w : A \times \Omega^3 \to [0, 1]$ fixing a probability with which conflict starts given a message vector in a given alliance structure. Hence, we must have that $\sum_{A \in \mathcal{A}} p(A|\mathbf{m}) = 1$, where $\mathbf{m}$ denotes the vector of type reports received by the arbitrator.

As in the benchmark model, we restrict our focus to the symmetric and "anonymous" solutions. By anonymous, we mean that the solution probability distributions only take into account the composition of a received type report, and not, for instance, if player 3 reported to be of high type. In other words, given an alliance configuration A, the type reports (M_H, M_L, M_L) and (M_L, M_H, M_L) are assigned to go to war with the same probability , i.e, $w(A, (M_H, M_L, M_L)) = w(A, (M_L, M_H, M_L))$. By symmetric, we mean those alliance configurations where the allied players have the same composition occur with the same probability. For example, consider the alliance structure where players 1 and 2 are allied and another where 1 and 3 are allied. Then if 1 reports to be of low type and 2 and 3 report to be of high type, the arbitrator assigns these players to those alliances with equal probability. From now onward, we will refer to each probability in the arbitrator's strategy following the notation described in figure 3.

Figure 3 – Visual Representation of Arbitration Strategy with Alliances



Notes: Each circle represents a player and each row represents a different composition of players. Between parenthesis we write the probability with which each triad is sent to war given their alliance structure.p_{ij} is the probability with the arbitrator fixes the alliance structure j to the vector of reported types i and w_{ij} is the probability with which the arbitrator establishes that the message vector with composition i with alliance structure j goes to war.

The objective of the arbitrator is to minimize the occurrence of conflict by choosing p and w, respecting participation and incentive compatibility constraints of the low and high type players. The forms of these constraints follow the same logic as in the benchmark case, but now the payoff associated with each possible composition of adversaries also takes into account the payoffs induced by the alliance structures available. Despite not being linear on either p or w, the objective function and constraints are linear on the products between each p and w, so that we can indeed characterize the arbitration problem as a linear programming problem.

To explain how to formulate the constraints, take, for instance, the position of a high type player who chooses to truthfully report her type. With probability $(1-q)^2$ her two adversaries are of low type. in such case, there are four possible alliance structures. The first is the scenario with no alliances whatsoever (the partition with all players

alone), which the arbitrator assigns to the (M_H, M_L, M_L) message vector with probability $p20 \equiv p(A_0|M_H, M_L, M_L)$. If these players go to war , which happens with probability $w_{20} \equiv w(A_0, (M_H, M_L, M_L))$, then the high type wins the spoils θ with probability $1/(2\kappa^\mu + 1)$. If instead, the arbitrator assigns her to a low type player – occurring with probability $2p_{22}$[1] $\equiv 2p(a_2; M_H, M_L, M_L)$–, then with probability $(1+\kappa)^\mu/((1+\kappa)^\mu + \kappa^\mu)$ she wins a fraction α of the spoils θ. We proceed with this reasoning to compose all terms of the players' expected payoff of participation under truth-telling.

If, however, a player decides to lie, she will face the probability distributions according to her reported type, but since war is a type revelation phenomenon, she will face the war payoffs according to the true distributions of types. For example, assume our player is of high type but reports to be of low type. With probability $2q(1-q)$ one of her opponents is low and the other, high. Nevertheless, the arbitrator believes to be in a situation where there is only one high type player. Then, he fixes reported low and high type players as allies with probability $2p_{22}$. As our original player reported to be of low type with probability p_{22} she will be in an alliance with the high type player. With probability w_{22}, conflict occurs, types are revealed and our high type player wins, in expectation, $[2^\mu/(2^\mu + \kappa^\mu)] \cdot \theta/2$. And we proceed this way to find the expected payoff for each player type.

As stated before, if players refuse arbitration, they automatically start the conflict and face the *ex-ante* expected war payoff, which assumes no alliances are formed.

Hence, the participation constraints are fulfilled if the players' expected payoff of participation under truth-telling is larger than their respective *ex-ante* expected war payoffs, and the incentive compatibility constraints are respected if the participation with truth-telling yields an expected payoff larger than participating and lying. Note that each player type has its own participation and incentive compatibility constraint. For the low type, we refer to them as LPC and LIC, respectively, and for the high type player, as HPC and HIC. For their specific functional forms, refer to appendix B.

Note that the benchmark model can be replicated in this new version if κ is set to zero, $\mu = 1$, to one and the arbitrator places all the weight on the alliance structure without any pairwise alliance (using our notation scheme, $p_{i0} = 1$ for every possible combination of types i).

[1] the factor of "2" refers to the fact that the arbitrator could ally the high type player to either one of the two low type players. See figure 3

4.3 Equivalence of Programs

In this chapter, we will discuss and show that despite gaining an additional instrument for negotiation, the possibility to fix alliances cannot improve over the benchmark scenario, in some cases. The main reason being, if alliances do not provide a sharing rule sufficiently favorable for the high type, then alliances cannot induce the high type player to accept a smaller incidence of conflict.

We illustrate this point by proving this result for the case where: the alliances do not provide any coordination advantage ($\mu = 1$), where the low type player receives a disproportionately large fraction of the spoils in relation to her relative military strength, and where the spoils from war (θ) assume intermediate values – more specifically, $\theta \in ((2\kappa + 1)/3, (2 + \kappa)/3)$[2].

Proposition 4.1. *The optimal arbitration program with alliances cannot decrease the probability of conflict any more than in relation to the program without alliances if they provide no synergies, and the high type share of spoils (α) is smaller than $1/(1 + k)$.*

The full proof of this proposition is available in the appendix A.2.

The idea behind this is to justify why the arbitrator would put all the weight of the distribution p on the empty alliance (or in a payoff equivalent alliance structure), given the goal of minimizing the probability of conflict.

Since it is not possible to make any transfer if a peaceful outcome occurs, then to minimize the probability of conflict, the arbitrator must put weight on the alliance structures that most increase the arbitration payoff of each player, or that most decrease their incentives to lie.

Given our parameters, we already know that the low type player has no incentive whatsoever to start a conflict from an ex-ante perspective. Since she is weaker than the high type and conflict destructs part of the surplus, the low type is content with the status quo, this means that the arbitrator does not need to increase the probability of conflict to appease the low type player. In addition, as in the benchmark model, there is no incentive to spur confrontation when all players report to be of low type, because if they are truthful, then this conflict yields a strictly smaller payoff than peaceful outcome and also because

[2] Note this assumption translates to the assumption adopted in the previous chapter, $\theta \in (1/3, 2/3)$, if the low type has an insignificant strength ($\kappa = 0$)

spurring conflict between reported low types increases the incentive for a high type to misreport since there would be a chance for the high type to end up in a profitable conflict.

Also, because of the higher strength of the high type and the lack of synergies between the allies, the remaining arbitration war payoffs of a low type player are at least as large as the war payoffs of entering the conflict all by herself. Thus, because arbitration removes the risk of conflict when all of them claim to be of low type, the low type participation constraint is not binding.

On the other hand, there is also no incentive to the high type player to lie about her type. That is because all the conflicts that can occur in case of a lie from the high type player are not profitable. In specific, if a high type player lies about her type, she will never enter a war against the players if they reveal to the arbitrator to be of low type. Meanwhile, if she reveals truthfully her type, there is a chance to end up in a conflict with two low type players.

So we can focus only on the two remaining constraints: the low type incentive compatibility and the high type participation.

Regarding the high type participation constraint, an interesting feature from the benchmark model is maintained: the arbitrator only fixes a positive probability of conflict if the *ex-ante* expected war payoff of the high type player is higher than the status quo payoff (1/3). The reasoning follows the same logic as before: since the incentive compatibility constraints are always fulfilled with a zero probability of conflict (i.e., by setting all w's to zero), we need only to ensure the participation of the high type. If war yields an expected outcome worse than the status quo, this means the high type player does not need to be convinced to put down arms, and so the arbitrator does not need to incur any conflict.

However, in the opposite case where the high type player wishes to start the wage war from an *ex-ante* perspective, the arbitrator must decrease the probability of conflict by providing her with the highest expected payoff in case of war. In other words, what the arbitrator tries to do is to keep the same *ex-ante* payoff of the high type, but concentrates weight on the most profitable conflict, and thus, reduces its frequency. The conflict with the highest profit for the high type happens precisely when the adversaries are two low type players and the high type player is left alone. So now, instead of occurring every time, there is at least one high type among the players, now, wars tend to occur more when the high type's adversaries are of low type. The low type players being alone or allied

do not alter, in this case, the probability the high type has of winning the conflict (since there are no coordination gains or losses and we fixed that whenever allies are of the same type, they split the spoils equally), and alliances with a low type player always hinder the high type, as a direct result of our assumption on parameters α and κ. Specifically, the condition $\alpha < 1/(1+\kappa)$ implies the following:

$$\alpha < \frac{1}{1+\kappa} \implies \alpha \frac{M_H + M_L}{M_H + 2M_L}\theta < \frac{M_H}{M_H + 2M_L}\theta \tag{4.2}$$

and, therefore, it is better for a high type player *not* to be allied with a low type player, independently of the types of the other players, but in particular if the remaining players are of low type. And, since the two alliance structures where the high type is left alone (either when all players are alone or the low types are allied) yield the same payoff and one of them is the empty alliance structure, then the arbitrator cannot decrease any further the probability of conflict through this channel.

However, this point alone is not enough to justify our result. We must also verify that the easiest way to ensure the low type reveals correctly her type is also without the use of alliances. Turns out that if we consider exclusively the low type's incentive compatibility constraint, this goal would be more easily achieved by allying a low and a high type when there are two reported high types and a single low type.

The reason is that our assumption $\alpha < 1/(1+\kappa)$ also implies that the low type player earns a higher fraction of the war spoils if allied to a high type than alone. Notice that because of the bounds on θ, an alliance with a high type is still worse for the low type than the status quo. That is, the H-L alliance purely implies that the loss from conflict for the low type will be smaller than if she had no allies. Since the low type player would end up allied to a high type if she chooses to reveal her type truthfully, she will lose a smaller fraction of the surplus if war is declared; and if she misreports her type to the arbitrator, then she will end up either alone or allied to a true low type player – situations in which our untruthful low type player loses a larger amount of the surplus. In contrast, if the arbitrator chooses not to use alliances then even if the low type reveals truthfully her type, she will lose a larger part of the surplus if a conflict occurs.

Nonetheless, the last paragraph's reasoning fails to account for the indirect effect that such alliance has on the participation of the high type. Because of the alliance, the high type has a decrease in her war payoff, which, in turn, presses the arbitrator to increase

the probability of conflict even further when the opponents are of low type. This indirect effect is large enough to offset the reduction in conflict generated by the use of alliances to ease the low-type revelation.

Finally, there is also no advantage in using alliances when all countries claim to be of high type because it is inconsequential if the players are reporting truthfully and it incentives the low type to lie about her type, as she would suffer a smaller loss from war because of the alliance with a high type.

In summary, since the arbitrator has no strict incentive to use alliances, he sees no other way to minimize the probability of conflict other than reproducing the benchmark program.

One discussion that remains is the extent our result depends on its assumptions. Notice that proposition 4.1 only characterizes a case in which the arbitrator is no better off by having at his disposal the possibility to fix alliances. We showed that if the share of spoils a high type receives if in an alliance with a low type is less than what the ratio high type's strength represents in the alliance then the arbitrator cannot achieve a smaller frequency of conflicts by allying the players. A corollary of this result is that, if we considered, such as in the previous chapter, that the strength of the low type player is zero ($\kappa = 0$), then for any sharing rule $\alpha < 1$, alliances are not useful.

Of course, the no synergies condition ($\mu = 1$) also plays a role. It ensures that the strengths of each player do not interfere with each other providing any benefit or disadvantage stemming purely from the alliance. As such, the most important impact of this assumption is that alliances between players of the same type increase the chance of success linearly. Combined with the fact that alliances between players of equal type share equally the spoils, it implies that the players only care about their alliances and not if the remaining players are allied or not.

5 When do alliances work?

In this chapter, we provide simulation exercises that confirm and broaden the results obtained in proposition 4.1 for a larger array of parameter values and, additionally, provide examples of cases where the arbitrator would indeed use alliance structures instead of sticking to the benchmark solution. Despite representing specific cases of the parametric space, they help to highlight which circumstances are sufficient to make alliance structures useful to increase the occurrence of peaceful outcomes.

We solve the linear programming problem of the arbitrator for the same parameter plane as the benchmark problem, $q \times \theta$ with $\theta \in [1/3, 2/3]$. We exhibit the results for the products $p_{ij}w_{ij}$ – recall that p_{ij} is the probability that message report i is put in the alliance structure j and w_{ij} is the probability that the players in such alliance structure are set to go to war. We compare the results of a simulation abiding by the assumptions of our proposition above ($\kappa = 0,\ \alpha = 0.5$ and $\mu = 1$, henceforth referred to as "standard scenario") with another five variations. For the first three, we keep the model without any coordination advantage or disadvantage ($\mu = 1$), and we study what happens when we vary separately κ and α. In the first two exercises, we show how changing α when κ remains zero does not alter qualitatively the results concerning the use of alliances, and on the third, we highlight a case where both κ and α attain a value high enough to induce the use of alliance structures.

In the last two simulations, we then, go on to verify what happens when we fix $\kappa = 0$ and $\alpha = 0.5$, but change the values for μ. We compare between the cases of coordination disadvantages ($\mu < 1$) and advantages ($\mu > 1$).

In these exercises, we abstain from plotting the results for reports where all players state to be of the same type (all low or all high types) because they exhibit no significant difference from the standard scenario.

5.1 Alliances without synergies

In our two first simulation exercises, we consider the cases where the high type player is infinitely stronger than the low type, and alliances do not hinder or improve the

success chance of any player, but we consider two extreme cases for the sharing rule. First, we consider the case where the high type player receives a smaller share of the spoils if allied to a low type – $\alpha = 0.1$ – and in the second, we assume, instead, she receives a much larger share –$\alpha = 0.9$. The arbitration solutions are compared in figures 6 and 7.

As we can see by the figure, there is no noticeable difference between the arbitration protocols with different sharing rules. More importantly, if the arbitrator receives the report with two low types and a single high type, he does as prescribed by proposition 4.1 and alternates between using the empty alliance structure and allying the two low types – as we discussed in the previous section, he is indifferent between using any of these alliance structures and this is why the simulation shows the arbitration protocol putting weight on both. The reason for this invariance is, as discussed above, that with $\kappa = 0$, there is no sense whatsoever in splitting any fraction of the war spoils with the low type player.

We can, indeed, obtain an arbitration strategy that places positive weights on alliance structures by changing the relative strength of the low type, and the amount of spoils received by the high type accordingly. Here, to exemplify, we chose $\alpha = 0.95$ and $\kappa = 0.1$. By looking upon figures 8 and 9, one can observe the stark contrast between the problem under the standard scenario and the new parameter values. Now, the arbitrator uses both the alliance between a high and a low type when there are two low types and a single high type ($p_{22}w_{22} > 0$) and the alliance between high types when there are two high types and a single low type.

The reason for the use of the first of the alliances aforementioned (H-L alliance when there are two low types and one high type) seems to be that the sharing rule and the relative strengths assume values that violate the assumption made on proposition 4.1. Specifically, the sharing rule in this exercise fixes the amount of spoils received by the high type to be greater than what her strength represents in a hypothetical conflict against a low type. In symbols, $\alpha = 0.95$ and $\kappa = 0.1$ are such that:

$$\alpha > \frac{1}{1+\kappa} = \frac{M_H}{M_H + M_L} \tag{5.1}$$

It happens that this inequality also implies that the amount of spoils a high type receives if she goes to war allied to a low type is larger than if she had gone by herself:

$$\alpha > \frac{M_H}{M_H + M_L} \iff \alpha \frac{M_H + M_L}{M_H + 2M_L}\theta > \frac{M_H}{M_H + 2M_L}\theta$$

As a result, the arbitrator can keep the participation constraint of the high type binding at the same time he decreases the conflict whenever he receives a report containing two low type and a single low type message.

In addition, the arbitrator uses the alliance between high types upon receiving a report containing a single L type and two H types, because it allows the low incentive compatibility constraint for the low type to bind with a smaller frequency of conflict. What happens, in this case, is that, with a sharing rule so unfavorable to the low type, it is worse if she lies about her type and she ends up allied with a high type if the remaining player is of low type. Also, because we have fixed $\mu = 1$, the high type is indifferent between remaining alone or allied to another high type if the other player is of low type. Thus, the arbitrator prefers to use alliances in this case.

It is also worth noticing that the region with a positive probability of conflict is smaller than in the case with $\kappa = 0$. This is because the larger is κ, the worst it becomes for the high type player to start a conflict based solely on her priors. In fact, it can be shown by algebra that if $\mu = 1$ and $\kappa = 0.5$, then the high type would weakly prefer to stick with the status quo payoff of 1/3 than engaging in conflict, and so the players always remain in peace.

5.2 Alliances with synergies

Now we turn to the analysis of the cases where alliances provide benefits or hinder the effectiveness of the players involved. In both simulations, we stick to the standard scenario conditions of the egalitarian sharing rule and $\kappa = 0$, but we compute the arbitration solution for both $\mu = 2$, where the strength of the alliance is greater than the sums of its components' strength, and $\mu = 0.5$, where each allied player hinders the effectiveness of the other. The results can be seen in figures 10 and 11.

It is interesting to see that even when the low type countries are completely ineffective ($\kappa = 0$), it is still better for the arbitrator to use alliance structures. Notice that instead of splitting between no alliances and to ally the low type players when the composition of the type report has two low types and a single high type (as in the case of

proposition 4.1), there is a defined region where the arbitrator uses either both or only one of these two alliance structures. In fact, the region where he is indifferent between these alliances corresponds to the region where the low incentive compatibility constraint does not bind, and the region where he chooses unequivocally either one or the other is the region where LIC binds. The high type participation constraint plays no part in this result. The reason is that, because $\kappa = 0$, the alliance between low types is inconsequential in the presence of a high type player. Nevertheless, as we explain below, $\mu \neq 1$ affects the war payoffs of a low type if she chooses to lie and the other players are also of low type.

For a small frequency of high types (low values of q), the low incentive compatibility constraint does not bind, and the sole reason for inducing war is to ensure the participation of the high type player. Since the low type player has zero strength, the high type player is indifferent between the empty alliance structure and an alliance between the other players if they are both of low type. On the other hand, when q becomes large enough, the incentive compatibility for the low type binds, and to avoid receiving a misreport of the low type, the arbitrator must choose the worst possible punishment for an untruthful low type. When there is a positive synergy between the allies ($\mu = 2$), the worst punishment will be to leave the untruthful low type alone and to ally the other two players. On the opposite case, where allies hinder each other ($\mu = 0.5$), the untruthful low type is put in a worse position if all players are left alone. In both cases, using alliances between a low and a high type are not advantageous to reduce conflict because they would simply reduce the participation payoff of a high type country, since the low type does not provide any benefit to the success in the conflict and yet again, the high type must share the spoils equally with her ally.

6 Conclusion

In this paper, we have studied the possible role that pairwise alliances between a total of three players can have in the arbitration of costly disputes an incomplete information environment with binary types, an unbiased third party with enforcement power and commitment, without the possibility of peaceful settlement agreements other than the equal split of the total surplus. We have shown that the greatest difficulty the arbitrator has to overcome in this scenario is to ensure a high type player accepts his negotiation strategy. As such, the arbitrator acts as an information designer that alerts the high player of profitable conflicts (when the other remaining players are of low type) while keeping the incentive to the low type players to remain truthful. In the sequence, we have shown through a formal result and simulation exercises that alliances are only better to diminish the occurrence of conflict if either the high type receives a disproportionately high share of the spoils or the alliance affects the coordination capabilities of its members.

Of course, many options to further investigate this issue remain open. First and foremost, the formal characterization of the parametric space where alliances are useful tools for conflict arbitration should be completed. In addition, it is also important to verify how our results change when the arbitrator can also propose splits for the surplus if he decides to leave players at peace (and thus, follow a more classical formulation of an arbitration problem, such as in Bester and Wärneryd (2006) and Balzer and Schneider (2019)). Another possible analysis is to make the model more applicable to situations where the players are sovereign entities and investigate how a mediator, who cannot enforce his recommendations, would use alliances among players.

Bibliography

Balzer, B. and J. Schneider (2019). Belief management and optimal arbitration. Technical report, Working Paper.

Benson, B. V. (2012). *Constructing international security: alliances, deterrence, and moral hazard*. Cambridge University Press.

Bester, H. and K. Wärneryd (2006). Conflict and the social contract. *Scandinavian Journal of Economics 108*(2), 231–249.

Bloch, F. (2012, April). Endogenous formation of alliances in conflicts. In M. R. Garfinkel and S. Skaperdas (Eds.), *Oxford Handbook of Economics of Peace and Conflict*. Oxford University Press.

Hennigs, R. (2019). Conflict prevention by bayesian persuasion. Working Paper of the Max Planck Institute for Tax Law and Public Finance No. 2019-16.

Hiller, T. (2017). Friends and enemies: A model of signed network formation. *Theoretical Economics 12*(3), 1057–1087.

Hörner, J., M. Morelli, and F. Squintani (2015). Mediation and peace. *The Review of Economic Studies 82*(4), 1483–1501.

Jackson, M. O. and S. Nei (2015). Networks of military alliances, wars, and international trade. *Proceedings of the National Academy of Sciences 112*(50), 15277–15284.

Johnson, J. C. and B. A. Leeds (2011). Defense pacts: A prescription for peace? *Foreign Policy Analysis 7*(1), 45–65.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*(6), 2590–2615.

Konrad, K. A. and F. Morath (2018). To deter or to moderate? alliance formation in contests with incomplete information. *Economic Inquiry 56*(3), 1447–1463.

König, M. D., D. Rohner, M. Thoenig, and F. Zilibotti (2017). Networks in conflict: Theory and evidence from the great war of africa. *Econometrica 85*(4), 1093–1132.

Leeds, B. A. (2003). Do alliances deter aggression? the influence of military alliances on the initiation of militarized interstate disputes. *American Journal of Political Science 47*(3), 427–439.

Manzini, P. and M. Mariotti (2009). Alliances and negotiations: An incomplete information example. *Review of Economic Design 13*(3), 195–203.

Senese, P. D. and J. A. Vasquez (2008). *The steps to war*. Princeton University Press.

Skaperdas, S. (1998). On the formation of alliances in conflict and contests. *Public Choice 96*(1), 25–42.

Snyder, G. H. (1984). The security dilemma in alliance politics. *World politics 36*(4), 461–495.

Tullock, G. (2001). Efficient rent seeking. In *Efficient rent-seeking*, pp. 3–16. Springer.

Vasquez, J. A. (1993). *The war puzzle*. Number 27. Cambridge University Press.

Appendix

APPENDIX A – Proofs

A.1 Proof of Proposition 3.1

First we rewrite the linear program so that we isolate each variable's coefficient the LHS of each constraint:

$$[LPC]: \ (1-q)^2(\theta-1)\pi_1 - 2q(1-q)\pi_2 - q^2\pi_3 \geqslant (1-q)^2(\theta-1) - 1 \tag{A.1}$$

$$[LIC]: \ (1-q)^2(\theta-1)\pi_1 + [-2q(1-q) - (1-q)^2(\theta-1)]\pi_2 - [q^2 - 2q(1-q)]\pi_3 + q^2\pi_4 \geqslant 0 \tag{A.2}$$

$$\begin{aligned}[HPC]: \ & (1-q)^2\left(\theta - \frac{1}{3}\right)\pi_2 + 2q(1-q)\left(\frac{\theta}{2} - \frac{1}{3}\right)\pi_3 + q^2\left(\frac{\theta}{3} - \frac{1}{3}\right)\pi_4 \geqslant \\ & (1-q)^2\theta + 2q(1-q)\frac{\theta}{2} + q^2\frac{\theta}{3} - \frac{1}{3}\end{aligned} \tag{A.3}$$

$$\begin{aligned}[HIC]: \ & -(1-q)^2\left(\theta - \frac{1}{3}\right)\pi_1 + \left[(1-q)^2\left(\theta - \frac{1}{3}\right) - 2q(1-q)\left(\frac{\theta}{2} - \frac{1}{3}\right)\right]\pi_2 + \\ & \left[2q(1-q)\left(\frac{\theta}{2} - \frac{1}{3}\right) - q^2\left(\frac{\theta}{3} - \frac{1}{3}\right)\right]\pi_3 + q^2\left(\frac{\theta}{3} - \frac{1}{3}\right)\pi_4 \geqslant 0\end{aligned} \tag{A.4}$$

Proof of (1):

Because $\theta < 1$, but $\theta > 1/3$, all coefficients associated with π_1 in the constraints are negative and thus whenever we decrease it, we relax all the restrictions. Thus the arbitrator will always choose $\pi_1 = 0$ □

Intuitively, this result comes from the fact that an increase in π_1 raises the chance of "unprofitable" wars occurring – if agents are truthful, we are raising the probability that a war between low types occur, which is worse than peace for them – and incentives the high type to lie.

Proof of (2):

Note that if we assume the solution is $\pi_1 = \pi_2 = \pi_3 = \pi_4 = 0$, we have that all LHS of the restrictions above become zero, so we only need to check for which values of q and θ this condition holds. Obviously, $0 \geqslant 0$ so LIC and HIC hold for all values of q and θ. Since $\theta < 1$, $(1-q)^2(\theta-1) < 0$ and thus $(1-q)^2(\theta-1) - 1 < 0$, therefore LPC also holds for every value of q and θ.

Hence the only condition we need to enable the solution $\pi_1 = \pi_2 = \pi_3 = \pi_4 = 0$ is that $0 \geqslant (1-q)^2\theta + 2q(1-q)\frac{\theta}{2} + q^2\frac{\theta}{3} - \frac{1}{3}$, which comes from HPC.

Because the arbitrator always wants to choose the smallest possible values for π_1, π_2, π_3 and π_4, if the constraints allow him to choose all π's to be equal zero, them he will choose to set them all to zero. □

Proof of (3):

In order to prove the statements, first we prove that the LPC constraint is not binding, then we prove our result for pi_2 and following, we show that HIC cannot bind. Finally we demonstrate our result for π_3 .

In order to show that LPC is not binding, note that without π_1, the left hand side of LPC is strictly decreasing in π_2 and π_3. Note that in the minimum point of the left hand side of LPC, we have that $-2q(1-q) - q^2 \geqslant (1-q)^2(\theta - 1) - 1$. Also note that $(1-q)^2 + 2q(1-q) + q^2 = 1$, but since we assume $\theta < 2/3 < 1$, we must have that $(1-q)^2(\theta-1) + 2q(1-q) + q^2 < 1$ and thus $-[(1-q)^2(\theta-1) + 2q(1-q) + q^2] > -1$. Hence we conclude that for $\theta < 1$, LPC is always inactive.

If the *ex-ante* expected conflict payoff of the high type is larger than 1/3, this requires the lhs of HPC to be strictly positive. But since the only π_i whose coefficient is positive is π_2 we must have that $\pi_2 > 0$ even if pi_3 and π_4 were zero.

Intuitively this makes sense because in order to make the arbitration appealing for the H-type player, he must earn more in expectancy participating on the arbitration than not, this translates as the arbitrator letting some profitable war occur.

The lhs of HIC comprises of the lhs of HPC, which we assumed to be positive, subtracted of terms depending on pi_2 and π_3 – given we have already established that $\pi_1 = 0$. Note that the coefficients of those terms will be positive, because if a high type lies, the wars spurred by π_2 and π_3 are not profitable, so increasing these probabilities serves as punishment for lying. Since we've shown above that if the rhs of HPC is positive so must be π_2, we conclude that HIC does not bind at the solution.

Thus, the problem can be restricted to the HPC and LIC constraints.

Regarding π_3, first note that the coefficient associated with π_2 in LIC is weakly positive whenever $q/(1-q) \leqslant (1-\theta)/2$. So, under this condition, the arbitrator can fulfill all the constraints only by using wars spurred by π_2. On the opposite case, where $q/(1-q) > (1-\theta)/2$, the coefficient associated with π_2 is negative, and so, in order to fulfill the LIC constraint, the arbitrator must use either π_3 or π_4, as they have positive coefficients in LIC.

□

A.2 Proof of Proposition 4.1:

Proof:

First we start by analyzing the relaxed problem where we find the optimal arbitration strategy only considering the low type incentive compatibility constraint and the high type participation constraint. We show that the war probability generated by this strategy cannot be higher than the war probability without the possibility of alliances considering also the low participation and high incentive compatibility constraints. Notice we still refer to the notation scheme explained in figure 3.

Because of the possibility of allocating between different alliance structures, the war probability (the arbitrator's objective function) becomes:

$$\begin{aligned} W(\mathbf{p},\mathbf{w}) = (1-q)^3 (w_{10}p_{10} + 3w_{11}p_{11}) + \\ +3q(1-q)^2 (w_{20}p_{20} + w_{21}p_{21} + 2w_{22}p_{22}) + \\ +3q^2(1-q)(w_{30}p_{30} + w_{31}p_{31} + 2w_{32}p_{32}) + \\ +q^3(w_{40}p_{40} + 3w_{41}p_{41}) \end{aligned} \tag{A.5}$$

where p_{ij} is the probability with the arbitrator fixes the alliance structure j to the vector of reported types i and w_{1j} is the probability with which the arbitrator establishes that the message vector with composition i with alliance structure j goes to war. $i = 1$ refers to the vector where all players report to be low (L, L, L); for $i = 2$, the arbitrator receives the vector with one player reporting high and the other two players to be of low type (L, L, H). $i = 3$ refers to the reported vector with two highs and a single low (L, H, H) and $i = 4$, to when all players report to be of high type (H, H, H).

$j = 0$ indicates the probability associated with no alliance structure, that is, the empty coalition. $j = 1$ refers to the alliance where either two low types or two high types are allied. $j = 2$ indicates alliances with players of different types; they appear twice in the distribution , because, for instance if player 1 is of high type and 2 and 3 are of low type, then an alliance of a H and an L type player could be fixed between either 1 and 2 or 1 and 3.

In the specific case of the vectors where all players are of a single type, because the arbitrator cannot distinguish between the players, alliances 1 and 2 must receive the same weight

The incentive compatibility constraint of the low type states that a low type player must not wish to participate in the arbitration program reporting he is of high type:

$$\begin{aligned} &(1-q)^2\left[\left(\frac{\theta}{3}-\frac{1}{3}\right)w_{10}p_{10} + \left(\frac{\theta}{3}-\frac{1}{3}\right)3w_{11}p_{11}\right] + \\ &+2q(1-q)\left[\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)w_{20}p_{20} + \left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)w_{21}p_{21} + \left(\frac{1}{2}\frac{\kappa+1}{2\kappa+1}(1-\alpha)\theta + \frac{1}{2}\frac{\kappa}{2\kappa+1}\theta - \frac{1}{3}\right)2w_{22}p_{22}\right] + \\ &+q^2\left[\left(\frac{\kappa\theta}{2+\kappa}-\frac{1}{3}\right)w_{30}p_{30} + \left(\frac{\kappa\theta}{2+\kappa}-\frac{1}{3}\right)w_{31}p_{31} + \left(\frac{\kappa+1}{2+\kappa}(1-\alpha)\theta - \frac{1}{3}\right)2w_{32}p_{32}\right] \\ &\geqslant (1-q)^2\left[\left(\frac{\theta}{3}-\frac{1}{3}\right)w_{20}p_{20} + \left(\frac{\theta}{3}-\frac{1}{3}\right)w_{21}p_{21} + \left(\frac{\theta}{3}-\frac{1}{3}\right)2w_{22}p_{22}\right] + \\ &+2q(1-q)\left[\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)w_{30}p_{30} + \left(\frac{\kappa+1}{2\kappa+1}(1-\alpha)\theta - \frac{1}{3}\right)w_{31}p_{31} + \left(\frac{1}{2}\frac{\kappa\theta}{2\kappa+1} + \frac{1}{2}\frac{\kappa\theta}{2\kappa+1} - \frac{1}{3}\right)2w_{32}p_{32}\right] + \\ &+q^2\left[\left(\frac{\kappa\theta}{\kappa+2}-\frac{1}{3}\right)w_{40}p_{40} + \left(\frac{\kappa\theta}{\kappa+2}-\frac{1}{3}\right)w_{41}p_{41} + \left(\frac{\kappa+1}{\kappa+2}(1-\alpha)\theta - \frac{1}{3}\right)2w_{41}p_{41}\right] \end{aligned}$$

Similarly, the high participation constraint for the high type establishes that a high type player will prefer to participate in the arbitration program reporting his type truthfully to engaging in conflict

against the other players alone:

$$
\begin{aligned}
&(1-q)^2\left[\left(\frac{\theta}{2\kappa+1}-\frac{1}{3}\right)p_{20}w_{20}+\left(\frac{\theta}{2\kappa+1}-\frac{1}{3}\right)p_{21}w21+\left(\frac{\kappa+1}{2\kappa+1}\alpha\theta-\frac{1}{3}\right)2p_{22}w_{22}\right]+\\
&+2q(1-q)\left[\left(\frac{\theta}{\kappa+2}-\frac{1}{3}\right)p_{30}w_{30}+\left(\frac{\theta}{\kappa+2}-\frac{1}{3}\right)p_{31}w_{31}+\left(\frac{1}{2}\frac{\theta}{\kappa+2}+\frac{1}{2}\frac{\kappa+1}{\kappa+2}\alpha\theta-\frac{1}{3}\right)2p_{32}w32\right]+\\
&+q^2\left[\left(\frac{\theta}{3}-\frac{1}{3}\right)p_{40}w40+\left(\frac{\theta}{3}-\frac{1}{3}\right)p_{41}w_{41}+\left(\frac{\theta}{3}-\frac{1}{3}\right)2p_{41}w_{41}\right]\\
&\geqslant(1-q)^2\frac{1}{2\kappa+1}\theta+2q(1-q)\frac{1}{\kappa+2}\theta+q^2\frac{\theta}{3}-\frac{1}{3}
\end{aligned}
$$

Claim 1: There is no war between low type players ($w_{1j}=0$)

All wars that occur when there are three low type players reduce the arbitration payoff for a low type player because he gains $\theta/3<1/3$, regardless of the alliance structure. Thus the arbitrator never uses these wars.

Claim 2: There exists a function $\overline{q}(\theta,\kappa)$ such that whenever $q<\bar{q}(\theta,\kappa)$ the high player's ex-ante expected war payoff is larger than 1/3, and as a consequence either $w_{20}>0$ or $w_{21}>0$.

The high type player *ex-ante* expected war payoff is given by:

$$(1-q)^2\frac{1}{2\kappa+1}\theta+2q(1-q)\frac{1}{\kappa+2}\theta+q^2\frac{\theta}{3}$$

Through some algebra we can show it is larger than 1/3 whenever

$$q<\overline{q}(\theta,\kappa)=\frac{3\theta-\sqrt{9\theta^2-2\theta(2+\kappa)(3\theta-2\kappa+1)}}{2(1-\kappa)\theta} \tag{A.6}$$

(see Figure 4, for a simulation considering $\kappa=0$). If q is indeed smaller than $\overline{q}(\theta,\kappa)$, then the arbitrator must encourage the high type to participate in the arbitration by warning the high type whenever there is a war with a higher payoff to be earned. Because only wars against low types are profitable and the $\alpha<1/(1+\kappa)$ assumption make the alliance with a low type a bad deal for the high type, the arbitrator can only use the alliance structures where the high type is left alone, 20 and 21.

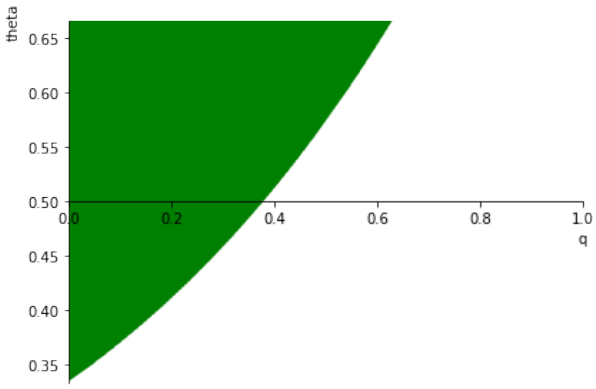
Claim 3: If q is low enough, LIC does not bind

More specifically, if q is such that

$$2q(1-q)\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)-(1-q)^2\left(\frac{\theta}{3}-\frac{1}{3}\right)\geqslant 0$$

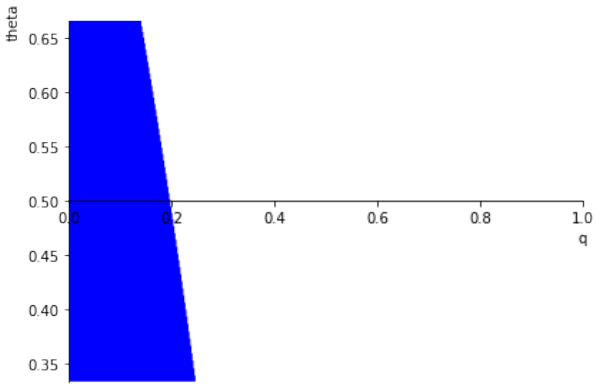
then the coefficients associated with $p_{20}w_{20}$ and $p_{21}w_{21}$ in LIC are positive and so, because of the previous claim, the fact that either $w_{20}>0$ or $w_{21}>0$ in order to fulfill HPC implies that LIC is

Figure 4 – Expected war payoff of H is $> 1/3$



automatically respected and so the arbitrator can set the other war probabilities to zero (See this region, for $\kappa = 0$, on Figure 5).

Figure 5 – Low type IC does not bind



As the alliance structure 21 (two L types, one H, the L types are allied) yields the same war payoff as the empty alliance structure, the overall war probability is the same as in the benchmark case.

Intuitively, what happens is that, in this region of the parameters, to have your adversaries to be of low type is far more likely than to have at least one high type. So, if a low player lies, chances are that the other reported types will be two other low types. As the arbitrator must incite an war between two low types and a single high type to ensure high's participation, the untruthful low player has a higher chance of entering a conflict where all players are of L type, which is bad for him.

On the opposite case, q is large enough so that the event where there is at least one high type becomes more likely, under this circumstance, a low type may prefer to lie in order to become less susceptible to conflict (up to this point, the arbitrator does not wish to promote wars where there are two high types because the outcome of this war decreases the participation payoff for the high type)

Claim 4: if q is large enough, that is,

$$2q(1-q)\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)-(1-q)^2\left(\frac{\theta}{3}-\frac{1}{3}\right)<0$$

then the arbitrator incites war between 2 H types and a single low type, with empty alliance structure

Under the condition stated, the coefficient associated with either $p_{20}w_{20}$ or $p_{21}w_{21}$ in LIC is smaller than zero and so to ensure LIC is respected the arbitrator must increase the probability of occurrence of conflict when there are two reported high types and a single low type. That is because q is high enough so that the event of at least one player being of high type becomes more likely than every player being of low type. As a low type player knows the arbitrator must increase the probability of conflict in this scenario, then the low type player has an incentive to lie. In order to dampen this incentive, the arbitrator then increases the probability of conflict if he receives the message vector with tow high types and a single low type.

Now, the arbitrator must decide which alliance structure to use. This is not a trivial decision because there are two forces at play.

From the low player's perspective, it would be better if the arbitrator used the alliance structure 32, where the low type player is allied to one of the other two high type players. The reason is that if the low type tells the truth, she would be allocated to a war in which he earns

$$\frac{\kappa+1}{2+\kappa}(1-\alpha)\theta$$

that, because of the assumption that $\alpha < 1/(1+\kappa)$, is an amount smaller than if she was in an alliance structure without any allies. If she lies, then with probability $p_{32}w_{32}$ she is the "high" type player allied to the real low type player and they go to war with the high type player. They are more likely to loose. Again, with probability $p_{32}w_{32}$, the untruthful low type ends up alone – the alliance is between the remaining two players– and as one of the adversaries is of high type, the untruthful player has a high chance of loosing.

The alliance 31 (two reported H types allied, reported L alone) will unambiguously never be used. The reason being this alliance increases the incentive for the low type to lie and, as a reported high type, become allied with another high type and get a smaller loss in case of war; and at the same time, the high type is indifferent between the conflict generated via 31 or 30.

From the high type player's perspective, the alliance 32 reduces the participation payoff because would be fighting a war against a high type player and having to share the war spoils with the low type player according to an unfavourable sharing rule.

So, in order to prove which of the alliances induces the least amount of war, I proceed according to the following steps. First, consider the case where LIC and HPC hold with equality. Let $\tilde{x}_{30} :=$

$\{(p_{30}w_{30})/(3q^2(1-q))$ and $\tilde{x}_{32} := (2p_{32}w_{32})/(3q^2(1-q))$[1]. Suppose, by way of contradiction, $\tilde{x}_32 > 0$ and assume $\tilde{x}_{30} \in [0,1)$. We will show that a reduction in $\tilde{x}_{32}$, accompanied by an increase in $\tilde{x}_30$ relaxes the problem. Assume that $\tilde{x}_{30}$, $\tilde{x}_{32}$ are initially such that LIC binds. We will disturb $\tilde{x}_{30}$ by an increase of $\epsilon > 0$ and $\tilde{x}_{32}$ by a decrease of $\delta > 0$ so that LIC continues to bind. Because LIC still binds we must have the following equality:

$$\left[\left(\frac{\kappa\theta}{\kappa+2}-\frac{1}{3}\right)\frac{1}{3(1-q)}-\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)\frac{2}{3q}\right]\epsilon = \left[\left(\frac{(\kappa+1)(1-\alpha)\theta}{\kappa+2}-\frac{1}{3}\right)\frac{1}{3(1-q)}-\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)\frac{2}{3q}\right]\delta$$

or,

$$\epsilon = \frac{\left[\left(\frac{(\kappa+1)(1-\alpha)\theta}{\kappa+2}-\frac{1}{3}\right)\frac{1}{3(1-q)}-\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)\frac{2}{3q}\right]}{\left[\left(\frac{\kappa\theta}{\kappa+2}-\frac{1}{3}\right)\frac{1}{3(1-q)}-\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)\frac{2}{3q}\right]}\delta \tag{A.7}$$

Because of the disturbances, the lhs of HPC is disturbed by:

$$\begin{aligned}
&\frac{2}{3q}\left(\frac{\theta}{\kappa+2}-\frac{1}{3}\right)\epsilon+\frac{2}{3q}\left(\frac{1}{2}\frac{\theta}{\kappa+2}+\frac{1}{2}\frac{\kappa+1}{\kappa+2}\alpha\theta-\frac{1}{3}\right)(-\delta) = \\
&\frac{2}{3q}\left(\frac{\theta}{\kappa+2}-\frac{1}{3}\right)\frac{\left[\left(\frac{(\kappa+1)(1-\alpha)\theta}{\kappa+2}-\frac{1}{3}\right)\frac{1}{3(1-q)}-\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)\frac{2}{3q}\right]}{\left[\left(\frac{\kappa\theta}{\kappa+2}-\frac{1}{3}\right)\frac{1}{3(1-q)}-\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)\frac{2}{3q}\right]}\delta+\frac{2}{3q}\left(\frac{1}{2}\frac{\theta}{\kappa+2}+\frac{1}{2}\frac{\kappa+1}{\kappa+2}\alpha\theta-\frac{1}{3}\right)(-\delta) = \\
&\left[\frac{2}{3q}\left(\frac{\theta}{\kappa+2}-\frac{1}{3}\right)\frac{\left[\left(\frac{(\kappa+1)(1-\alpha)\theta}{\kappa+2}-\frac{1}{3}\right)\frac{1}{3(1-q)}-\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)\frac{2}{3q}\right]}{\left[\left(\frac{\kappa\theta}{\kappa+2}-\frac{1}{3}\right)\frac{1}{3(1-q)}-\left(\frac{\kappa\theta}{2\kappa+1}-\frac{1}{3}\right)\frac{2}{3q}\right]}-\frac{2}{3q}\left(\frac{1}{2}\frac{\theta}{\kappa+2}+\frac{1}{2}\frac{\kappa+1}{\kappa+2}\alpha\theta-\frac{1}{3}\right)\right]\delta
\end{aligned} \tag{A.8}$$

Thus, in order to prove our claim, we need to show that the coefficient of δ in A.8 is strictly positive. Through some algebra, and by considering that q does not assume any extreme value (i.e, $q \neq 0, 1$), we can reduce such coefficient to the following expression:

$$\frac{\theta\left(\alpha(1+\kappa)-1\right)\left(-6\kappa\theta+16\kappa+q\left(12\kappa\theta-22\kappa+3\theta-11\right)+8\right)}{27q^2\left(1-q\right)\left(\kappa+2\right)\left(2\kappa+1\right)} \tag{A.9}$$

In order to prove our claim, we only need to show that expression A.9 is positive. Note that since $\kappa \in [0,1]$, $q \in (0,1)$ and $\theta > 1/3$, the only two factors that determine the signal of the expression are $(\alpha(1+\kappa)-1)$ and $(-6\kappa\theta+16\kappa+q\left(12\kappa\theta-22\kappa+3\theta-11\right)+8)$. Because of our initial assumption that $\alpha < 1/(1+\kappa)$, it is trivial to see that $\alpha(1+\kappa)-1 < 0$, and so, for expression A.9 to be positive, we need that $(-6\kappa\theta+16\kappa+q\left(12\kappa\theta-22\kappa+3\theta-11\right)+8) < 0$. This is true whenever

$$q < \hat{q}(\theta,\kappa) = \frac{2\left(3\kappa\theta-8\kappa-4\right)}{12\kappa\theta-22\kappa+3\theta-11} \tag{A.10}$$

[1] These adjustments are made only to normalize each choice variable with their weight on the objective function

Again, through some algebra, it is possible to show that $\bar{q} \leqslant \hat{q}$, and as a result, whenever the arbitrator must use the $3j$ alliances to impede the misreport by a low type, he will, indeed use the 30 alliance structure.

Claim 5: The arbitrator never uses alliances whenever he receives messages of 3 high types

The justification for this claim follow the same idea from above. First note that a high player is indifferent between the conflicts occurring when his opponents are of high type. in every case, he gains, in expectation, $\theta/3$. Now, a low type has an increased incentive to lie if the arbitrator uses alliances when all players report to be of high type, because there would be a chance the untruthful player would end up allied to a high type and thus he would suffer a smaller loss in case of war, as he would share the spoils with a high type.

In sum, since the arbitrator cannot do better in the reduced program with alliances in relation to the relaxed problem without alliances, we can conclude that, in our scenario, he cannot reduce war probability any further than in the benchmark case. □

APPENDIX B – The General Problem with Alliances

The general form of the objective function of the arbitration problem with alliances is given by:

$$\begin{aligned} W(\mathbf{p},\mathbf{w}) = (1-q)^3 \left(w_{10}p_{10} + 3w_{11}p_{11}\right) + \\ +3q\,(1-q)^2 \left(w_{20}p_{20} + w_{21}p_{21} + 2w_{22}p_{22}\right) + \\ +3q^2\,(1-q)\left(w_{30}p_{30} + w_{31}p_{31} + 2w_{32}p_{32}\right) + \\ +q^3\left(w_{40}p_{40} + 3w_{41}p_{41}\right) \end{aligned} \tag{B.1}$$

The restrictions faced by the arbitrator are:

- Low participation constraint (LPC):

$$\begin{aligned} &(1-q)^2\left[\left(\frac{\theta}{3}-\frac{1}{3}\right)w_{10}p_{10} + \left(\frac{\theta}{2^\mu+1}-\frac{1}{3}\right)w_{11}p_{11} + \left(\frac{2^\mu}{2^\mu+1}\frac{\theta}{2}-\frac{1}{3}\right)2w_{11}p_{11}\right] + \\ &+2q(1-q)\left[\left(\frac{\kappa^\mu\theta}{2\kappa^\mu+1}-\frac{1}{3}\right)w_{20}p_{20} + \left(\frac{(2\kappa)^\mu\theta}{(2\kappa)^\mu+1}-\frac{1}{3}\right)w_{21}p_{21} + \right. \\ &\left.+\left(\frac{1}{2}\frac{(\kappa+1)^\mu}{\kappa^\mu+(\kappa+1)^\mu}(1-\alpha)\theta + \frac{1}{2}\frac{\kappa^\mu}{\kappa^\mu+(\kappa+1)^\mu}\theta - \frac{1}{3}\right)2w_{22}p_{22}\right] + \\ &+q^2\left[\left(\frac{\kappa^\mu\theta}{\kappa^\mu+2}-\frac{1}{3}\right)w_{30}p_{30} + \left(\frac{\kappa^\mu\theta}{2^\mu+\kappa^\mu}-\frac{1}{3}\right)w_{31}p_{31} + \left(\frac{(\kappa+1)^\mu}{(\kappa+1)^\mu+1}(1-\alpha)\theta-\frac{1}{3}\right)2w_{32}p_{32}\right] \\ &\geqslant (1-q)^2\frac{\theta}{3} + 2q(1-q)\frac{\kappa^\mu}{2\kappa^\mu+1}\theta + q^2\frac{\kappa^\mu}{\kappa^\mu+2}\theta - \frac{1}{3} \end{aligned}$$

- Low incentive compatibility constraint (LIC):

$$\begin{aligned} &(1-q)^2\left[\left(\frac{\theta}{3}-\frac{1}{3}\right)w_{10}p_{10} + \left(\frac{\theta}{2^\mu+1}-\frac{1}{3}\right)w_{11}p_{11} + \left(\frac{2^\mu}{2^\mu+1}\frac{\theta}{2}-\frac{1}{3}\right)2w_{11}p_{11}\right] + \\ &+2q(1-q)\left[\left(\frac{\kappa^\mu\theta}{2\kappa^\mu+1}-\frac{1}{3}\right)w_{20}p_{20} + \left(\frac{(2\kappa)^\mu\theta}{(2\kappa)^\mu+1}-\frac{1}{3}\right)w_{21}p_{21} + \right. \\ &\left.+\left(\frac{1}{2}\frac{(\kappa+1)^\mu}{\kappa^\mu+(\kappa+1)^\mu}(1-\alpha)\theta + \frac{1}{2}\frac{\kappa^\mu}{\kappa^\mu+(\kappa+1)^\mu}\theta - \frac{1}{3}\right)2w_{22}p_{22}\right] + \\ &+q^2\left[\left(\frac{\kappa^\mu\theta}{\kappa^\mu+2}-\frac{1}{3}\right)w_{30}p_{30} + \left(\frac{\kappa^\mu\theta}{2^\mu+\kappa^\mu}-\frac{1}{3}\right)w_{31}p_{31} + \left(\frac{(\kappa+1)^\mu}{(\kappa+1)^\mu+1}(1-\alpha)\theta-\frac{1}{3}\right)2w_{32}p_{32}\right] \\ &\geqslant (1-q)^2\left[\left(\frac{\theta}{3}-\frac{1}{3}\right)w_{20}p_{20} + \left(\frac{\theta}{2^\mu+1}-\frac{1}{3}\right)w_{21}p_{21} + \left(\frac{2^\mu}{2^\mu+1}\frac{\theta}{2}-\frac{1}{3}\right)2w_{22}p_{22}\right] + \\ &+2q(1-q)\left[\left(\frac{\kappa^\mu\theta}{2\kappa^\mu+1}-\frac{1}{3}\right)w_{30}p_{30} + \left(\frac{(\kappa+1)^\mu}{\kappa^\mu+(\kappa+1)^\mu}(1-\alpha)\theta-\frac{1}{3}\right)w_{31}p_{31} + \right. \\ &\left.+\left(\frac{1}{2}\frac{(2\kappa)^\mu\theta}{(2\kappa)^\mu+1} + \frac{1}{2}\frac{\kappa^\mu}{\kappa^\mu+(\kappa+1)^\mu}\theta - \frac{1}{3}\right)2w_{32}p_{32}\right] + \\ &+q^2\left[\left(\frac{\kappa^\mu\theta}{\kappa^\mu+2}-\frac{1}{3}\right)w_{40}p_{40} + \left(\frac{\kappa^\mu\theta}{2^\mu+\kappa^\mu}-\frac{1}{3}\right)w_{41}p_{41} + \left(\frac{(\kappa+1)^\mu}{(\kappa+1)^\mu+1}(1-\alpha)\theta-\frac{1}{3}\right)2w_{41}p_{41}\right] \end{aligned}$$

- High participation constraint (HPC):

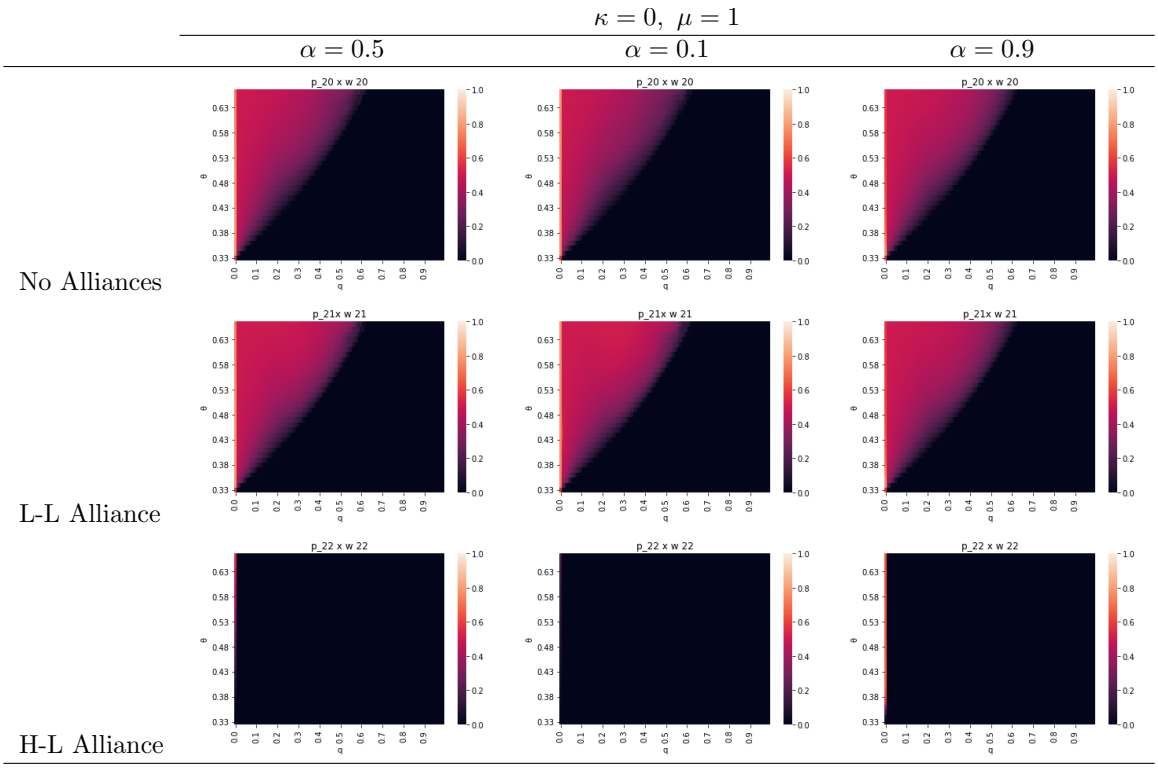
$$
\begin{aligned}
&(1-q)^2\left[\left(\frac{\theta}{2\kappa^\mu+1}-\frac{1}{3}\right)p_{20}w_{20}+\left(\frac{\theta}{(2\kappa)^\mu+1}-\frac{1}{3}\right)p_{21}w21+\left(\frac{\theta\,(\kappa+1)^\mu}{\kappa^\mu+(\kappa+1)^\mu}\alpha-\frac{1}{3}\right)2p_{22}w22\right]+\\
&+2q(1-q)\left[\left(\frac{\theta}{\kappa^\mu+2}-\frac{1}{3}\right)p_{30}w_{30}+\left(\frac{2^\mu\theta}{2\,(2^\mu+\kappa^\mu)}-\frac{1}{3}\right)p_{31}w_{31}+\right.\\
&\left.+\left(\frac{\theta}{(\kappa+1)^\mu+1}-\frac{1}{3}\right)p_{32}w_{32}+\left(\frac{(\kappa+1)^\mu}{(\kappa+1)^\mu+1}\alpha\theta-\frac{1}{3}\right)p_{32}w_{32}\right]+\\
&+q^2\left[\left(\frac{\theta}{3}-\frac{1}{3}\right)p_{40}w_{40}+\left(\frac{\theta}{2^\mu+1}-\frac{1}{3}\right)p_{41}w_{41}+\left(\frac{2^\mu\theta}{2\,(2^\mu+1)}-\frac{1}{3}\right)2p_{41}w_{41}\right]\\
&\geqslant(1-q)^2\frac{1}{2\kappa^\mu+1}\theta+2q(1-q)\frac{1}{\kappa^\mu+2}\theta+q^2\frac{\theta}{3}-\frac{1}{3}
\end{aligned}
$$

- High Incentive compatibility constraint (HIC):

$$
\begin{aligned}
&(1-q)^2\left[\left(\frac{\theta}{2\kappa^\mu+1}-\frac{1}{3}\right)p_{20}w_{20}+\left(\frac{\theta}{(2\kappa)^\mu+1}-\frac{1}{3}\right)p_{21}w_{21}+\left(\frac{\theta\,(\kappa+1)^\mu}{\kappa^\mu+(\kappa+1)^\mu}\alpha-\frac{1}{3}\right)2p_{22}w_{22}\right]+\\
&+2q(1-q)\left[\left(\frac{\theta}{\kappa^\mu+2}-\frac{1}{3}\right)p_{30}w30+\left(\frac{2^\mu\theta}{2\,(2^\mu+\kappa^\mu)}-\frac{1}{3}\right)p_{31}w_{31}+\right.\\
&\left.+\left(\frac{\theta}{(\kappa+1)^\mu+1}-\frac{1}{3}\right)p_{32}w_{32}+\left(\frac{(\kappa+1)^\mu}{(\kappa+1)^\mu+1}\alpha\theta-\frac{1}{3}\right)p_{32}w_{32}\right]+\\
&+q^2\left[\left(\frac{\theta}{3}-\frac{1}{3}\right)p_{40}w_{40}+\left(\frac{\theta}{2^\mu+1}-\frac{1}{3}\right)p_{41}w_{41}+\left(\frac{2^\mu\theta}{2\,(2^\mu+1)}-\frac{1}{3}\right)2p_{41}w_{41}\right]\\
&\geqslant(1-q)^2\left[\left(\frac{\theta}{2\kappa^\mu+1}-\frac{1}{3}\right)p_{10}w_{10}+\left(\frac{\theta}{(2\kappa)^\mu+1}-\frac{1}{3}\right)p_{11}w_{11}+\left(\frac{\theta\,(\kappa+1)^\mu}{\kappa^\mu+(\kappa+1)^\mu}\alpha-\frac{1}{3}\right)2p_{11}w_{11}\right]+\\
&+2q(1-q)\left[\left(\frac{\theta}{\kappa^\mu+2}-\frac{1}{3}\right)p_{20}w_{20}+\left(\frac{(\kappa+1)^\mu}{(\kappa+1)^\mu+1}\alpha\theta-\frac{1}{3}\right)p_{21}w_{21}+\right.\\
&\left.+\left(\frac{\theta}{(\kappa+1)^\mu+1}-\frac{1}{3}\right)p_{22}w_{22}+\left(\frac{2^\mu\theta}{2\,(2^\mu+\kappa^\mu)}-\frac{1}{3}\right)p_{22}w_{22}\right]+\\
&+q^2\left[\left(\frac{\theta}{3}-\frac{1}{3}\right)p_{30}w_{30}+\left(\frac{\theta}{2^\mu+1}-\frac{1}{3}\right)p_{31}w_{31}+\left(\frac{2^\mu\theta}{2\,(2^\mu+1)}-\frac{1}{3}\right)2p_{32}w_{32}\right]
\end{aligned}
$$

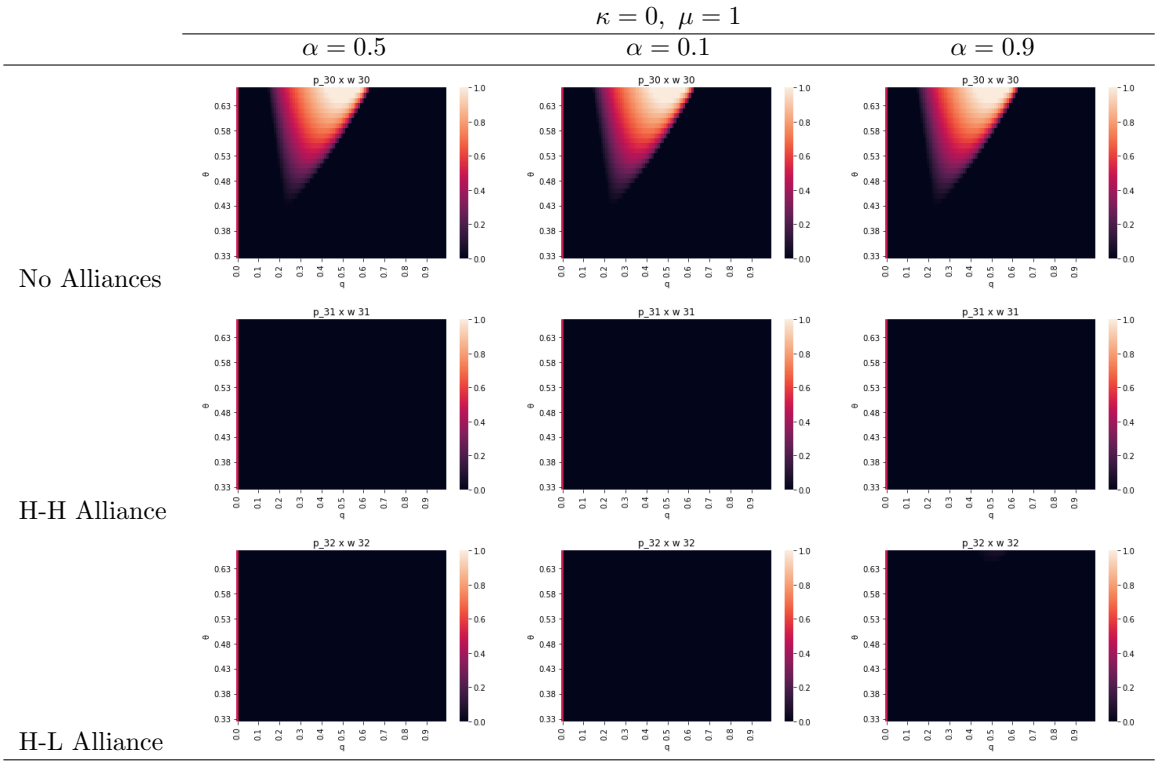
APPENDIX C – Simulation Results – Figures

Figure 6 – Simulation results on α for report with two L and a single H



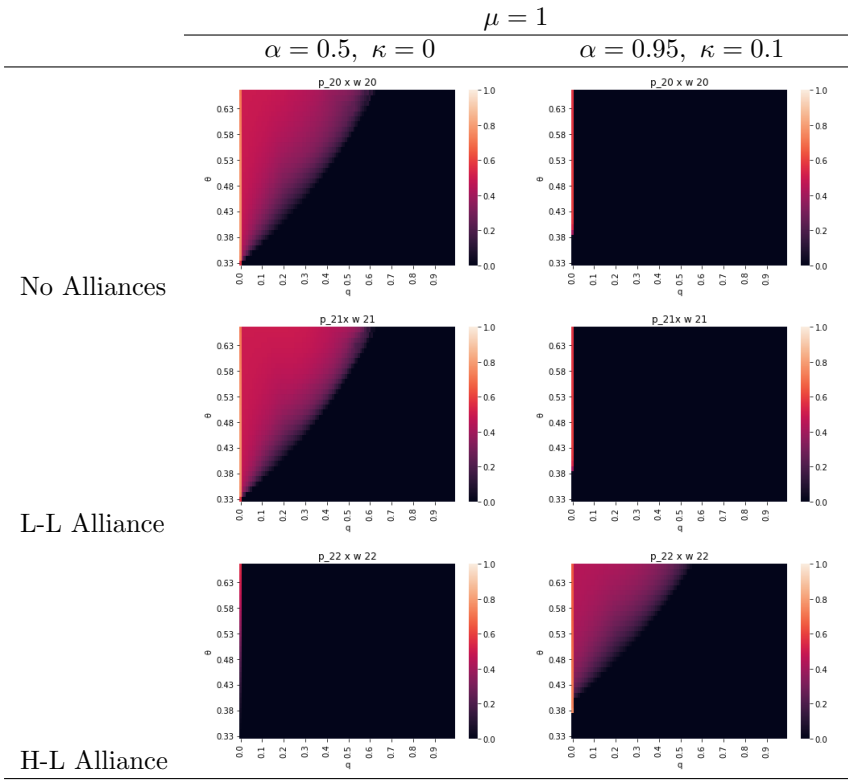
Notes: Each graph corresponds is the simulation result of the linear programming arbitrator problem for a given choice variable, the product $p_{2j}w_{2j}$. in a $(q \times \theta)$ plane. Recall that the index 2 corresponds to the report containing two low types and a single high type, and the index j correspond to each possible alliance configuration – $j = 0$ means no alliances, $j = 1$, alliance between low types and $j = 2$, alliance between a low and a high type. Each row represents the results for each alliance structure. Each colmun represents a different value for the parameter α. Throughout the exercise we keep $\kappa = 0$ and $\mu = 1$. Ligther colors means higher values. By the nature of the problem each variable must assume values between 0 and 1.

Figure 7 – Simulation results on α for report with two H and a single L



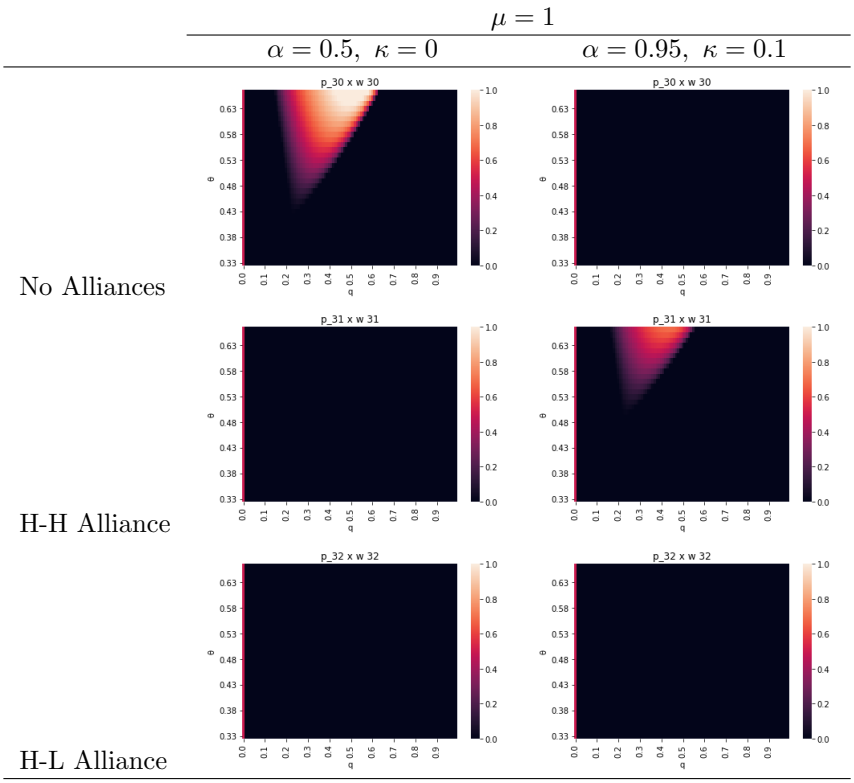
Notes: Each graph corresponds is the simulation result of the linear programming arbitrator problem for a given choice variable, the product $p_{3j}w_{3j}$. in a $(q \times \theta)$ plane. Recall that the index 3 corresponds to the report containing two high types and a single low type, and the index j correspond to each possible alliance configuration – $j = 0$ means no alliances, $j = 1$, alliance between low types and $j = 2$, alliance between a low and a high type. Each row represents the results for each alliance structure. Each colmun represents a different value for the parameter α. Throughout the exercise we keep $\kappa = 0$ and $\mu = 1$. Ligther colors means higher values. By the nature of the problem each variable must assume values between 0 and 1.

Figure 8 – Simulation results on α and κ for report with two H and a single L



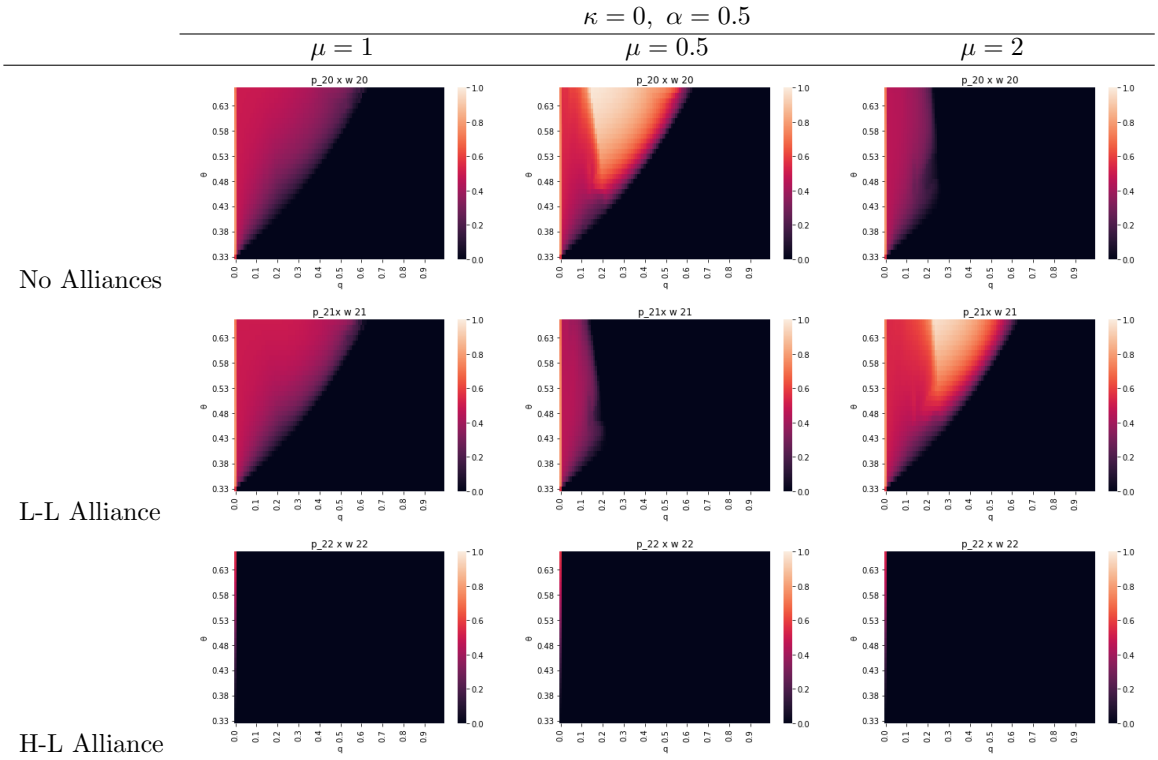
Notes: Each graph corresponds is the simulation result of the linear programming arbitrator problem for a given choice variable, the product $p_{2j}w_{2j}$. in a $(q \times \theta)$ plane. Recall that the index 2 corresponds to the report containing two low types and a single high type, and the index j correspond to each possible alliance configuration – $j = 0$ means no alliances, $j = 1$, alliance between low types and $j = 2$, alliance between a low and a high type. Each row represents the results for each alliance structure. Each colmun represents a different value for the parameter α. Throughout the exercise we keep $\kappa = 0$ and $\mu = 1$. Ligther colors means higher values. By the nature of the problem each variable must assume values between 0 and 1.

Figure 9 – Simulation results on α and κ for report with two H and a single L



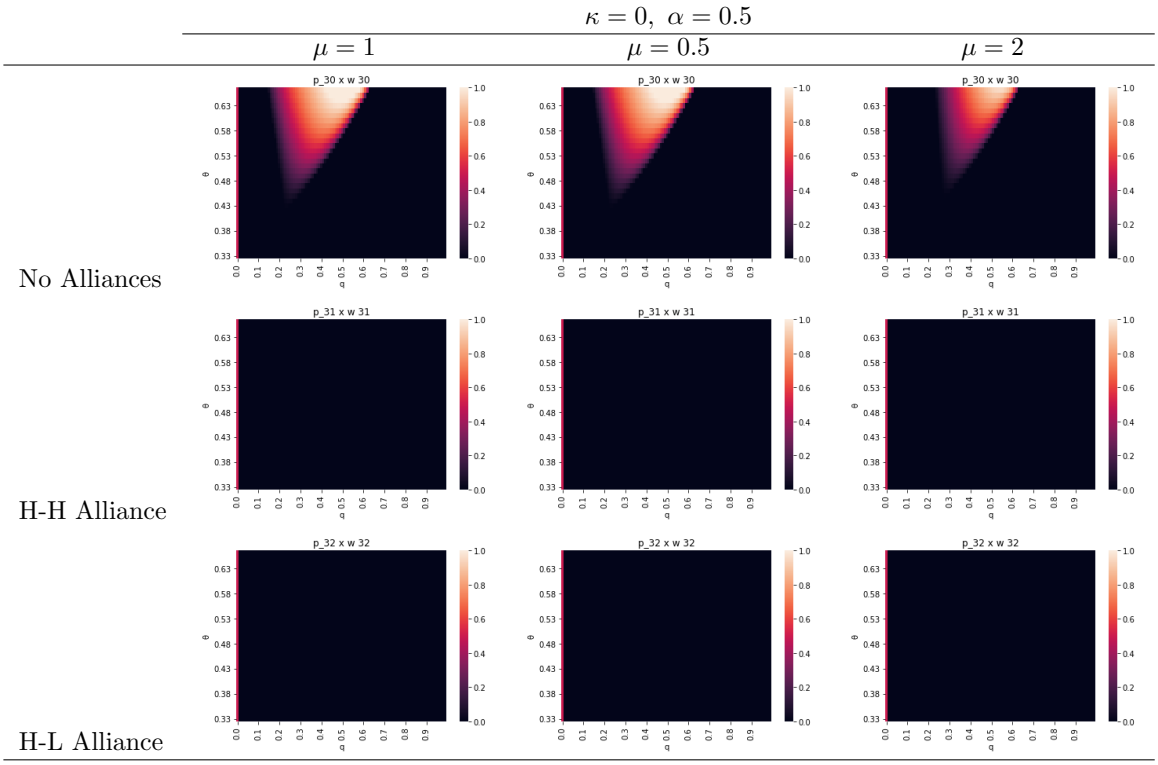
Notes: Each graph corresponds is the simulation result of the linear programming arbitrator problem for a given choice variable, the product $p_{3j}w_{3j}$. in a $(q \times \theta)$ plane. Recall that the index 3 corresponds to the report containing two high types and a single low type, and the index j correspond to each possible alliance configuration – $j = 0$ means no alliances, $j = 1$, alliance between low types and $j = 2$, alliance between a low and a high type. Each row represents the results for each alliance structure. Each colmun represents a different value for the parameter α. Throughout the exercise we keep $\kappa = 0$ and $\mu = 1$. Ligther colors means higher values. By the nature of the problem each variable must assume values between 0 and 1.

Figure 10 – Simulation results on μ for report with two L and a single H



Notes: Each graph corresponds is the simulation result of the linear programming arbitrator problem for a given choice variable, the product $p_{2j}w_{2j}$. in a $(q \times \theta)$ plane. Recall that the index 2 corresponds to the report containing two low types and a single high type, and the index j correspond to each possible alliance configuration – $j = 0$ means no alliances, $j = 1$, alliance between low types and $j = 2$, alliance between a low and a high type. Each row represents the results for each alliance structure. Each colmun represents a different value for the parameter α. Throughout the exercise we keep $\kappa = 0$ and $\mu = 1$. Ligther colors means higher values. By the nature of the problem each variable must assume values between 0 and 1.

Figure 11 – Simulation results on μ for report with two H and a single L



Notes: Each graph corresponds is the simulation result of the linear programming arbitrator problem for a given choice variable, the product $p_{3j}w_{3j}$. in a $(q \times \theta)$ plane. Recall that the index 3 corresponds to the report containing two high types and a single low type, and the index j correspond to each possible alliance configuration – $j = 0$ means no alliances, $j = 1$, alliance between low types and $j = 2$, alliance between a low and a high type. Each row represents the results for each alliance structure. Each colmun represents a different value for the parameter α. Throughout the exercise we keep $\kappa = 0$ and $\mu = 1$. Ligther colors means higher values. By the nature of the problem each variable must assume values between 0 and 1.