Essays in the Experimental Analysis of Conflict

by

Humera Khan

A thesis submitted to the University of Birmingham for the degree of
DOCTOR OF PHILOSOPHY

Department of Economics
Birmingham Business School
College of Social Sciences
University of Birmingham
January 2017

Abstract

The thesis consists of three chapters with Chapters 2 and 3 providing experimental evidence on the role of cheap-talk and a third party recommendation in reducing or aggravating conflict.

Chapter 1 surveys the theoretical, empirical and experimental literature on the determinants of conflict.

Chapter 2 considers an experiment based on Baliga and Sjostrom (BS, 2004) to investigate whether communication reduces the probability of an arms race. We find that communication does indeed reduce the possibility of using strategies that lead to an arms race, even when the unique Bayesian Nash equilibrium without communication has both players playing a strategy that leads to an arms race.

Chapter 3 considers a set of experiments based on Baliga and Sjostrom (BS, 2012) to understand if third parties can provoke conflict. We adapt their model to experimentally test if a third party recommendation can trigger conflict. While in some treatments with recommendation, more players do choose an aggressive strategy compared to the treatment without, none of them are statistically significant. We propose a number of explanations for why provocation may not necessarily increase conflict in this environment.

This thesis is dedicated to my grandfather, Mohammed Aqeel Siddique.

Acknowledgements

I would like to acknowledge my main supervisor, Siddhartha Bandyopadhyay, for his support throughout my research. For his encouragement and kindness, my gratitude extends beyond the completion of these essays. I thank Michalis Drouvelis, for his patience and direction, particularly for the organization of the experiments. I am also thankful to Bibhas Saha and Anindya Banerjee, the examiners of this thesis, for their valuable comments and suggestions.

I am appreciative of Daniel Derbyshire's advice, assurance, friendship and for never making any request feel a bother. More friends I am indebted to, who have been there from the very beginning, are Joseff Eynon-Freeman, Joseph Beeby, Patrick Surlis and Steven Savidge. They have always been quick to remind me of the bigger picture, and to try and enjoy the process, as much as possible.

Lastly, I would like to thank my family. I thank my mother for her belief in me, and my father for his infinite positivity. I thank my paternal grandparents for their love and words of motivation. I thank my maternal grandparents for saying they were proud of me, before I had even started.

Table of Contents

List of Figures

List of Tables

Summary of Chapters

The thesis consists of three chapters with Chapters 2 and 3 providing experimental evidence on the role of cheap-talk and a third party recommendation in reducing or aggravating conflict.

Chapter 1 surveys the theoretical, empirical and experimental literature on the determinants of conflict.

Chapter 2 considers an experiment based on Baliga and Sjostrom (BS, 2004). They build a model of an arms race where the players' propensity to arm is private information and find that if a 'Multiplier Condition' holds, there is a unique Bayesian Nash equilibrium involving an arms race with probability one. However, they show that there exists an equilibrium in a cheap-talk extension where the probability of an arms race is close to zero. We adapt the Baliga Sjostrom framework to conduct an experiment to investigate whether when the 'Multiplier Condition' holds, cheap-talk reduces the probability of an arms race. We find that communication does indeed reduce the possibility of using strategies that lead to an arms race, when players can engage in pre-play communication, even when there is a unique Bayesian Nash equilibrium without communication where both players play a strategy that leads to an arms race.

Chapter 3 considers a set of experiments based on Baliga and Sjostrom (BS, 2012) to understand if third parties can provoke conflict. BS analyse a stag hunt game of incomplete information and find that when actions are strategic complements, there exists an equilibrium when a third party message can increase the probability of conflict by triggering some players to switch from a non-aggressive to an aggressive strategy. We adapt a discrete version of the BS model to experimentally test if a third party recommendation can trigger conflict. While in some treatments with recommendation, more players do choose an aggressive strategy compared to the treatment without, none of them are statistically significant. This may occur for a variety of reasons: players may simply be playing a 'babbling' equilibrium and ignoring recommendations or the payoff pair from the non-aggressive strategy may look like a 'safe strategy' with implicit co-ordination between players, they may be inequality averse or they might be playing the socially optimal outcome.

Introduction

History provides many examples where fear and distrust, sparked by uncertainty about the opponent's motives, appears to have triggered arms races. In the Peloponnesian War, Sparta and its allies interpreted a treaty of mutual protection between Athens and Corcyra as an act of provocation. Although the Athenians and Spartans had been opposed to one another for many years prior, the treaty caused Sparta to cancel its peace treaty with Athens and the war began in 431BC when Sparta's ally, Thebes, tried to seize Potidea. A lack of trust in Germany, caused Britain to join the war against them in World War I. Belgium ports were close to the British coast and the German control was a threat. The Cold War was based on a situation of distrust between the Soviet Union and the Allied Powers, just as the India-Pakistan arms race is between India and Pakistan. The growing literature on what provokes conflict or causes an escalation is mixed. This thesis contributes to the understanding of conflict and the role of communication in escalating and deescalating aggression between players.

Experimental methods within laboratories are applied for three purposes (Garfinkel and Skaperdas, 2012). Firstly, they can check theoretical models. They permit us to observe behaviour and enable us to capture the psychology underpinning theory more successfully. Secondly, they can substitute field data that is often unavailable, particularly when studying a topic such as conflict. Lastly, even in situations where some field data is available, laboratory data can be gathered in tandem with field data. This is strongly complementary as one method's weakness is compensated by the other's strengths. Field data is gathered in real life and is therefore more realistic. However, it suffers from noise, identification problems and a lack of control. On the other hand, laboratory data allows the use of a controlled environment so we can test variations of individual factors, keeping all others constant - endogeneity problems do not arise. We conduct analysis in a controlled setting such as this to isolate the role of fear in the propagation of conflict.

The best possible scenario would be to have a combination of theory, field data and laboratory data, all in support of each other, and we hope to contribute in this way.

The role of private information is currently a fundamental issue. There are studies such as Levy (2004) and Ottaviani and Sorenson (2006) where there exists asymmetric information on key parameters, in order to gain a competitive advantage against their opponent. Our research in Chapter Two contributes to this area. There is also a growing interest in the mechanisms of third-party mediators (Eisenkopf, 2009), which we contribute towards in Chapter Three. There is much work within the experimental literature on communication and the treatments differ in various ways, for example; timing, who sends the messages, who observes the messages, what type of messages can be sent. Both chapters can provide more insight into this too as, using experiments, we look for evidence on the role of communication in a class of games with imperfect information. In particular, we look at whether mutual aggression can be reduced by bi-lateral communication and whether mutual aggression can be increased by third party provocation.

In the second chapter of this thesis, we show that communication between players helps to avoid an adverse equilibrium by creating credible information about the players' types. This research helps us to understand the effectiveness of an 'arms control treaty' because in the absence of a 'World Court', any agreement would be pure cheap-talk. If all types send the same message, talk is not informative and cannot prevent an arms race spiral. However, we show that information can actually be generated, which could be used to reduce the likelihood of an arm's race. It is an experiment based on Baliga and Sjostrom (2004) and the design consists of a one-shot stag hunt game with four treatment conditions: Treatment 1 serves as a baseline, Treatment 2 gives the pure 'Multiplier Condition' effect, Treatment 3 gives the pure communication effect and Treatment 4 shows the full interaction. We have two hypotheses. Hypothesis 1 is that when the condition holds, we expect the frequency of the aggressive strategy to increase. Hypothesis 2 is that when the condition holds, we expect communication to reduce the frequency of the aggressive strategy.

In the third chapter, we investigate how a third party can manipulate conflict and attain an adverse equilibrium by acting as a provocateur. Their talk would not be informative if they always sent an aggressive message but by exploiting the behaviour of the dominant strategy types, they make a recommendation that would cause just one player to play aggressively, maximizing the Provocateur's payoffs. The study of provocation contributes to tackling extremism. An improved understanding of extremist behaviour will help understanding how to approach extremists in situations such as after terrorist attacks, during negotiations and before elections. This experiment is based on Baliga and Sjostrom (2012) and the design consisted of a one-shot stag hunt game with two treatment conditions: Treatment 1 serves as a baseline and Treatment 2 gives the pure recommendation effect. Again, we had two hypotheses. Hypothesis 1 is that we expect a recommendation to increase the frequency of the aggressive strategy. Hypothesis 2 is that we expect a recommendation to increase the frequency of the aggressive strategy played by the "weak moderate" type, Type C.

This thesis has three key findings. Firstly, without communication, players play mutually aggressive strategies; both when that is the unique equilibrium and also when there are other equilibria. Secondly, costless bi-lateral communication reduces play of mutually aggressive strategies. Lastly, in a Strategic Complements game, provocation by a third party does not significantly affect the probability of playing aggressive strategies which led us to conduct a second set of experiments with modified payoffs. There could be several reasons to explain this finding. Players may be maximizing the sum of their payoffs (Englemann and Strobel, 2004), inequity averse (Fehr and Schmidt, 1999) or having (pro-)social preferences (Kahneman, Knetsch and Thaler, 1986). The experiment suggests that provocation may not always be successful; if the gains from not engaging in conflict are sufficient, the peaceful strategy may be stable enough to withstand provocation.

These findings support experiments of a similar nature, such as Charness (2000) who shows that without communication, efforts to achieve efficiency are unsuccessful, and Clark et al. (2001) who show how agreements to play a Nash equilibrium are fragile when players have a strict preference over their opponent's strategy choice and that informative communication does not always lead to the Pareto-dominant outcome. We elaborate on the external validation of this thesis within the Conclusion.

Chapter 1

Literature Review

1.1 Theoretical and Empirical Research

1.1.1 Determinants of Conflict

The causes of conflict are as varied as its nature. Its roots are multifaceted and have important historical contexts. Historians have identified many factors as potential instigators but they believe it is often a combination that can trigger conflict (Carecon, 2011). These factors include; colonial legacy, militaristic cultures, ethnicity, religion, unequal development, inequality, poverty, bad leadership and polity frailties.

The Marxian paradigm states that through an economic class struggle, economic inequality is the root cause of conflict (see e.g. Rummel, 1977). This, however, lacks robust empirical support and is unlikely to be the primary mechanism behind conflict. Inequality is managed in a variety of ways by societies, there are different kinds of inequality and the transmission mechanisms that enable a "peaceable" inequality to turn into a violent conflict are important factors to be considered (Cramer, 2003). Along with insufficient data, the heterogeneity of transmission mechanisms have made it even more difficult to conduct convincing empirical work, particularly inter-country comparison. One of many examples is an empirical study by Muller and Seligson (1987) who compare countries within the Central American region to refute the "land maldistribution hypothesis" and subsequently contradict the findings of Midlarsky and Roberts (1985). Midlarsky and Roberts' paper supported the hypothesis that in El Salvador and Nicaragua conflicts may be caused by rising land and economic inequality that is led by population increases and failed agricultural policies. Muller and Seligson, however, found Costa Rica and Panama were characterized by similar preconditions during the same period as a study on El Salvador and Nicaragua, but remained peaceful.

Blattman and Miguel (2010) is another example of a study that weakened the evidence for the role of income inequality as a sole primary causal factor. They look at the contest model, a seminal theory within conflict literature, where two opponents (a rebel group and a government) allocate their resources to either production or appropriation. The most robust prediction of contest models, with broad empirical support, is that the odds of winning increase with the relative effectiveness of that player's fighting technology. However, when developing the model from unitary actors to many citizens, problems arise with individual participation; soldiering rises only as the opportunity cost of fighting falls and, therefore, households require motivation to coordinate. In their paper, they discuss this issue with 'wealth effects'; a low-income population fights because appropriation is more valuable than production, and a rich-income population desire conflict because there is more for them to fight for. These opposing effects can cancel out in some models (Fearon, 2007) if state revenues are drawn entirely from income taxation and income may, thus, have no net effect in conflict unless there are non-linearities in the form of the utility function or revenue collection function.

Many economic analyses of conflict model conflict as a contest. The contest is a game in which players spend resources on arming to increase their probability of winning a conflict if one was to occur, otherwise improving their bargaining position or using the arms as a deterrent (Garfinkel and Skaperdas, 2012). The key ingredient to these models are the "technologies on conflict" that show how the likelihood of winning varies with different levels of arming. Garfinkel and Skaperdas (2012) discuss the different functional forms that have been used and review their axiomatic foundations or stochastic derivations. Appropriative conflicts occur between groups but also within groups. Munster (2007) studies simultaneous inter- and intra-group conflicts, where individuals in groups can utilize their resources for production, for appropriation in a contest between groups and for appropriation in a contest within their own group. Contractual incompleteness is an exogenous parameter and this leads to the absence of well-defined and enforced property rights, thereby enabling a trade-off between production and appropriation. The technology of conflict is also parameterized. The model shows a 'group cohesion effect' where if the contest between the groups becomes more decisive, the individuals spend fewer resources on intra-group conflict. There is a 'reversed group cohesion effect' if the contest within the individuals becomes less decisive, and they spend more resources on the inter-group contest. Munster gives conditions for when the increase in the number of groups leads to more productive and less appropriative activities. He also shows that there is an optimal size of the entire organization, which is determined by a trade off between increasing costs of appropriation and increasing returns to scale in production.

Following the skepticism with income as an important causal factor, new theories of economic motives emerged, trying to explain the instigation of conflict. Some discussed how it could be due to having a capital-intensive economy (Dal Bo and Dal Bo, 2004), others focused on the condition of a country's terms of trade shifting (Besley and Persson, 2008) or how easily their wealth could be expropriated (Bellows and Miguel, 2008).

Dal Bo and Dal Bo's paper (2004) followed from the theories of mitigating 'wealth effects' and looked at positive shocks on capital-intensive and labour-intensive sectors. Their model states that if appropriation is more labour intensive than production, then shocks to labour-intensive industries will reduce conflict whereas shocks to the capital-intensive industries will increase it. With a similar framework, Besley and Persson (2008) model terms of trade volatility, investigating how the impacts of an increase in import prices or an increase in export prices, cause a higher risk of conflict. They found this occurred by import prices suppressing the real wage and export prices boosting the size of the government revenue pie. The expropriation of wealth is studied by Bellows and Miguel (2008) by linking it with geographic patterns. The theory is supported by their empirical findings that significantly more armed clashes occurred in areas with greater diamond wealth. They studied the civil war in Sierra Leone from 1991-2002 using nationally representative household data on conflict experiences, postwar economic outcomes, local politics and collective action, to understand local collective action after a group had experienced violence. They state that not only did armed groups compete for control over the diamond mines but, as large-scale diamond smuggling was made easier by chaos, there was less incentive for groups to not prolong the conflict.

These explanations were all backed by economic motives; however, some researchers argued that conflict could only achieve expression with non-economic markers such as national origins, ethnicity and religion (Esteban and Ray, 2008). Ray and Esteban study ethnic divisions and discuss how to distinguish between the concepts of fractionalization and polarization when researching the causes of conflict. Fractionalization is widely indexed by the degree to which a society is split into distinct groups, however, polarization has been conceptualized by a variety of specifications – some cruder than others. Collier (2001) uses perhaps the crudest, dominance, to describe polarization. Societies are considered 'dominated' if the largest group contains between 45% and 90% of the population. Esteban and Ray (1994) use 'antagonisms' which are the result of an inter-play between the group-size of a group that identify with each other (they have similar characteristics/attributes) and the inter-group distance (their sense of alienation with respect to the other groups). Unless there are just two groups within a society, the contrast is large; fractionalization is maximal when each individual is different, whereas, polarization is maximal when there are two only groups of individuals (Ray and Esteban, 2008).

Literature on ethnic divisions, such as Esteban and Ray (2008), still relied on economic assumptions, some still vindicating income as an important factor. They showed how poverty could potentially inhibit a successful revolution while wealth may allow for the resources for engagement in successful conflict but the marginal gains for the wealthy are low. However, within an ethnic group, income differences may allow for specialisation with the rich providing the money and the poor their labour for conflict across ethnic groups.

Mitra and Ray (2014) study intergroup conflict led by economic changes within the groups and apply their model to data on Hindu-Muslim violence in India. Empirically, their main result is that increases in Muslim well-being (measured by per capita expenditures) cause a large and significant increase in future religious conflict, however, an increase in Hindu well-being has a negative or insignificant effect. They use a three-period Indian panel with region and time effects, and find their result is robust in many dimensions. They explain that this difference between the groups is possibly due to the origins of the violence in post-Independence India.

The multitude of unresolved topics within conflict is still large, with leading theories still failing to have been rigorously tested empirically (Blattman and Miguel, 2010). At the start of this section, I stated that historians believe that wars are often sparked by a combination of factors. A factor, which holds a persuasive argument to the extent that it is often included within formal models that have been empirically tested, is fear.

1.1.2 Conflict and Fear

The 2005 Nobel Memorial Prize in Economic Sciences was awarded to Robert Aumann and Thomas Schelling for “having enhanced our understanding of conflict and cooperation through game-theory analysis”. One of Schelling’s major pieces of work was ‘The Strategy of Conflict’ (1960), in which he discussed a ‘fear spiral’ now termed “Schelling’s dilemma”. When players are unsure of whether their opponents are aggressive or peaceful, having actions which are strategic complements leads to mutual fear and aggression. The origin of these beliefs could be irrational but when types are private information, they can spread through rational contagion. Schelling’s book gives an illustration of how uncertainty may lead to an unwanted attack by using the case of a homeowner confronting a burglar, both of whom are armed. The players may have no desire to kill yet a surprise attack would be the best response when they believe the other will attack.

Within pure coordination games, he introduces the concept of a “focal point”, which is a solution that people will tend to use in the absence of communication. They may use such points because they seem natural or special to the player and Schelling explains how they are highly useful in negotiations, in particular, when players cannot completely trust their opponent’s words. Mehta et al. (1994) support Schelling’s findings in the United Kingdom and the insight has been traced as far back as Hume in 1739. Many papers have investigated the concept since, for example, Sugden (1995) who gives a theoretical framework that explores the implications of collective rationality for a range of games, by distinguishing between the strategic structure of a game and the way the players describe the games to themselves.

Schelling’s thoughts were formulated into a model by Jervis (1976). He explained that the predominant underlying problem of the ‘fear spiral’ was neither human psychology nor human nature, but the fact we are living in a Hobbesian Jungle. Without a sovereign, each country is only protected by its own strength and, therefore, decision-makers worry about the most implausible threats. Even if leaders believe that others are not currently acting aggressive, there is no guarantee that they will not develop bellicosity later on. Of course, so-called paranoia may turn out to have a factual basis. For years, historians believed Frederick the Great of Prussia was paranoiac because he insisted the Seven Years’ War was preceded by a foreign conspiracy against his state. After accessing secret archives, it turned out that this was true.

Jervis (1978) adds that the lack of a sovereign not only permits war but it inhibits a state’s ability to achieve common interest goals when they are satisfied with the status quo. This is because the policies of cooperation that bring mutual awards, if conducted, could bring disaster should one side renege. Therefore, they are all worse off than they could be in the absence of sovereignty. This is also echoed in the literature on property rights (see Besley and Ghatak, 2010).

This situation of bilateral threat has been examined by many authors since (Janis and Mann, 1977; King et al., 2001; Axelrod, 2006). Janis and Mann (1977) build a “conflict model” on the actual decision making process, which can also be applied to study the decision to engage in a conflict, itself. They give an order of stages that an individual goes through, from identifying problems to implementing the decision, and state that the stress associated with the process affects the individual’s behaviour; evaluating risks causes a person stress which in turn causes them to make more inappropriate evaluations. However, an absence of stress could lead to a less productive search for information and also lead to unvigilant evaluations. Axelrod (2006) gives a more general discussion, using the round-robin computer tournaments with a repeated prisoner’s dilemma game to understand the evolution of cooperation. He uses this to show the potential extensions towards games with varying degrees of responses, with three or more players or with uncertainty of players’ objectives like the aforementioned studies.

The repeated prisoner's dilemma game is particularly interesting due to its application as an example of a concept called Folk theorem. The theorem suggests that anything is possible if it is feasible and individually rational (Friedman, 1971). Folk theorems can be compared in several aspects, such as whether the game is repeated finitely (Benoit and Krishna, 1985) or infinitely (Friedman, 1971), whether the equilibrium type is Nash (Aumann and Shapley, 1994) or Subgame-perfect (Rubinstein, 1979), or whether the utility of a player is an arithmetic mean (Benoit and Krishna, 1985) or discounted sum (Fudenberg and Maskin,1986). In the one-shot prisoner's dilemma game, the only Nash equilibrium is that both players defect. In an infinitely repeated version of the prisoner's dilemma game, the Folk theorem states that sufficiently patient players mean there is a Nash equilibrium where both players cooperate on the equilibrium path.

Kydd (1997) focuses on how military preparations show a desire for aggression, even when they are merely defensive. The action-reaction cycle may occur even without a factual basis; when a state assumes the opposition's escalation is a sign of aggression even though they, themselves, are only escalating out of a desire for security. Although Jervis explains this is due to psychological bias, Kydd's paper proposes conditions under which such escalation occurs as well as when 'security seekers' may refrain from escalating. In particular, using a model of incomplete information he shows that states that have a high cost of an arms race may rationally show their peaceful intent by deliberately not building weapons to signal their peaceful intent.

The key to reducing fear may be addressing the past actions of the players. It has been suggested that conflict is developed from the entities' past actions (Rahim, 2010). Rahim discusses the definition of conflict and poses the question of whether it is a process or a behaviour, before concluding conflict is an interactive process between social entities, which is manifested by incompatibility, disagreement or dissonance. He looks at conflict management within organizations, drawing from when the United States came under more intense competition from abroad. He investigates intra- and inter- personal, as well as group models to develop a design for effective management at different levels in an organization. Parties have recognized opposing interests in what they perceive to be a zero-sum situation, they believe the other would act against them and this belief is likely to be justified by previous actions taken against them.

To understand the interactive process between social entities, I think it is important not to discount the impact of politics within the process of conflict development. Therefore, I will now discuss the relationship between conflict and political systems in more detail.

1.1.3 Conflict and Political Systems

For some time, an influential school of thought was that conflict boomed at the end of the Cold War due to the sudden introduction of a post-Cold War international system. It was then found that, to truly make sense of what was happening, we needed to acknowledge the significance of insurgents; a form of militia who harness assorted political agendas. Conditions favouring insurgency also favoured conflict (Fearon and Laitin, 2001). From studies such as these, it was understood that politics had a significant role in the understanding of war.

The impact of political systems is widely considered to be an important factor. Economic theory has looked at the political use of wars linked to economic performance, election cycles and war decisions (Hess and Orphanides, 1995). Observing US politics, it was found that an incumbent leader suffering from bad economic performance would initiate a war in order increase chances of re-election, even with information symmetry and voter rationality. This was due to the heads of state diverting voters who would have judged them otherwise on their (poor) economic performance. This causes them to test their leadership skills in a war. Using data on state involvement in international war from 1816 to 1975, de Mesquita and Siverson (1995) tested four of seven proposed hypotheses, expecting a change in a leader’s term-length after the commencement of a war. They found that although war was risky regarding re-election, the risk was mitigated if they had been in power for some time, particularly for authoritarians. Thus, in congruence with Hess and Orphanides’ study, democratic leaders chose wars more likely to be won.

Downs and Rocke (1994) use the phrase ‘gambling for resurrection’. They perform formal analysis of strategies that could prevent overly aggressive and overly passive leaders. This enables them to show how statesmen will continue a lost war to avoid removal from office, whether it is rational to or not. Fearon (1994) explains that this is because the leader will suffer audience costs, which increases proportionally with public confrontation. They use this framework to conduct equilibrium analysis showing how these costs allow agents to figure out their opposition's true preferences; whether they wish to settle or wage war. The paper finds a democracy will be able to signal their intentions more clearly and credibly but they are also less likely to back down. Therefore, democratic countries can reduce the 'fear spiral', however, once they have committed, they will not stop. Mukand and Majumdar (2004) look at how leaders persist with inefficient policies in the latter part of the electoral cycle. Bandyopadhyay and Oak (2010) show that under asymmetric information the electoral process can itself exacerbate conflict with leaders trying to signal competence through adopting Hawkish positions.

While the electoral process and re-election motives can trigger conflict, there is a co-called 'democratic peace hypothesis' prevalent in the literature; a suggestion that democracies will rarely fight against each other. Studies on this have been summarized by a paper relating regime type and conflict, by Tangeras (2008). He finds that democratic electorates can effectively punish leaders for bad decisions, making democratic leaders more selective of which wars to participate in. They find that, on average, these countries win more of the wars they have instigated themselves, and so, democracies must be the most peaceful of dyads. Not only are democracies less likely to start war but also they are the most unattractive of targets (Bueno de Mesquita et al., 1999). The leaders' fear of punishment causes them to exert additional levels of effort in winning. This military advantage stemming from their institutional constraints is studied via eight empirical regularities that comprise democratic peace in a game-theoretic setting. Levy and Razin (2004) also find that with informational asymmetries and strategic complements, conflicts between two democracies differ from other dyads as they have a higher probability of peaceful resolution.

Further game-theoretical studies have shown that the 'democratic peace hypothesis' is not the only relationship we can identity within political systems. Despite Tangeras (2008) finding autocratic dyads the least peaceful of duos, there some evidence to refute this, claiming there is a non-monotonic relationship between democracy and peace. When investigating the effects of domestic institutional environments on the prospect of cooperation, it was found that not only were democratic dyads more cooperative but so were autocratic dyads (Leeds, 1999). They suggest that this is because of similar abilities to commit credibly, to adjust to changes in the environment and to accept costs when policies fail. Peceny et al. (2002) note that this research must work from a more sophisticated notion of authoritarianism. They study three forms of autocratic regime – personalist, military and single-party dictatorships – and find some types are more peaceful than others, towards each other. Although single-party dyads were the only ones to have experienced war, with no personalist dyads or military dyads having gone to war since 1945, all were still more peaceful than a mix. They agree that the evidence is not as robust as democratic peace (shown by Bennett, 2006) and more research is needed on autocratic dyads.

Epstein et al. (2006) argue that partial democracies are the most important and least understood, not autocracies. Mansfield and Snyder (2005) concentrate on these regimes and state that when institutions fail to mature before mass participation, democratizing states are the most war-torn. Much theory has developed on the intermediate political system and this bolstered the non-monotonicity argument; Baliga et al. (2011) use an empirical analysis of militarized disputes in the nineteenth and twentieth centuries that fit their theoretical model and find limited democracies are the most aggressive regime of all. The paper uses Polity data to classify countries as full democracies, limited democracies and dictatorships, then uses Correlates of War data from 1816 to 2000, to suggest that while full democratization might advance peace, limited democratization might advance the cause of war. Their explanation for this is that in limited democracies and dictatorships, citizens are unable to overthrow a Hawkish leader, however, the leader of a limited democracy risks losing power by appearing too dovish.

It has been shown that when countries' motivations to go to war depend on their 'political bias', war may be inevitable. This 'bias' is a leader's risk-to-reward ratio of a war, compared to the country's ratio, and it is the strongest with limited democracies. When there is no bias, a war can be avoided but a bias on either side could result in the worst outcome (Jackson and Morelli, 2005).

Psychologists think that the beliefs that trigger conflict can be identified as five predominant domains; superiority, injustice, vulnerability, distrust and helplessness (Eidelson and Eidelson, 2003). This strong link between economic theory and psychology is what is making experimental economics research more popular nowadays allowing for testing of alternate hypotheses in a controlled setting.

1.2 Experimental Research

Since the late 1990s, experimental studies on conflict have become increasingly popular. As most primary games used in experimental research are not exclusive to one situation, they can often model conflict as well. Every public good game can be understood as the predicament faced when mobilizing forces for war. Every bargaining model could provide insights towards post-conflict negotiations. Every risky choice task can be seen as the decision between a peaceful settlement and the risk of launching an attack (Garfinkel and Skaperdas, 2012). This creates a difficult process of separating 'conflict experiments' from a completely arbitrary selection of general-purpose models, and made selecting our framework a particularly challenging task.

The use of a first-strike game has been used in the experimental literature in order to model a hostile environment (Abbink and de Haan, 2011). This game is of interest because of its deliberate removal of economic gains as a motive to attack, enabling us to isolate the effects of fear. In the game, two players accumulate earnings over a series of rounds. A player is able to deactivate their opponent in any round, causing the opponent to lose all earnings and have their future payoffs reduced, whilst further ensuring that the deactivator isn't deactivated, themselves, in future rounds; deactivation is hostile but also defensive. The experiment has four treatments that disentangle spite, fear of spite and trust. They find a lack of spite but a significant fear of it. They also find mutual trust reduced hostility. In our experiments, we use the stag hunt game, which is also referred to as the "trust dilemma", to model fear spirals. The justification for their use in our experiments is provided in Section 2.2.

In ‘The Handbook of Experimental Economics’ (1997), Abbink states that there are around fifteen major experimental economics topics on conflict and papers within these have been identified by those mentioning ‘conflict’ in their titles or abstracts. For this section, I will discuss some of these topics by arranging them into the following categories; anarchy, coordination problems, prevention and post-conflict behaviour.

1.2.1 Anarchy

The theme of the first experimental papers on conflict was anarchy; where there is no central state guaranteeing property right and individuals must protect their property by force.

Durham et al. (1998) started by investigating the ‘Paradox of Power’ (Hirshleifer, 1991), hoping to gain some understanding on how even though the richer half of a population should have more power and ability to exploit the poor, redistribution ultimately makes them the exploited. They found that, on average, their predictions of POP were confirmed. The poor failed to allocate all their effort to fighting, as theory would predict, but the authors justified this as any error from the corner solution of devoting all resources (by the poor) to fighting, would show up as a reduction in conflict. POP has also been studied with an asymmetric assignment of roles (Carter and Anderton, 2001), using a model by Grossman and Kim (1995). There are two players investing in production and arms, with only one player’s production vulnerable to seizure. This player can only invest in defence, whilst their opposition invests in an attack. The paper also supports theoretical predictions.

Grossman and Kim’s model was modified by a paper, which only allowed players to pick either production or predation (Duffy and Kim, 2005). Using large groups instead of individual players, this introduced a coordination problem. They also introduced a Government agent for their second treatment, who imposes a minimum-spend on defence and has the incentive to maximise production. This implied that, in equilibrium, it will choose just enough defence to deter predation and no predation happens. The data converged to their predictions and it was faster in the second treatment, although it was hampered slightly by their low number of observations and the presence of some heterogeneity across sessions.

Not all experimental models are based on trying to see if subjects play a static Nash equilibrium. Some experiments are in continuous-time and are considered more natural; however, the results are much harder to interpret. They try to see how much of resources people set aside for offence and defence actions. The less stylized environments have user interfaces similar to those of a videogame, creating a huge amount of heterogeneity of outcomes. In a study by Smith et al. (2009), the fraction of potential earnings actually realised ranged from 14% to 71%, but these studies do provide useful insight into the way subjects protect their endowment or try to use resources to steal other people's resources both in environments where co-operation is possible and others where it is not.

1.2.2 The Coordination Problem

These studies switch from anarchy to models of dictatorship and the manifestation of power. They wish to understand how rulers oppress and exploit, yet still stay in power. In order to understand this, researchers have modeled a coordination problem to determine the behaviour of the citizens. A sufficient number must physically join force for a challenge to be successful and it doesn't matter how many have the desire to, if nobody does anything. Thus, if everyone believes that others will not take action, then no challenge occurs, while if people believe others will join, successful challenges occur. Thus, even when everyone dislikes a regime, a revolution is a rare event when people have successfully coordinated on their beliefs to take action. This can help explain the unpredictability of uprisings, the reason why empirical data is so sparse i.e. revolutions are rare events.

Abbink and Pezzini (2005) study this requirement of a simultaneous belief that enough people will join the revolution. When and how these beliefs occur is unobservable but empirical evidence can show what factors underlie the action. They combine this data with a three-stage experiment of seven agents to investigate the link between preferences and actions. There is a governor who makes a decision on three policy dimensions and citizens choose whether or not they wish to oppose, after a stage of communication. If more than five oppose, the governor is replaced by a randomly selected successor. They found that greedy governors were likely to be overthrown but this threat was deteriorated by restrictions on communication. Therefore, their results suggest that a successful dictator should not take too much but should be tough on opposition.

Inequality has also been studied within these experiments, using players who are advantaged or disadvantaged at earning an income in a rent-seeking game (Abbink et al., 2008). They vary the degree of inequality in three treatments and find that the disadvantaged are more likely to riot. They also found a more unexpected result indicating that more inequality actually reduced the number of revolutions. By allowing the participants to match after each round, they investigated whether this was down to fear of a counterattack or to resignation. They found it to be the latter.

Leaders can also undergo coordination problems, when a dictator gathers support from group leaders for aid in attacking other groups. This has been termed as ‘Divide and Conquer’ and studies have been done on what conditions effect group leaders’ decisions to join forces against a dictator (Weingast, 1997). It has also been investigated whether these group leaders can stop a dictator from transgressing altogether (Cason and Mui, 2007, 2009). These results show that private ex-ante communication and a probabilistic stopping rule can effectively reduce transgression.

1.2.3 Conflict Prevention

These experiments try to analyse which environments make conflict more likely or more intense, thus providing some insight into preventative measures.

In light of recent history, experimental economists have taken an interest in counterterrorism strategies. Colombier et al. (2009) investigate the public good nature of these strategies, with agents being able to choose national (e.g. border control efforts) or international (e.g. cooperative military action) measures. Unless there are punishment or reward opportunities, countries will decrease their own probability of being hit but won't help others. It is also found that punishments are not as efficient as rewards, however, all of this information still seems far from an actually implementable counterterrorism plan as a reward system with the huge efficiency gains they suggest would be difficult to create outside of an experimental setting. Abbink and Brandts (2009) investigate independence struggles of minorities in multi-ethnic countries, using a rent-seeking game. Their work on separatist struggles finds that contribution to the 'war chest' is more than four times the sub-game perfect equilibrium prediction, suggesting their strategic environment resulted in a strong emotional charge. The laboratory environment had limited real-world context, which many argue, renders experiments such as these, worthless. McBride and Skarpedas (2006) agree, stating that conflicts between countries have a strong dynamic component that cannot be modeled by techniques like repetition.

1.2.4 Post-Conflict Behaviour

Experiments can be used as a measurement tool of conflict's impact on human behaviour.

Studying post-war Bosnia, Whitt and Wilson (2007) used fairness norms in a dictator game to determine interethnic relationships between Bosnjaks, Serbs and Croats. They found that after seven years, these norms have been re-established. However, when comparing the results relative to those given within this experiment only, there is still substantial discrimination. The inferences are not clear-cut as no one conducted an experiment before the war, therefore we do not know if the war is exclusively to blame. It may just be the effect of in-group identification.

This sort of technique is still valuable and is becoming increasingly pivotal in the evaluation of the policy impact. For example, Coleman and Lopez (2010) evaluate the "Peace and Development" programme in Columbia using games that measure "insular" and "inclusive" cooperation.

1.2.5 Political systems

There is surprisingly very little on the topic of different political systems and measuring their effects.

These experiments generally have a similar set up and require many participants, such as in Bausch (2013), where most sessions included 21 subjects. This paper, like many others, only considered a democracy and a dictatorship- it would be interesting to include treatments with a limited democracy as well. A coin was flipped to determine which regime the group existed under and a leader was randomly assigned as soon as the game began. They investigated the role of communication and proved that democracies allow agents to organise more efficiently. They also showed that, on average, democratic citizens were better off but their leaders would benefit from becoming autocratic. However, most experiments within this area of political systems do not vary in how they model and implement regimes (see Duffy and Kim, 2005).

Papers in this area discuss the efficiency of regimes but, from what I have found so far, they only concentrate on studying intra-group relationships (effects of introducing a government agent, communication) and not the consequences of some external source of interaction. In Chapter 3, I look at the impact of third party recommendation on the escalation of conflict.

1.3 Conflict and Communication: A Short Literature Review on Cheap Talk

Cheap talk is a concept within game theory, in which the communication between players does not directly affect their payoffs. There is an assortment of literature on information asymmetry, discussing the problems that may be caused by an inability to coordinate. Like other issues, such as the lack of property rights and the nonexistence of democracy, this asymmetry could be eased by communication.

Crawford and Sobel (1982) introduce the theory of cheap talk in their seminal paper. They study a situation involving a degree of conflict, where there is some correlation between an agent's true type and their preference over the receiver's true beliefs. This model has been used as the basis of many research papers, including that of Krishna and Morgan (2004) who use more extensive communication with several rounds and a number of experts. Aumann and Hart (2003), and Forges (1990), show the effects of multiple rounds of costless signaling, in order to expand equilibrium payoffs. The model has also been applied to the theory of coalition formation (Ghosh et al., 2011), showing how partisan media, via voting mechanisms, can affect political outcomes. Additionally, there has been noteworthy work regarding mediated talk, reported by Ray and Ganguly (2006) in a survey article. In my experiment of bi-lateral threat possibilities I analyse the role of information transmission. An open area of experimental research is how such communication possibilities can affect the stag hunt game of incomplete information. We investigate this when there is an arms race equilibrium without communication in Chapter 2.

In experimental economics, communication treatments differ not only in timing but also in who observes the messages. The role of private information is currently a fundamental issue. Players are likely to possess asymmetric information on key parameters, in order to gain a competitive advantage against their opponent. Levy (2004) and Ottaviani and Sorenson (2006) have studied reputation and how when agents are concerned about their image, it can affect the reliability of their advice. These studies, along with Ray and Ganguly's (2006) survey paper, open up the opportunity to consider the introduction of a mediator in my experiments. It is interesting to deliberate on what may happen if a player with a conflict of interest enters our regimes. A modern day example of this player could be the media, who will produce partially informative messages that are actually very strong influentially but may have a bias. Chapter 3 explores this in a stag hunt game of incomplete information.

Chapter 2

Conflict and communication: can cheap-talk reduce aggressive behaviour?

2.1 Introduction

We experimentally analyse the effect of non-binding communication to see if non-binding pre play communication can reduce the play of aggressive strategies in a scenario akin to an arms race. This sheds some light on the effectiveness of an "arms control treaty". As there is no "World Court", any agreement would be pure cheap-talk. Aggressive play occurs because of the fear that the opponent will play aggressively and we look at whether cheap-talk can reduce that. In order to investigate the role of communication, we first show how an arms race type scenario can be triggered within an experimental laboratory. We do this using an adaptation of a model by Baliga and Sjostrom (2004), in which they show how even a small chance that the opponent may be building weapons dramatically increases the chances of everyone building, a so-called "arms race spiral".

It is not obvious that cheap-talk would be successful in this kind of situation because if all types send the same message, talk is not informative and cannot prevent an "arms race spiral". However, Baliga and Sjostrom (BS, 2004) show that talk is indeed informative. This occurs because while all types want to reduce the probability that their opponent arms, all but the players who are dominant types (i.e. who always arm) also prefer to know the type of their opponent. Thus, they have incentives to send different messages. This creates a scenario (discussed in more details in Section 2.3.3) where all but the dominant types do not play aggressive strategies. The communication between players helps by creating credible information about the players' types.

The experimental design, based on BS, 2004, consists of a stag hunt game with incomplete information. In the Stag Hunt game, each player has to choose amongst two strategies: *Defect* is a safer, risk dominant strategy that yields at least c. The strategy *Cooperate* might yield the higher payoff of d – if the opponent chooses *Cooperate* as well – but it is risky since it yields only $b<c$ if the opponent chooses to *Defect*. It is easily seen that there are two Nash equilibria, (*Cooperate*, *Cooperate*) and (*Defect*, *Defect*). Hence, when playing this game, players face a coordination problem. The dilemma is whether, in view of the strategic uncertainty, one should rather play according to the safer equilibrium. Figure 1.1 shows a generic stag hunt game.

	Defect	Cooperate
Defect	a , a	c , b
Cooperate	b , c	d , d

Figure 2.1: Game g(x) ($d > c \geq a > b$)

Many game theorists, including Schelling (1960), would argue that the Pareto-dominant equilibrium (*Cooperate*, *Cooperate*) is the natural focal point in the game. This outcome has the largest payoff with each player better off than if they chose to *Defect*. Schelling's principle of tacit bargaining proposes that since the players know that if they could communicate they would talk each other into (*Cooperate*, *Cooperate*), they will also be able to reach this conclusion without actually communicating. However, these arguments in favour of (*Cooperate*, *Cooperate)*, which postulate collective rationality, are not wholly convincing in reality. It is possible to argue that this payoff-dominant strategy isn't even always the intuitively most appealing solution.

In BS, 2004, they give the example of two states having to decide whether or not to invest in a new weapons system (called the 'Build' strategy). Each state knows the best possible outcome is for neither side to invest in new weapons, but the worst possible situation is to be unarmed when the other side is armed. Therefore, no state wants an arms race, but each state will acquire new weapons (‘Build’) if it believes the other state will acquire them. With these preferences as common knowledge, there exist two pure strategy Nash equilibria: an “arms race equilibrium” in which both states acquire new weapons (*Build, Build*) and a “détente equilibrium” in which neither state acquires new weapons (*Not Build, Not Build*). Rational players should be able to coordinate on the Pareto dominant détente equilibrium, perhaps using communication (O’Neill, 1999). However, Baliga and Sjostrom shows that if each state assigns some very small probability to the event that the opponent is a truly aggressive type for whom acquiring weapons is a dominant strategy, then under some conditions, this gives a *multiplier effect*; an escalating cycle of pessimistic expectations towards mutual armaments leading to everyone playing the ‘Build’ strategy, leading to an arms race.

Baliga and Sjostrom suggest the role of communication to see if this can lead to an equilibrium where all but dominant types play the ‘Not Build’ strategy. They consider an extension in which the players are allowed to send a costless message in the first stage. Thus, before making the decision to ‘Build’ or not, each state can send a Hawkish (aggressive) message or a Dovish (conciliatory) message to the opponent. These messages are pure cheap-talk and the players are free to arm themselves (‘Build’) regardless of what messages were sent. Since all types are better off if their opponent does not arm, one may believe that all types would send a Dovish message as it is most likely to persuade the opponent not to arm; thus, rendering cheap-talk uninformative, and not reducing the likelihood of an arms race whatsoever. However, cheap-talk can be informative. For the types who are not truly aggressive, it is not only the probability that the opponent defects that matters, but it is also important to be able to coordinate with their opponent. They show that if the dominant types are sufficiently rare then, with cheap-talk, there exists an equilibrium where the probability of an arms race is close to zero. Thus, communication can expand the set of equilibria and reduce the probability of an arms race.

Our experiments are designed to analyse two questions in the BS, 2004 environment. First, we consider whether there are differences in play of the aggressive strategy (‘Build’) when there is a unique equilibrium (this holds when a so called ‘Multiplier Condition’ is present that leads to a spiral in the presence of a dominant type) vs.

when there are other equilibria present (i.e. the ‘Multiplier Condition’ does not hold). Next, we consider the introduction of nonbinding, pre-play communication to see if there is a reduction in play of the aggressive strategy. We now describe the experiment in more detail.

The experiment has a 2x2 between-subject design. Overall, the design consists of a one-shot stag hunt game with four treatment conditions.

Table 2.1: Four treatments of the experiment

	Multiplier Condition Holds	Cheap-Talk Enabled
Treatment 1		
Treatment 2	x	
Treatment 3		x
Treatment 4	x	x

Treatment 1 serves as a baseline, Treatment 2 gives the pure ‘Multiplier Condition’ effect, Treatment 3 gives the pure communication effect and Treatment 4 shows the full interaction i.e. when the ‘Multiplier Condition’ is present and there is communication.

Subjects were only allowed to participate in one treatment and played with an anonymous opponent. The sessions did not last more than forty minutes each. There were a total of 160 participants, as each treatment included two sessions with 20 participants in each. Treating each game (pair of participants) as an independent observation, this gave 10 independent observations per session and a total of 20 independent observations per treatment. In Treatments 3 and 4, costless free-form messages were exchanged simultaneously before the players submitted their decisions. Players submitted their decisions simultaneously and then their beliefs about their opponent’s behaviour were elicited. The payoff matrices determined their payoffs, the outcomes were disclosed at the end of the experiment and all participants were paid privately.

BS (2004) that if the ‘Multiplier Condition’ is satisfied, there is a unique Bayesian Nash Equilibrium where all types choose to acquire weapons. Thus, our first hypothesis is that when the condition did not hold in Treatment 1, we expected at least some types to play the ‘Not Build’ strategy while when the ‘Multiplier Condition’ held in Treatment 2, we expected all types to choose ‘Build’.

For our second hypothesis, we considered the choice of individuals in the first and last treatments. BS, 2004 showed that if the probability that the player is a dominant type is very small, then, using cheap-talk, there is an equilibrium where the probability of an arms race is close to zero. We compared both the baseline of Treatment 1 and Treatment 2 where the ‘Multiplier Condition’ holds with Treatment 4, which showed the full interaction of the ‘Multiplier Condition’ and cheap-talk. We predicted that in the latter treatment, all types (except the dominant Hawks) would be more likely to play the ‘Not Build’ strategy.

We present our analysis in two parts. First, we give a descriptive analysis; checking whether there are differences in the frequency of the ‘Build’ strategy between treatments and formally examining whether or not there are significant differences in the observed frequencies among treatments by using a non-parametric Wilcoxon ranksum test. Then, as the non-parametric statistical test does not control for other characteristics that may affect an individual’s decision, we also use a multivariate regression to address our hypotheses. We use a probit model as the dependent variable is binary.

We found there is no significant difference in the frequency of playing ‘Build’ between Treatment 1 and Treatment 2; participants do not choose ‘Build’ more when the ‘Multiplier Condition’ holds. However, there is a significant difference in the frequency of ‘Build’ between Treatments 1 and 2 and Treatment 4; participants choose ‘Build’ more when the ‘Multiplier Condition’ holds and when cheap-talk is enabled. We briefly investigate the individual effect of cheap-talk using data from Treatment 3, to check whether the effect is most likely from the full interaction, which includes the ‘Multiplier Condition’, or possibly from the communication alone.

The remainder of this paper is organized as follows. Section 2.2 provides a literature review; discussing papers with similar information structures, the role of cheap-talk in this area, broader models of arms races and relevant experimental literature. Section 2.3 explains the model and Section 2.4 describes the experimental design. In Section 2.5, we give our hypotheses. In Section 2.6, we give our results and discuss them. In Section 2.7, we conclude.

2.2 Literature Review

Schelling (1960) first introduced the multiplier effect discussed by Baliga and Sjostrom. His paper explained how a non-zero-sum conflict could exist. It was known that conflict could coexist with common interest but Schelling suggests that both are aspects of an integral structure and aren't separable. He wrote that "the best choice for either player depends on what he expects the other to do, knowing that the other is similarly guided, so that each is aware that each must try to guess what the second guesses the first will guess the second to guess and so on, in the familiar spiral of reciprocal expectations".

He sets up the problem and illustrates the multiplier effect, using an example of mutual distrust. He describes the situation of finding yourself face-to-face with a burglar, where you're both armed and where there is the danger of an outcome which neither of you desires. Whether or not either of you want him to leave peacefully, he may think that you want to shoot him. Even worse, he may believe you think that he wants to shoot you. Or he may think you think he thinks you think… and so on. If the gains from shooting first are less than the gains from not attacking at all, there is no basis of attack for either player and there is little justification of self-defence. However, introducing the smallest temptation may proliferate through a process of interacting expectations.

This payoff matrix can be exemplified by a number of scenarios. It is equivalent to any problem of two partners who lack confidence in each other, for example, if some members of a gang are being closed in on by the police, the rest might be tempted to kill them, further tempting those being pursued to become snitches. Schelling builds an explicit model in which two rational players are victims of the logic that governs their expectations of each other. He analyzes whether a multiplier effect can be generated and then how this phenomenon can arise through a rational choice of strategies and the calculation of probabilities.

Baliga and Sjostrom's paper differs from previous literature both in terms of the incomplete information structure and the cheap-talk extension. We now discuss this in detail, stating other game-theoretical papers that have developed Schelling's information structure. We also look at papers that have studied equilibrium selection in the Stag Hunt game plus experimental papers based on these developments.

A paper by Rubenstein (1989) applies Schelling's argument and discusses "almost common knowledge", where there are only a finite (but large) propositions of the type "I know that you know that I know… that the game is G." He analyses this using a game with incomplete information and Halpern's (1986) example of the "coordinated attack problem". In this problem, there are two divisions that will win a battle if they both attack simultaneously, whereas they will be defeated if only one attacks. The commanding general of the first division wants to coordinate and he knows neither will attack unless he is sure the other will attack with him. The divisions can only communicate via a messenger but it is possible he may get lost or captured, and if he does make it to the camp, the general will now require a reply in order for him to know his message was received. The returning messenger would also risk not getting across and the second general will also require a response to know his confirmation had been received.

Rubenstein investigates how long it would take them to coordinate an attack, using his own game called "Coordination Through Electronic Mail". Two players want to coordinate on the state of nature (either *a* or *b*) and only the first is told the true state. The players are in two different locations and can only communicate by signals. The sending of messages is not a strategic decision by the players, it is automatically carried out by the computers; a message is only sent if the state of nature is *b* but due to "technical difficulties", there's a small probability that the message doesn't arrive. The computers send a confirmation automatically if any message is received, including the confirmation of the initial message, and the confirmation of the confirmation, and so on. The players are only able to choose *a* or *b* after the communication phase has ended and since there are only a finite number of messages transferred, the players never have common knowledge of the state of nature. They were given the number of messages their machine had sent and, sometimes facing higher order statements, they made their decisions. With the results from this game, he convinces us that in the "coordination attack problem", even if there is no common knowledge of the attack at the end of the information transmission, there is a possibility with a positive probability that they will both attack at dawn, especially if the probability of messenger failure is very small.

Jervis (1978) highlights that there is not much distinction between offensive and defensive weapons because even if a country arms for defensive purposes, they will make other countries feel less secure. He calls this the "security dilemma", where the suspicion that a state may arm, for whatever purpose, makes others more likely to arm in self-defence, creating a multiplier effect (or an "arms race spiral") that makes everyone worse off. He based this discussion on a stag hunt game, where the best possible outcome is for nobody to arm, but each player prefers to arm if they think the opponent will. If it is common knowledge that the pay-offs are those of a stag hunt game, it is not clear why rational players could not refrain from an arms build-up. Jervis argued that it, therefore, must be driven by irrationality, although it is possible it could just be a coordination failure.

Many papers have studied the Stag Hunt game and equilibrium selection. Carlsson and Van Damme (1993) highlight the hazards of taking short cuts in matters of equilibrium selection and they implore for a more fundamental approach where solutions are derived from individualistic assumptions. The use of an informational set-up implying a considerable weakening of the common knowledge assumption was shown to generate a model with different equilibrium selection properties. They compute and compare the solutions of the equilibrium selection theories by Harsanyi and Selton (1988), Guth and Kalkofen (1989), and Guth (1992), for the Stag Hunt game.

A common feature of those theories is that the selection of a particular equilibrium results from the individual players' strategic uncertainty. Kydd (1997) also suggested incomplete information about the opponent's preferences by introducing states that want war, and some that want peace, into Jervis' model. His discrete type space meant there was no contagion and he discussed how initial armaments could signal a player's type in a repeated game. Baliga and Sjostrom showed how, with a continuous type space, a *multiplier effect* could be triggered by an arbitrarily small probability that a player "wants war". Although in theoretical models, mixed strategies are commonly studied in games of multiple equilibria, we focus exclusively on pure strategies. We consider Kyd's method and use a discrete version of Baliga and Sjostrom's model, where there is still a unique Bayesian Nash Equilibrium.

However, there are still major differences in the papers discussed by Carlsson and Van Damme; the Harsanyi/Selton approach selects a specific form of uncertainty, a uniform prior in some fixed scheme of expectations formation, whereas in Carlsson and Van Damme, the uncertainty results from more fundamental assumptions about how the players' information is generated; they assume players signals/types (and payoffs) consist of a common shock plus a noise term. They derive the selection rule within a strictly non-cooperative framework by perturbing the game to be solved and inserting it in a game of incomplete information, deeming Harsanyi and Selton's approach to have a somewhat ad hoc character.

In this 1993 paper, Carlsson and Van Damme showed that for two-player 2x2 games, Harsanyi and Selten's risk dominance criterion could be justified with slight payoff uncertainty. In their later paper, they demonstrated that some of their previous model's properties, which only covered 2x2 games, could be generalized to a class of n-person symmetric binary choice problems. They showed that introducing slight payoff uncertainty allows the equilibrium selection problem to be resolved by a process of iterative elimination of strictly dominated strategies, because when the noise is small, the players' types are very highly correlated and there is a unique equilibrium. However, although they found that if there are two players, all solutions from the equilibrium selection theories coincided, if the number of players exceeds two, then these approaches no longer yielded the same solutions.

Cabrales et al. (2007) ran an experiment based on Carlsson and Van Dammes' theory of global games, studying what happens when they introduce uncertainty, similar to how we have implemented uncertainty within our experimental design. However, the uncertainty in their game was incorporated by sending each subject a noisy signal about the true payoffs. They find the Carlsson/Van Damme results do not hold under all circumstances. The game had a unique strategy profile that survives the iterative elimination of strictly dominated strategies and thus had a unique Nash equilibrium. On average, this equilibrium outcome coincided with the risk-dominant equilibrium outcome of the underlying coordination game. In the baseline game, they found behaviour converged to theoretical prediction after 50 periods. The data suggested that this could be explained by learning. Other research in this area is by Heinemann et al. (2004) who also find a small but significant difference between complete and incomplete games and Cornand (2006) who extends their paper by introducing a public signal in addition to the private signal.

Like these papers, subjects participating in experiments frequently do not succeed in coordinating on the Pareto-dominant equilibrium. Van Huyck et al. (1990) also use experiments to approach the equilibrium selection problem and study a class of tacit pure coordination games with multiple equilibria, which are strictly Pareto ranked. They provide further evidence on how human subjects make decisions under strategic uncertainty by studying the conflict between efficiency and security. They ran seven experiments with a total of 107 students and used incentivized belief elicitation. The difference between the experiments were whether the pairings were random or fixed, the periods in which subjects made decisions and the number of periods; the game could be repeated up to 27 times. The minimum action was the only common historical data available to the subjects. They found coordination failure was due to strategic uncertainty because some subjects concluded that it was too risky to choose the payoff-dominant action and most subjects focused on outcomes in earlier period games. Their results showed that the payoff-dominant equilibrium was extremely unlikely, even in repeated play. They concluded that only the secure, inefficient equilibrium indicates behaviour that subjects were likely to coordinate on, in repeated play, when the numbers of players is large.

Straub (1995), Schmidt et al. (2003) and Clark et al. (2001) looked at the risk and payoff properties of games in more depth. Straub (1995) suggested that the existence of a payoff dominated risk dominant equilibrium was a necessary but not sufficient condition for coordination failure. He presented a trade off between risk dominance and payoff dominance, and then linked this to the speed of convergence towards an equilibrium. He ran eleven sessions that lasted at least nine periods and he tested seven different payoff matrices. In each session, ten subjects repeatedly played the same single-period two-person game against a different opponent. Schmidt et al. (2003) found that changes in risk dominance did significantly affect subjects, whereas changes in payoff dominance did not. In their experiments, observing the history of play greatly influenced the participants' behaviour, whether they were re-matched or not. Participants were given complete information about the payoff structure and after each round, all were told their opponent's decision and their own payoffs. They did not receive information on the decisions of the other pairs nor were they told the number of rounds. The experiment consisted of four different games with four sessions conducted per game. Three groups had no previous experience of the game, however, one subject group contained subjects randomly selected from those who had played in previous sessions. Clark et al. (2001) studied repetition and showed that although participants generally failed to attain the efficient equilibrium of a one-shot game, they could in the repeated version. They found the participants were using actions to signal future intentions in the repeated games. They ran four treatments; a Coordination Game in two treatments and a Mechanism Game in two treatments, and they conducted two sessions per treatment. For each game, they ran a one-shot treatment and a Repetition treatment. In the Repetition treatment, the subjects either played a ten-period repeated games or sequences of ten one-shot games.

We can also apply Schelling's argument to macroeconomic phenomena such as currency attacks and liquidity crises. Morris and Shin (2003) state that if spectators believe that a currency will come under attack, their actions in anticipation accelerate the crisis, while if they believe that a currency is not in danger of imminent attack, their inaction spares the currency and vindicates their initial beliefs. They say that the actual onset of an attack cannot merely be attributed to the self-fulfilling nature of beliefs. Conventional papers explain the process on forces outside of the theoretical model, ignoring the role of speculators' beliefs about other speculators' behaviour. They assume that speculators know exactly what other spectators will do.

Their paper acknowledges that although this isn't realistic, they may have some idea, and so it takes neither extreme. As there is some uncertainty about equilibrium, the spectators' behaviour depends somewhat on what they believe they will do. Self-fulfilling currency attacks lead to multiple equilibria when the fundamentals are common knowledge, but they demonstrate a unique equilibrium when speculators face a small amount of noise in their signals about the fundamentals.

Cooper (1999) provides a broad survey of similar macroeconomic coordination game literature. In his first chapter, he also uses experimental evidence to convince readers of the relevance of coordination games to macroeconomics. He lists the aforementioned (e.g. Van Huyck) papers which use experiments to research complementarities, isolating the various sources of complementarity and exploring their implications, and applies their findings to macroeconomic problems, for example, how increased effort by other agents leads the remaining agent to follow suit – becoming the basis for multiple equilibria and also giving rise to multiplier effects. Using these papers, he explores some of the theories of selection and he relates the key points to the multiplicity of Nash equilibria and their Pareto ranking. He states the most important of these points is confidence rather than conflict and that overall, the evidence points to the fact that Pareto domination does not provide a natural focal point; coordination failures can arise. Complementarities have been applied in a variety of other contexts too, including development, in which Hoff and Stiglitz (2001) and Ray et al. (2007) provide excellent discussions.

A common shock and correlated types may be a good approximation of some interactions. However, in many situations the costs and benefits from taking a certain action may be specific to a certain player. These authors assume that players' types are correlated in a way that would be unnatural in our context. Baliga and Sjostrom obtain a multiplier effect without any correlation of types. For example, in the bank-run model, liquidity shocks may be uncorrelated. In the arms-race model, the costs and benefits from acquiring nuclear weapons depend on many psychological, moral and political considerations that may be specific to a certain leader. The cost of developing nuclear weapons could also be much lower for a nation if they have access to technical expertise and fissile material. Whether a state has such access will depend on a variety of idiosyncratic events that may be independent across countries and will be hard for an opponent to verify.

Harsanyi (1973) studied incomplete information games with independent types. He showed that every Nash equilibrium of a complete information game is the limit of pure strategy Bayesian-Nash equilibria of any sequence of incomplete games that are close-by. However, Baliga and Sjostrom find conditions under which only the Pareto-inefficient “arms race” equilibrium is a Bayesian Nash equilibrium outcome of the game with incomplete information. Their results differ from Harsanyi’s because even though they assume each player is almost sure that the opponent does not have a dominant strategy, there is substantial uncertainty about the actual numerical payoffs and therefore no ordinal ranking of the outcomes. Thus, even though the underlying complete information games have multiple equilibria with high probability, when there is a common shock but little noise or there is no common shock but a large amount of noise, there is a unique Bayesian Nash equilibrium; it can be the result of either sufficient correlation of types or sufficient uncertainty about independent types.

This intuition has been made very precise by Morris and Shin (2002). Baliga and Sjostrom refer to their paper for a detailed discussion of the relationship between common shocks, independent types, noise and multiplicity. Morris and Shin first introduce a simple binary action example with random matching. They then discuss alternative interpretations, including incomplete information.

Two players are randomly chosen from a population. Each Player *i* is characterized by a payoff parameter x_i. They assume that the payoff parameter is normally distributed in the population. However, the draws from the population are not independent: two players are more likely to be chosen to interact if they have similar payoff parameters. Thus, x_1 and x_2 are jointly normally distributed with a correlation coefficient that measures their parameters’ similarities. The player knows the payoff of an action but there is strategic uncertainty over the opponent’s action, due to the uncertainty over the opponent’s payoff parameter. The player’s private value x_i is essentially their type. They have a uniqueness condition as when x_i is either very high (greater than one) or very low (less than zero), Player *i* has a dominant action, and strategic uncertainty is only relevant when x_i lies in between zero and one. When the condition is satisfied, each player chooses action 1 only if his type is above some threshold. If the correlation coefficient were identically equal to zero, there was complete information, and if the players’ common type was between 0 and 1, there would be multiple equilibria. Thus, there is a unique equilibrium either if there is sufficient variance of players' private values or if those private values are sufficiently closely correlated.

They also show how strategic multipliers can be interpreted across different interaction settings by studying an interaction game with the local heterogeneity, giving a useful qualitative distinction between different kinds of strategic multipliers. It highlights that the extreme sensitivity in the limit of a global game is closely related to the jumps between complete information equilibria that must occur if there is not common knowledge. In settings where there is a small amount of local heterogeneity, the local sensitivity is largest when heterogeneity is small, and it helps to explain why small actions by, say, the Government can have a large effect on outcomes.

We have given an assortment of literature on information asymmetry, and there is much more to discuss the problems that may be caused by an inability to coordinate. Like other issues, such as the lack of property rights and the nonexistence of democracy, this asymmetry could be caused by communication. So far, none of these papers consider the role of cheap-talk when Baliga and Sjostrom's focus is on the cheap-talk extension of their model. We will now introduce the concept of cheap-talk, discuss literature in this area and give existing experimental findings.

Cheap-talk is a concept within game theory, in which the communication between players does not directly affect their payoffs. Crawford and Sobel (1982) introduce the theory of cheap-talk in their seminal paper. They study a situation involving a degree of conflict, where there is some correlation between an agent's true type and their preference over the receiver's true beliefs. This model has been used as the basis of many research papers, for example, using more extensive communication with several rounds and a number of experts (Krishna and Morgan, 2004).

Most articles study the sender-receiver games where the informed party takes no further role in the game after sending a report to the receiver. McDaniel (2011) investigates the effects of uncertainty, however, this is one-sided, with one player having perfect information. They show that the benefit of cheap-talk depends on the relationship between payoffs and risks, and cheap-talk only benefits informed players when the stag hunt payoffs exhibit low risks. The experiment is 2x2 and consisted of two games; each randomised between prisoner's dilemma and stag hunt (one with riskier payoffs than the other). For each game they varied the amount of communication allowed. Either players simultaneously made choices at the start of each round or the player with perfect information was allowed to send a non-binding message to their opponent, and then they simultaneously made their choices. Subjects were matched with the same partner for 16 rounds and changed player types every four rounds. In contrast, Baliga and Sjostrom's paper assumes both players have private information, send messages and take actions. There is a small set of models with a related structure (Baliga and Morris (2002), Matthews and Postlewaite (1989), Austen-Smith (1990)).

Duffy et al. (2010) specifically look at costly communication and how messages available to the sender imperfectly describe the state of the world. This is more in conjunction with Baliga and Sjostrom's cheap-talk extension. They find that larger communication costs are associated with worse outcomes for both the sender and receiver. They find that there is over-communication and that the response time of both the sender and receiver are positively related to their payoffs.

Most closely related to our research is Banks and Calvert's (1992) paper, which uses a battle-of-the-sexes game. In their model, there is also two-sided incomplete information, and each player has his own favourite outcome but his type (high or low) determines the intensity of his preference. Without communication, there are efficient asymmetric Bayesian–Nash equilibria where one player always gets his favourite outcome, and the only symmetric equilibrium involves an inefficient randomization. However, Banks and Calvert show how communication can produce a symmetric equilibrium where coordination occurs more frequently. Their favourite action is also more likely to be chosen when the intensity of their preference is high.

In the battle-of-the-sexes game, players just want to take the same action, whilst in a stag hunt game, players always want their opponent to choose a particular strategy (not arm). Thus, the nature of communication is different in the two games. Aumann (1990) states that this makes communication particularly difficult, with each player wanting their opponent to take the same action whatever action they take themselves. He used a variant of the stag hunt game to show that an agreement to play the most efficient outcome actually conveys no information about what the players will do. Many experiments had been conducted on coordination in 2x2 games, but few with uncertainty similar to his. Since then, much research has been made in this area to show that, in simple coordination games, the effectiveness of communication is affected by many factors. These include the order of messages and actions, message credibility and relative payoff differences between Pareto ranked equilbria.

One of the many papers contesting Aumann was by Farrell and Rabin (1996), who said cheap-talk could achieve efficiency. They stated a message could be self-committing if it created incentives for the signaler to fulfill it and if the other player considered the message credible. Charness (2000) tested this and found coordination when the sender first chooses a signal and then an action. Without communication, he found that efforts to achieve efficiency are unsuccessful. He also found that reversing the order, and choosing the action first, led to a change in the players' behaviour. The experiment used 10 periods and subjects were in groups of six. Within their groups, the subjects were told that their partners and their roles would be re-matched after every round. Non-binding signals were made by writing "I intend to play [A or B]" on a piece of paper which were anonymously passed to their opponents. Different payoff calibrations were used to investigate the sensitivity to risk-dominance.

Clark et al. (2001) showed inconclusive results testing Aumann's conjecture. They investigated whether the credibility of subjects' messages is important to achieve socially optimal outcomes and looked at the sensitivity of communication to the structure of payoffs. They found that agreements to play a Nash equilibrium are fragile when players have a strict preference over their opponent's strategy choice. They also used a game where subjects had a positive incentive to represent themselves truthfully and found that informative communication does not always lead to the Pareto-dominant outcome. Their main experiment had four treatments, two games with varying payoffs structures, each played with and without communication. They ran treatments without communication for three reasons; it enabled the investigation of the robustness of behaviour to payoff variation, they could validate their procedures by comparing their results to existing literature, and it provided a baseline against which the effects could be judged. They used simultaneous, fixed messages and subjects were re-matched after every period for ten rounds. Similar to

Charness (2000), signals were made by sending “I intend to play [A or B]” and subjects were reminded that they were not required to make the choice that they had announced.

The aforementioned article by Banks and Calvert (1992) show how communication can produce a symmetric equilibrium and efficiency, if there is a mediator. In experimental economics, communication treatments differ not only in timing but also in who observes the messages. Players are likely to deliver asymmetric information on key parameters, in order to gain a competitive advantage against their opponent. Levy (2004) and Ottaviani and Sorenson (2006) have studied reputation and how when agents are concerned about their image, it can affect the reliability of their advice. These studies, along with Ray and Ganguly (2006), open up the opportunity to consider the introduction of a mediator (a “World Court”) in our experiments. In Baliga and Sjostrom’s model, they can work out first-best efficiency without a mediator when the dominant strategy types are rare. They also state that the issue of whether a mediator is useful when the dominant strategy types are not rare is an interesting topic for future work.

Schelling’s stated players may attack each other inadvertently because of a false alarm. If the underlying problem is an imperfect warning system, then Baliga and Sjostrom state that cheap-talk cannot be the solution. They showed that cheap-talk can be useful when the underlying problem is incomplete information about the opponent’s preferences and one can interpret their model as a model of war initiations; each nation must simultaneously decide whether to attack or not.

Their results are also irrelevant if the perceived first-mover advantage is large as then the true pay-off matrix is more likely to be a prisoner’s dilemma than a stag hunt game; as with the battle-of-the-sexes game, they state that pre-play communication is not useful in the former but is in the latter.

2.3 Model

2.3.1 Baliga and Sjostrom’s Stag Hunt Game

In the Stag Hunt game, there exist two pure strategy Nash equilibria; an “arms-race” equilibrium in which both governments acquire new weapons (*Build*, *Build*) (the equivalent of (*Defect*, *Defect*) in game g(x)) and a “détente” equilibrium in which neither government acquires any weapons (*Not Build*, *Not Build*) (or (*Cooperate*, *Cooperate*) in g(x)). The Nash equilibrium (*Not Build*, *Not Build*) Pareto dominates the other.

The game we use is essentially a coordination game augmented with incomplete information about players’ independently drawn types, and the true payoff matrix is a stag hunt game with a high probability. The payoffs are given as below where c_i is the cost of building weapons for Player i, with one type having low enough c for ‘Build’ to be a dominant strategy. We denote by c_i the cost of aggression for player i; if the cost of building is lower, then the player is more aggressive. We call d the disutility of being less advantaged with an inferior weapon system and μ the gain of being more advantaged.

	Build	Not Build
Build	$-c_i$	$\mu - c_i$
Not Build	$-d$	0

Figure 2.2: Payoff matrix of Player i in Baliga and Sjostrom’s (2004) model

2.3.1.1 Parameters

$-c_i > -d$ ensures that 'Build' is always a best response against 'Build'. However, 'Not Build' is a best response against 'Not Build' for Player i, **if and only if $c_i \geq \mu$**. Thus, if $c_i < \mu$, Player i is an "aggressive" type and 'Build' is a strictly dominating strategy for them. μ is small so that the probability of being the "aggressive" type, $F(\mu)$, is close to 0.

These parameters give payoffs that mean that the "aggressive" type, who has the highest propensity to arm, will always play 'Build' and arm. The "fairly aggressive" type has a medium propensity and prefers to take whatever action he thinks his opponent is most likely to take. Therefore, he will play 'Not Build' if assured facing peaceful ("normal") opponent, otherwise, he will play 'Build'. "Normal" types have the lowest propensity. They will only arm if they are almost sure their opponent will and even so, their payoff differential is not very high from playing 'Build'.

2.3.2 The Multiplier Effect

As mentioned earlier, introducing the uncertainty affects our decision on which one of the two Pure Nash equilibria to select. Assigning a small but positive probability on the event that the opponent is a truly aggressive type, for whom acquiring weapons is a dominant strategy, can trigger an arms race.

The situation is modeled as follows. Each state has a type parameterizing their propensity to arm. The state's true type is its private information, and types are independently drawn from a continuous distribution. At one end of the distribution are the "aggressive" types; they prefer to arm regardless of the opponent's actions (a fraction $\varepsilon > 0$ of the players). At the other end of the distribution are the "normal" types who prefer to arm only if they are virtually sure the opponent will arm. In between are the "fairly aggressive" types (a fraction $\delta > 0$ of the players), who prefer to arm when the opponent arms with even the smallest probability. The "aggressive" types have a dominant strategy to arm; they will certainly arm; but this triggers a *multiplier effect*, as the "fairly aggressive" ("almost dominant strategy types") must arm in equilibrium but then, all types that prefer to arm when the opponent arms with at least probability δ + ε must arm, etc. The spiral of pessimistic expectations and fear causes more and more types to arm. Therefore, even when each state thinks it is extremely unlikely that the opponent is the dominant strategy type, the unique Bayesian Nash equilibrium appears to involve an arms race with probability one.

BS, 2004 shows that when a so-called *'Multiplier Condition'* holds, it means that there can be no symmetric Bayesian-Nash equilibrium where 'Not Build' is chosen with positive probability. Therefore it guarantees that each type will strictly prefer to choose 'Build' whenever he thinks all types with a lower cost than him will choose 'Build', so the contagion to play 'Build' will infect the whole population.

Formally, the distribution satisfies the 'Multiplier Condition' **if $F(c)d \geq c$ for all** c. This means for each type: given all types with lower costs play aggressively, they prefer to play aggressively too, because the expected disutility of not being aggressive when all those with lower costs are aggressive (**$F(c)d$**), is bigger than the cost of their aggression (**c**). We choose parameters in Treatments 2 and 4 such that for each type, if they know that types with costs lower than them play the aggressive strategy, they would also play the aggressive strategy.

2.3.3 Cheap-Talk

Baliga and Sjostrom show there exists an equilibrium of the cheap-talk extension where the probability of an arms race is close to zero. Thus, communication can expand the set of equilibria.

They consider an extension where before making the decision to arm, each state sends a Hawkish (aggressive) message or a Dovish (conciliatory) message to the opponent. These messages are pure cheap-talk and a state is free to arm itself regardless of what messages were sent. The cheap-talk can be informative as although all types want to **reduce the probability that the opponent arms**, some players (the non-aggressive types) also want to **reduce the uncertainty of their opponent's action** to avoid a coordination failure. As the different types will trade off these two objectives at different rates, it may be possible to induce different types to send different messages, where the “fairly aggressive” type put the highest value on resolving the uncertainty about the opponent’s action, whereas the “aggressive” and “normal” types are more interested in reducing the probability of their opponent arming.

Essentially, the separation of the “aggressive” and “fairly aggressive” types prevents the *multiplier effect*, so the model requires these two types to send different messages. Thus, BS, 2004 construct a separating equilibrium. A Dovish message minimizes the probability the opponent arms and a Hawkish message yields a higher probability the opponent arms but then there is no ambiguity about the opponent’s action. Most types send the Dovish message. The exception is a small number of “fairly aggressive” types who have a high propensity to arm but not so high as to be dominant strategy types. They have the incentive to send a Hawkish message; it enables them to coordinate with fellow “fairly aggressive” types and play ‘Build’ with the “aggressive” types.

The “aggressive” types, on the other hand, have incentive to send a Dovish message as it enables them to pool with “normal” types and arm unilaterally. “Normal” types have incentive to send a Dovish message, as they will pool with “aggressive” types if they are faced sufficiently infrequently. If the number of “aggressive” types is very small, then the number of peaceful types is closer to one, so the probability is close to one that both states are Dovish and refrain from acquiring weapons. Ex ante, the “normal” types are better off trusting an opponent who appears Dovish. The “fairly aggressive” types, however, are not willing to take this gamble. The only way to prevent “aggressive” types arming unilaterally against “normal” opponents would be to get the “aggressive” types to reveal their true nature, alerting the opponent that they should arm too. This would not be incentive compatible though, since the “aggressive” type does not want his opponent to arm.

Therefore, “aggressive” and “normal” types minimize the probability their opponents arm whereas the “fairly aggressive” types minimize the probability of coordination failure. This is a non-monotonic equilibrium where types with the highest and lowest costs pool with the intermediate type separating out.

In summary; if the two states send different messages, the result is an arms race and both states acquire new weapons; if both states sent Hawkish messages, then neither acquires any weapons; if both states sent Dovish messages, then the “aggressive” types acquire new weapons, while the remaining types do not.

2.4 Experiment

We consider a discrete version of the Baliga Sjostrom model. There are three type of players; “aggressive” (type A, with costs c_L), “fairly aggressive” (type B with costs c_M) and “normal” (type C with costs c_H). The game and payoff are described below.

2.4.1 Parameters within the Experiment

Treatment 1: ‘Multiplier Condition’ does not hold

We have set c_H equal to 9 for “normal” types, c_M is 5.1 for “fairly aggressive” types and c_L is 3 for “aggressive” types. The disutility of being less advantaged, *d*, is 11 and the gain of being more advantaged, μ, is 5. These ensure that player with costs c_M will play ‘Build’ on the anticipation that c_L will play ‘Build’. This creates a contagion where a player with cost c_H will also play ‘Build’.

	Build	Not Build
Build	$-c_i$	$5-c_i$
Not Build	-11	0

Figure 2.3: Payoff matrix of Player i with values of d and μ inserted

Treatment 2: 'Multiplier Condition' holds

When the 'Multiplier Condition' holds, ***F(c)d* ≥ *c* for all *c*,** and so; c_M is 5.1 for "fairly aggressive" types, c_L is 3 for "aggressive" types but we lower c_H equal to 8 for "normal" types.

The probability of being an "aggressive" type is 0.1, the probability of being a "fairly aggressive" type is 0.4 and the probability of being a "normal" type is 0.5 across both treatments. These ensure that even if all c_M play 'Build'. The types with costs c_H may co-ordinate and play 'Not Build'.

Treatments 3 and 4 introduce pre play communication with and without the 'Multiplier Condition'.

2.4.2 Experimental Design

The experiment has a 2x2 between-subject design. Overall, the design consists of a one-shot game with four treatment conditions (See Table 2.1). In Treatment 1, there is no communication and the 'Multiplier Condition' does not hold. In Treatment 2, there is no communication but the 'Multiplier Condition' does hold. In Treatment 3, there is two-sided communication (cheap-talk is enabled) and the 'Multiplier Condition' does not hold. During the pre-play communication phase, each subject in a pair was asked to send free form message s. In Treatment 4, there is both cheap-talk and the 'Multiplier Condition' holds too. Treatment 1 serves as a baseline, Treatment 2 gives the pure 'Multiplier Condition' effect, Treatment 3 and 4 look at the effect of cheap-talk both when the 'Multiplier Condition' holds (Treatment 4) and when it does not (Treatment 3).

Subjects were only allowed to participate in one session and, therefore, only one treatment, with an anonymous opponent. We chose to conduct separate subjects per treatment, to rule out any dependence on what happened in the previous treatment. There were a total of 160 participants, as each treatment included two sessions with 20 participants in each. Treating each game as an independent observation, this gave

10 independent observations per session and a total of 20 independent observations per treatment.

With a prior probability, the subjects were assigned a type; Type A (the low-cost type in the model), Type B (the medium-cost type) or Type C (the high-cost type). They were not simply given their possible payoffs from each action which depended on what their opponent played. Furthermore, their actions were not referred to as ‘Build’ or ‘Not Build’, but rather as ‘Option X’ or ‘Option Y’, respectively. Thus, instead of ‘Build’ or ‘Not Build’; in the experiment neutral framing was used to describe these two actions. The payoff structure of the game, including the percentage of types in the experiment, was known. However, the type of each player was private information. Thus, the players knew their own type but not the other player’s type. The types were drawn from a computer programmed with a probability of being either one of the three with probabilities given as above.

In order to avoid the participants facing negative outcomes, the payoffs were scaled up by 14 points to ensure positive payoffs for everyone. The exchange rate was 1 point equaling £0.50p.

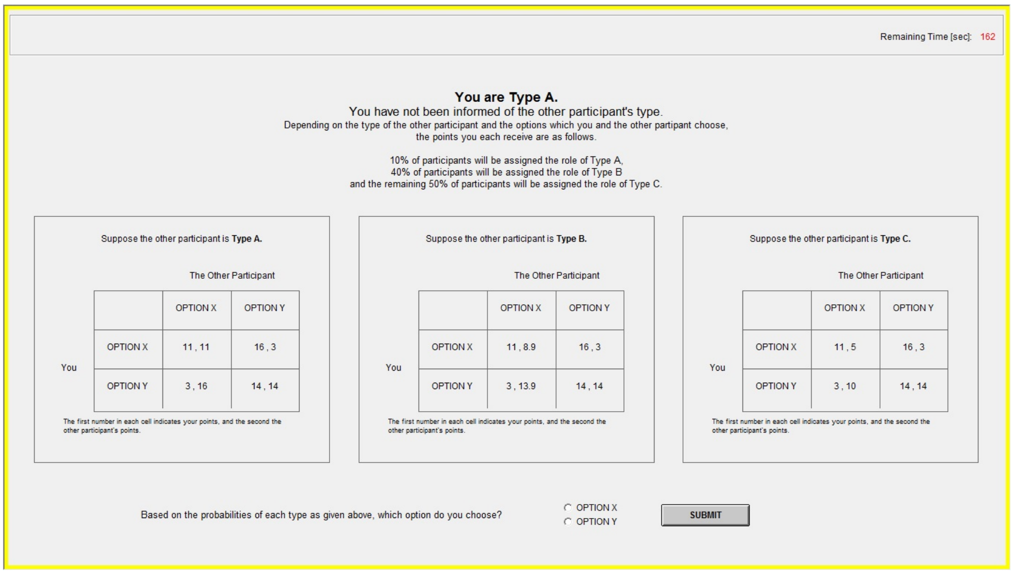
	Build	Not Build
Build	$14\text{-}c_i$	$19 – c_i$
Not Build	3	14

Figure 2.4: Payoff matrix for Player i when scaled

When the players were allowed to communicate, free-form costless messages were exchanged simultaneously in a time limited communication structure. While this was somewhat different from the messages in the model by Baliga and Sjostrom, this still captured the idea of using costless communication.

First, players were given instructions and asked to complete a quiz to make sure they had understood the information given. An experimenter read the instructions aloud and the experiment only began once all of the participants had answered the quiz's questions correctly. Then, they were assigned their type.[1]

In the next stage, players were matched in groups of two. As the players' payoffs differed depending on what type they were exogenously assigned, we provided a number of payoff matrices to show all of the possible outcomes for the participants from facing each type. The sets of payoff matrices are shown in Figure 5, Figure 6 and Figure 7. We have provided the screenshots from the experiment[2].



You are Type A.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B
and the remaining 50% of participants will be assigned the role of Type C.

Suppose the other participant is **Type A.**

The Other Participant

You		OPTION X	OPTION Y
	OPTION X	11 , 11	16 , 3
	OPTION Y	3 , 16	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is **Type B.**

The Other Participant

You		OPTION X	OPTION Y
	OPTION X	11 , 8.9	16 , 3
	OPTION Y	3 , 13.9	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is **Type C.**

The Other Participant

You		OPTION X	OPTION Y
	OPTION X	11 , 5	16 , 3
	OPTION Y	3 , 10	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

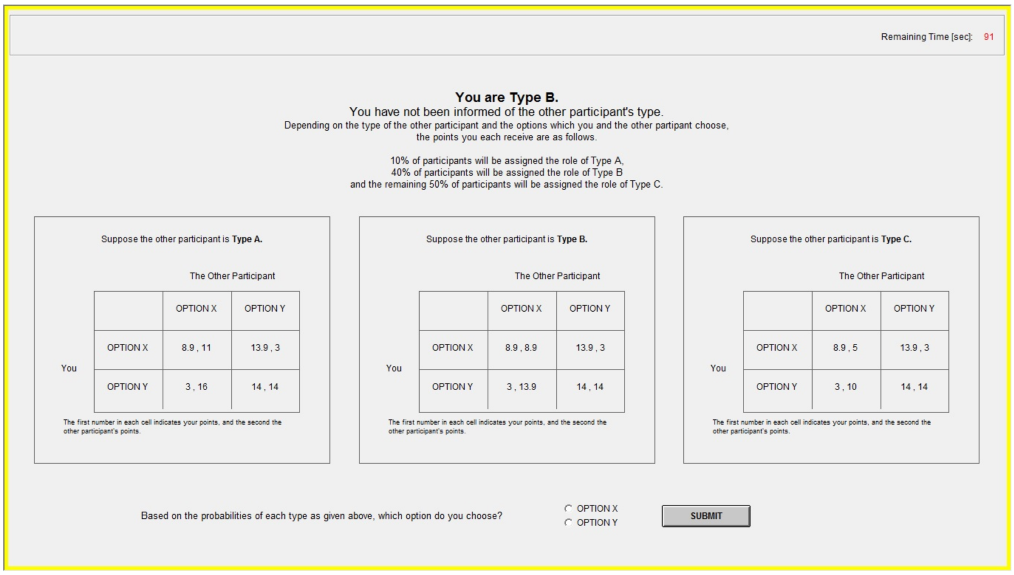
Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

Figure 2.5: Payoff matrices for Type A players in Treatment 1

---

[1] Full instructions are included in the Appendix.
[2] Payoff matrices for treatments with and without the 'Multiplier Condition' are in the Appendix.



Remaining Time [sec]: 91

You are Type B.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B
and the remaining 50% of participants will be assigned the role of Type C.

Suppose the other participant is **Type A.**

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	8.9 , 11	13.9 , 3
	OPTION Y	3 , 16	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is **Type B.**

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	8.9 , 8.9	13.9 , 3
	OPTION Y	3 , 13.9	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is **Type C.**

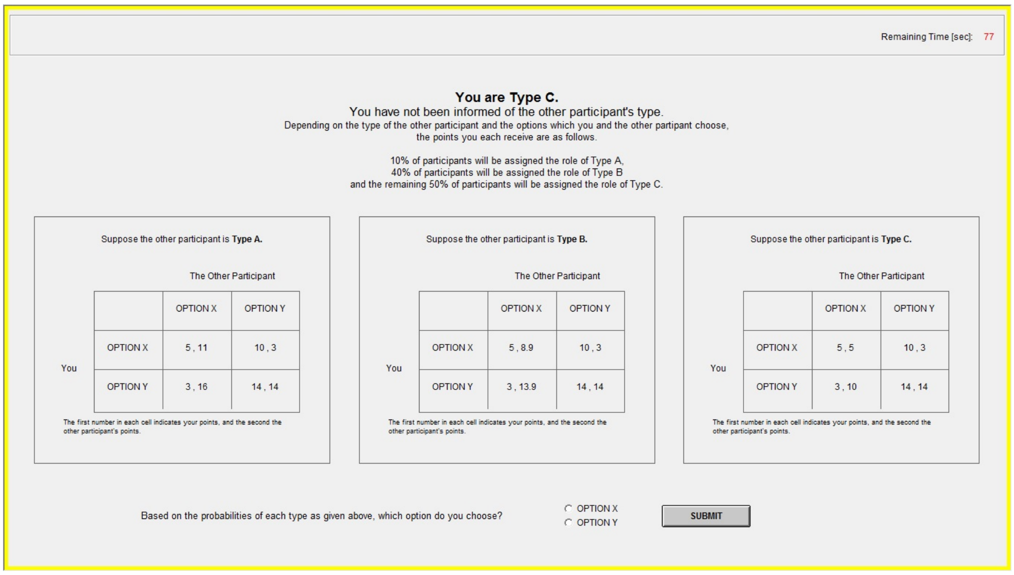
		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	8.9 , 5	13.9 , 3
	OPTION Y	3 , 10	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

Figure 2.6: Payoff matrices for Type B players in Treatment 1



Remaining Time [sec]: 77

You are Type C.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B
and the remaining 50% of participants will be assigned the role of Type C.

Suppose the other participant is **Type A.**

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 11	10 , 3
	OPTION Y	3 , 16	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is **Type B.**

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 8.9	10 , 3
	OPTION Y	3 , 13.9	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is **Type C.**

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 5	10 , 3
	OPTION Y	3 , 10	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

Figure 2.7: Payoff matrices for Type C players in Treatment 1

In Treatments 3 and 4, free-form costless messages were exchanged simultaneously before participants submitted their decisions.

Belief elicitation: Beliefs about their opponent’s behaviour were elicited. They were asked to indicate which option (X or Y) they thought each type they could be matched with, will choose. If they correctly estimated the actual choice made by their opponent, they received £2, thus there were incentives for correct elicitation of beliefs.

At the end of the game, the outcomes were disclosed to each person and everyone was paid privately. The final screen reported their type and their choice, their opponent's type and their choice, the total points they earned, the total points they earned in pounds, their earnings from the belief elicitation, their £2.50 show up fee, plus their corresponding final pound payment. Before they left the laboratory, participants were asked to complete a short questionnaire providing information such as their gender and how clearly they thought the instructions had been.[3]

2.4.3 Experimental procedures

The experiment was programmed with z-Tree software (Fischbacher, 2007) and all subjects were recruited at the University of Birmingham, using ORSEE software (Greiner, 2015). The experiment was conducted at the Birmingham Experimental Economics Laboratory. In total, 160 subjects participated in the experiment with 40 subjects per treatment. We ran two sessions per treatment, therefore, there were 20 subjects in each session. At the end of each session, subjects were privately paid according to their total amount of experimental currency units (ECUs), using an exchange rate of £0.50 per ECU. Average earnings (excluding a show- up fee of £2.50) were £7.10. Sessions lasted 40 minutes, on average.

2.5 Hypotheses

In this section, we present our theoretical hypotheses and discuss which ones hold and which do not.

Our first hypothesis is to check if people play the aggressive strategy more frequently when the 'Multiplier Condition' holds than when it does not. As shown in Baliga and Sjostrom, if the 'Multiplier Condition' is satisfied, there is a unique Bayesian Nash Equilibrium where all types choose to acquire weapons whereas when it does not, other equilibria exist. When the 'Multiplier Condition' holds in Treatment 2, we

---

[3] The questionnaire is included in the Appendix.

expect all types to choose X, including Type C while we don't expect this to be true in Treatment 1.

Hypothesis 1
Null Hypothesis:
The frequency of X in Treatment 2 is at most the frequency of X in Treatment 1.
Alternate Hypothesis:
The frequency of X is greater in Treatment 2 compared to Treatment 1.

In order to reject the null, the frequency of X must be greater in Treatment 2 than in Treatment 1.

BS, 2004 also predict that if the probability that the player is an "aggressive" type is very small, then, using cheap-talk, there is an equilibrium where the probability of an arms race is close to zero. We test two hypotheses in regards to communication. Firstly, we consider the choice of individuals in the first and last treatments. In Treatment 1, there is no communication nor does the 'Multiplier Condition' hold. Treatment 4 shows the full interaction of the 'Multiplier Condition' and communication, and we predict more types will then play Y.

Hypothesis 2
Null Hypothesis:
The frequency of X in Treatment 4 is at least the frequency of X in Treatment 1.
Alternate Hypothesis:
The frequency of X is lesser in Treatment 4 compared to Treatment 1.

In order to reject the null, the frequency of X must be lesser in Treatment 4 than in Treatment 1.

Furthermore, we test whether communication even with the 'Multiplier Condition' holding, as the BS theorem states in Cheap-Talk model, that if the condition holds, more people will play 'Not Build'. We make a comparison between the data from Treatment 2 and Treatment 4. In Treatment 2, the 'Multiplier Condition' does not hold but there is communication. Whereas, in Treatment 4, the ‘Multiplier Condition; holds and there is communication.

Hypothesis 3
Null Hypothesis:
The frequency of X in Treatment 4 is at least the frequency of X in Treatment 2.
Alternate Hypothesis:
The frequency of X is lesser in Treatment 4 compared to Treatment 2.

In order to reject the null, the frequency of X must be lesser in Treatment 4 than in Treatment 2.

2.6 Results

In this section, we present the findings from our experiment and see if our hypotheses hold.

We present our analysis in two parts. First, we give a descriptive analysis; checking whether there are differences in the frequency of X between treatments and formally examining whether or not there are statistically significant differences in the observed frequencies among treatments by using a non-parametric Wilcoxon ranksum test. Then, we use multivariate probit regressions to see if the differences hold after we include a host of control variables that are likely to affect the dependent variable we consider.

2.6.1 Hypothesis 1: The effect of the ‘Multiplier Condition’

Based on the model by Baliga and Sjostrom, we can predict that the ‘Multiplier Condition’ will affect whether all types prefer to play X. The paper states that if the ‘Multiplier Condition’ is satisfied, there is a unique Bayesian Nash Equilibrium where all types choose to acquire weapons. To test this hypothesis, we consider the frequency of X in the first two treatments as Treatment 1 is the baseline and

Treatment 2 shows the pure effect of the ‘Multiplier Condition’. In order to reject the null, the frequency of X must be greater in Treatment 2 than in Treatment 1.

We look at the observed frequencies of how often the individual’s chose X. Each treatment consisted of 80 subjects. Our data shows that there is an increase of four participants choosing X when the ‘Multiplier Condition’ held.

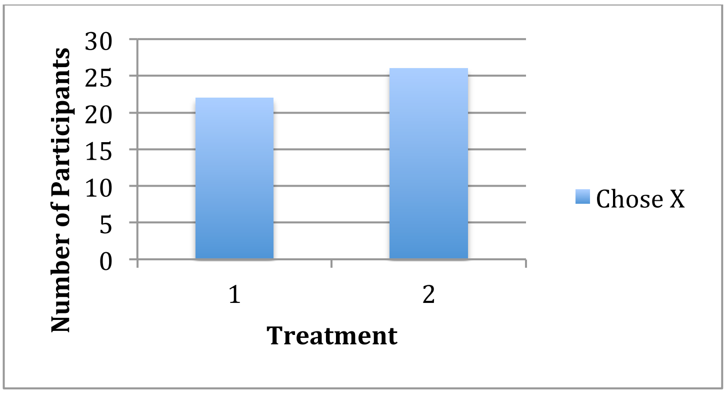


Figure 2.8: Frequency of X in Treatments 1 and 2

We can also look at the proportion of X within the individual observations. An individual observation consists of a group of two subjects. We measure the proportion of X by showing the percentage of each group that chose X; so if both members chose X, the group is assigned 100. The frequency of the outcome (X,X) increased from six groups (out of 20) to nine groups. This represents the number of groups in which both subjects chose X, and shows an increase of 50%. The frequency of the outcome (Y,Y) decreased from 18 individual observations to 14, meaning it has fallen by 22%. These results seem to support our first hypothesis.

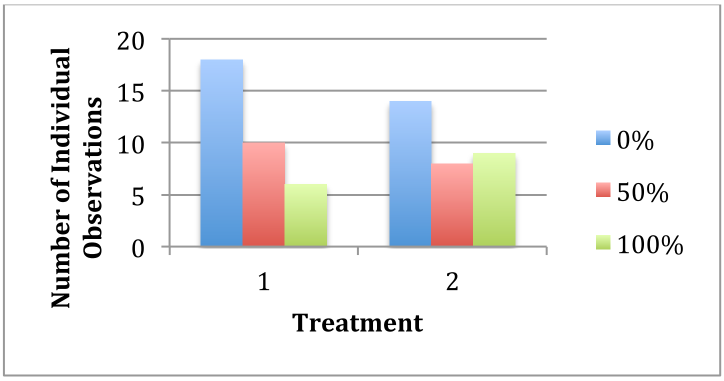


Figure 2.9: Proportion of X within individual observations in Treatments 1 and 2

We consider the proportion of X within the individual observations and present, in the table below, the means of Treatment 1 and 2.

Table 2.2: Mean proportion of X within individual observations in Treatment 1 and 2

	Treatment 1	**Treatment 2**
Mean	55	65

If all participants chose X, the mean would be 100. From Table 2.2, we observe that the mean proportion of choosing X in each group is 55 in Treatment 1, and 65 in Treatment 2. Therefore, Treatment 2 has a mean that is 27% higher, supporting our hypothesis. This indicates that there are differences in the frequency of X between treatments. We now formally examine whether the differences in the observed frequencies in our treatments are statistically significant, by using a non-parametric Wilcoxon ranksum test, and report the corresponding p-values from the test performed for a pairwise comparison. It tests the hypothesis that two independent samples (that is, unmatched data) are from populations with the same distribution.

For Treatment 1 and Treatment 2, the p-value for differences in choosing X was 0.3634. We find that we cannot reject the null and, therefore, it is statistically insignificant and we cannot reject that the two populations have similar distributions, even at a 10% level of significance. As the frequency of X has not differed greatly between the treatments, there is no significant effect from the 'Multiplier Condition' alone.

Result 1 There is no significant difference in the frequency of X between Treatment 1 and Treatment 2; participants do not choose X more or less when the 'Multiplier Condition' holds.

2.6.2 Hypothesis 2: The effect of the 'Multiplier Condition' and Communication

We first broadly look at the effect of communication. To test this hypothesis, we consider the frequency of X in the first and last treatments. Our experiment also includes Treatment 3, which shows the pure effect of communication. Further in our results, we shall analyze this separately. For now, in order to not reject the null, the frequency of X must be lesser in Treatment 4 than in Treatment 1.

The observed frequency of X decreases from 22 individuals to three when the 'Multiplier Condition' holds and the participants are allowed to communicate. This is a reduction by 86%, which supports our hypothesis.

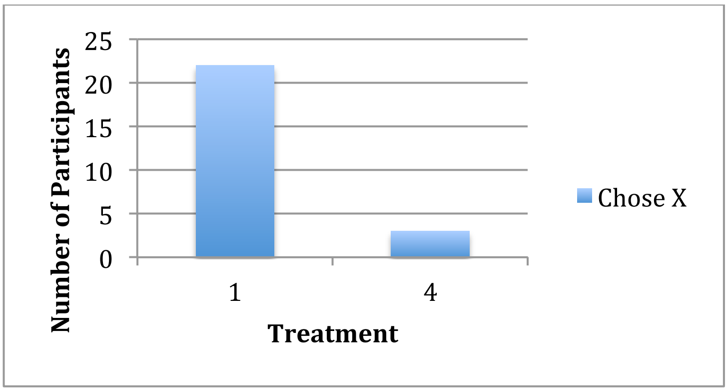


Figure 2.10: Frequency of X in Treatments 1 and 4

The number of groups that chose the outcomes (X,X) and (X,Y) decreased significantly. No groups chose (X,X) in Treatment 4, and only three chose (X,Y). The number of groups that chose (Y,Y) increased over double from 18 groups to 37. All of this is conducive to our second hypothesis.

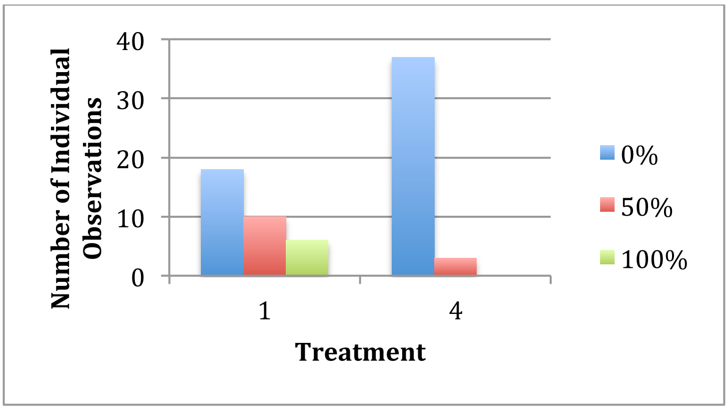


Figure 2.11: Proportion of X within individual observations in Treatments 1 and 4

From Table 2.3, we observe that the mean proportion of choosing X within an individual observation decreases from 55 to 7.5. This means it is almost 86% lower in Treatment 4 than in Treatment 1.

Table 2.3: Mean proportion of X within individual observations in Treatment 1 and 4

	Treatment 1	**Treatment 4**
Mean	55	7.5

We also conducted a non-parametric Wilcoxon ranksum test for Treatment 1 and Treatment 4. The p-value for differences in choosing X was 0.0000. We find that we can reject the null and, therefore, there is a significant difference between the two treatments in this respect, even at a 1% level of significance. This means that Treatment 1 and Treatment 4 do not have similar distributions, the frequency of X differs greatly between the treatments and there is a significant effect from the 'Multiplier Condition' when there is communication.

Result 2 There is a significant difference in the frequency of X between Treatment 1 and Treatment 4; participants choose X less if they can communicate and the 'Multiplier Condition' holds.

2.6.3 Hypothesis 3: The effect of Communication

In their Cheap-Talk model, Baliga and Sjostrom's theorem states that if the 'Multiplier Condition' holds, a higher frequency of people will play 'Not Build'. We make a comparison between the data from Treatment 2 and Treatment 4. In order to not reject the null, the frequency of X must be lesser in Treatment 4 than in Treatment 2.

There is a reduction of 23 participants choosing X; from 26 individuals to three between treatments. This is a reduction by 88%, which supports our hypothesis.

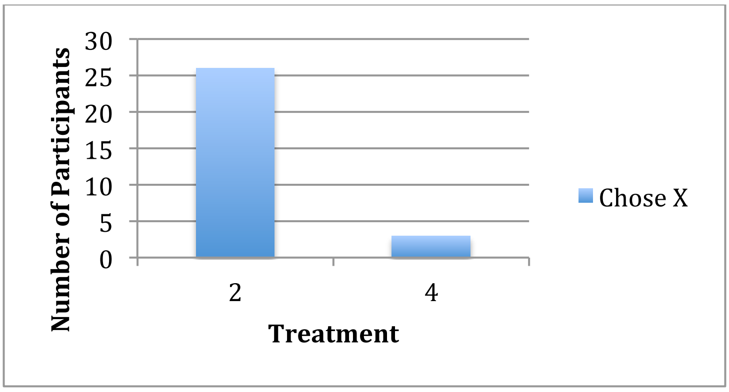


Figure 2.12: Frequency of X in Treatments 2 and 4

Similar to Hypothesis 2, the number of groups that chose the outcomes (X,X) and (X,Y) decreased significantly, whereas the number of groups that chose (Y,Y) increased from 14 groups to 37. All of this is conducive to our third hypothesis.

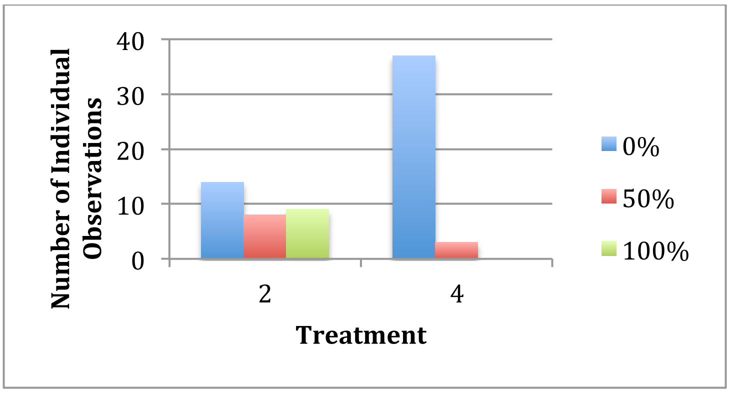


Figure 2.13: Proportion of X within individual observations in Treatments 2 and 4

The mean proportion of choosing X within the individual observations decreased by 88%; from 65 to 7.5. The p-value from the non-parametric Wilcoxon ranksum test was 0.0000. We reject the null and find that there is a significant difference between the two treatments at a 1% level of significance.

Table 2.4: Mean proportion of X within individual observations in Treatment 2 and 4

	Treatment 1	Treatment 4
Mean	65	7.5

This means that Treatment 2 and Treatment 4 do not have similar distributions, the frequency of X differs greatly between the treatments and there is a significant effect from communication when the 'Multiplier Condition' holds.

Result 3 There is a significant difference in the frequency of X between Treatment 2 and Treatment 4; when the 'Multiplier Condition' holds, participants choose X less when they can communicate.

2.6.4 The effect of the 'Multiplier Condition'

Our results have shown that the difference between Treatments 1 and 2 and Treatment 4 is statistically significant. To gain a better understanding of the underlying processes, we can investigate the individual effect of communication by using Treatment 3, particularly as the 'Multiplier Condition' does not give statistically significant results from when it does not hold (See Hypothesis 1). This will enable us to question whether the effect is from communication alone. The exact frequencies of all treatments are presented in Table 2.5 below.

Table 2.5: Frequency of X in all treatments

	Treatment 1	Treatment 2	Treatment 3	Treatment 4
Frequency of X	22	26	3	3

As shown in Table 2.5, there is no difference between Treatment 3 and 4, when comparing the frequency of X. The mean proportions (within individual observations) of all treatments are presented in Table 2.6, and we can see this doesn't differ either. It remains at 7.5 when there is communication, whether the 'Multiplier Condition' holds or not.

Table 2.6: Mean proportion of X within individual observations in all treatments

	Treatment 1	**Treatment 2**	**Treatment 3**	**Treatment 4**
Mean	55	65	7.5	7.5

We performed Wilcoxon ranksum tests for each pair of treatments. The results are given below in Table 2.7.

Table 2.7: p-values from Wilcoxon ranksum tests for all treatments

	Treatment 2	**Treatment 3**	**Treatment 4**
Treatment 1	0.3634	0.0000	0.0000
Treatment 2	-	0.0000	0.0000
Treatment 3	0.0000	-	0.6886

We can reject that the data from Treatment 1 (and 2) and 3 have similar distributions, so there is an effect from communication. We can also reject that the data from Treatments 1 and 2 and 4 have similar distributions, showing that there is an effect from full interaction. However, as the data from Treatment 3 and 4 have similar distributions, and as Result 1 means there is no effect from the 'Multiplier Condition', this indicates that the change in frequency is likely to be caused by communication alone.

2.6.5 Regression Analysis

As the non-parametric statistical test does not control for other characteristics that may affect an individual's decision, we also use a multivariate regression to address our hypotheses. To assess the differences among the treatments, we first run a probit regression using the raw data for each individual observation. We use a probit model as the dependent variable is binary. We will estimate the probability that an observation with particular characteristics will fall into a specific one of the categories. Probit models, which employ a probit link function, are most often estimated using the standard maximum likelihood procedure. This is what we shall use to estimate our regression.

The dependent variable *Choice* is binary; it can have only two possible outcomes, which we will denote as 1 and 0. *Choice* takes a value of 1 if the individual chose X and a value of 0 if the individual chose Y.

In our first regression, Model 1, we include the following independent variables:

1. *Treatment2, Treatment3, Treatment4*: binary variables that takes value 1 when the data belongs to that treatment.
2. *Type*: that takes a value of 1 if the individual is of Type A, a value of 2 if the individual is of Type B and a value of 3 if the individual is of Type C.

We have 20 observations from each treatment, giving a total of 80 overall. We present the coefficients for these variables from the probit regression with robust standard errors clustered on independent matching groups.

Table 2.8: Probit regression results for Model 1

No. of Obs.: 160; Dependent variable: *Choice* =1 if X and =0 ow.	
Treatment2	0.27 (0.29)
Treatment3	-1.66 *** (0.37)
Treatment4	-1.64 *** (0.37)
Type	-0.51 *** (0.18)
Constant	1.37 *** (0.49)
Pseudo R^2: 0.3092	

There are limited ways in which we can interpret the individual regression coefficients. A positive coefficient means that an increase in the predictor leads to an increase in the predicted probability. So, for example, the coefficient of *Type* is -0.51. This means that a decrease in *Type* (being Type A rather than Type C) could increase the predicted probability of X being chosen. The results also show that the predicted probability of X being chosen decreased from Treatment 1 to Treatment 3, as well as from Treatment 1 to Treatment 4. The predicted probability increased from Treatment 1 to Treatment 2 though, so the probit regression suggests more people would choose X when the 'Multiplier Condition' holds, unless there is any communication, when the number of people is more likely to fall.

We use the probability the z test statistic would be observed under the null hypothesis that a particular predictor's regression coefficient is zero, given that the rest of the predictors are in the model. For a given alpha level, it determines whether or not the null hypothesis can be rejected. If it is less than alpha, then the null hypothesis can be rejected and the parameter estimate is considered statistically significant at that alpha level. We use, following standard notations, * to denote statistical significance at the 10% level, ** at the 5% level and *** at the 1% level. For example, with *Treatment2*, when we set our alpha level to 0.05, we fail to reject the null hypothesis and we can conclude that the regression coefficient for Treatment 2 has not been found to be statistically different from zero, given the other regressors are in the model.

In Table 2.8, we are mainly interested in the *Treatment* variable. As *Treatment2* does not seem to have an effect on the predicted probability of the frequency of X, we can conclude that Result 1 still holds. The table clearly indicates significant differences between Treatments 3 and 4 with the baseline. As there is a negative coefficient, the regression above shows that the frequency of X significantly decreases when communication is enabled. Hypothesis 2 has not been rejected; we find supporting evidence for the hypothesis that the frequency of X increases when they can communicate and the 'Multiplier Condition' holds.

As *Type* increases, that is, as a participant becomes a less aggressive type, the predicted probability of the frequency of X falls. This variable is significant, even at a 1% level.

Table 2.9: Probit regression results for Model 2, 3 and 4

No. of Obs.: 160; Dependent variable: *Choice* =1 if X and =0 ow.			
	Model 2	Model 3	Model 4
Treatment2	-1.01 (1.34)	0.08 (0.35)	-1.35 (1.45)
Treatment3	-2.82 * (1.44)	-1.45 *** (0.49)	-3.69 ** (1.61)
Treatment4	-3.02 ** (1.46)	-1.40 *** (0.49)	-3.93 ** (1.61)
Type	-1.09 *** (0.38)	-0.6 *** (0.23)	-1.15 *** (0.40)
T2	0.54 (0.52)	-	0.58 (0.56)
T3	0.70 (0.61)	-	0.95 (0.66)
T4	0.85 (0.60)	-	1.09 * (0.65)
Constant	2.05** (0.94)	0.10 (0.71)	2.47 ** (1.12)
	Pseudo R^2: 0.4298	Pseudo R^2: 0.4670	Pseudo R^2: 4852

We then conducted another regression, which included interactive terms. These variables deal with the non-independence of the types and the treatments. The relationship between the participants' types and the frequency of X, changes depending on the treatment they were in. For example, according to our model, a Type C ("normal") participant should play Y in Treatment 1, whereas they should play X in Treatment 2.

We now include the following additional independent variables:

1. *T*2: Treatment2*Type
2. *T3*: Treatment3*Type
3. *T4*: Treatment4*Type

We present the results of the probit regression of differences in choosing X for each treatment with interactive terms in Table 2.9, labeling it Model 2.

Treatment3, *Treatment4*, and *Type* are still significant, however, none of the interactive terms appear to be significant. Statistically, the treatment variables and interactive terms are correlated. It is possible that the treatment variable may be picking up all the effects of the interactive terms.

We then conducted further regressions to include control variables. We included the following:

1. *Gender*: a binary variable that takes value 1 when the participant is female.
2. *Studies*: a binary variable that takes value 1 when the participant has studied Economics previously, and may therefore be familiar with the game.
3. *Religion*: a binary variable that takes value 1 when the participant reports to have any religious inclination.
4. *Instruction*s*:* that takes a value of 0 if the individual found the instructions 'Very Unclear', a value of 1 if the individual found them 'Unclear', a value of 2 if the individual found them 'Neutral', a value of 3 if the individual found them 'Clear' and a value of 4 if they found the instructions 'Very Clear'. No participants in this experiment reported that they had found the instructions 'Unclear' or 'Very Unclear'.
5. *Belief*: a binary variable that takes value 1 when the participant believed their opponent would play X. When participants elicited their beliefs, we checked if they thought the type they were actually matched with, would play X.

In Table 2.9, the results of the probit regression of differences in choosing X for each treatment with controls, is labeled Model 3. There does not appear to be any obvious changes by introducing the controls, without any interactive terms. However, looking at Model 4, the results of the probit regression of differences in choosing X for each treatment with interactive terms and controls, *Treatment4* gets statistically stronger and *T4* also becomes significant at a 10% level. This supports our second result, as the type of a participant and the frequency of X is affected by whether they participated in Treatment 4, in which there was a full interaction.

Gender also appeared significant at a 10% level with a negative probability, suggesting that males were more likely to choose X. *Religion* is significant at a 5% level and had a positive coefficient. This means that people who did not identify themselves with a religion behaved less aggressively. *Belief* was statistically significant at a 1% level with a positive coefficient, meaning that if a participant believed their opponent would play X, they would choose X too. *Studies* also had a negative coefficient, as did *Instructions* when the interactive terms were included, but neither of these regressors had statistical significance at an acceptable level of significance.[4]

2.7 Conclusion

We conducted a set of experiments to show how an arms race scenario can be triggered and the role of communication in reducing the possibility of such a race. We used two sets of parameters, based on the Baliga and Sjostrom paper, to model when the 'Multiplier Condition' didn't hold and when it did hold. We found under both treatments, a high frequency of subjects chose to build, indicating the possibility of an arms race spiral regardless of whether the Condition held. We then showed how non-binding communication significantly lowers the probability of an arms race. Participants seemed to have created credible information about their opponents' types via communication.

---

[4] A table showing these regression outputs is included in the Appendix.

Thus, the empirical results supported Hypotheses 2 and 3 but not Hypothesis 1. There was no significant difference in the frequency of X between Treatment 1 and Treatment 2, meaning that participants do not choose X more when the ‘Multiplier Condition’ holds. It is possible that the difference between the treatments is not statistically significant because the results from the ‘Multiplier Condition’ requires a high order iteration of beliefs that the participants of the experiment have not been able to do. For the ‘Multiplier Condition’ to hold, a player needed beliefs and beliefs over beliefs, etc. due to the lack of common knowledge available on each other’s types (as in, Rubenstein, 1989). They also need common knowledge of beliefs. All rational players will believe dominant strategy types will play X, and they will also believe that the “almost” dominant strategy types will play X if they know the dominant strategy types will do so. The contagion takes hold as they have the same beliefs of the “almost-almost” dominant strategy types, and so on. Moreover, it should be noted that there is a Bayesian Nash Equilibrium where all types choose to acquire weapons, without the ‘Multiplier Condition’ holding. Thus, it could simply be the case that subjects ended up playing the equilibrium in which they all build (play X) even when there was another equilibrium.

There was a significant difference in the frequency of X between Treatments 1 and 2 and Treatment 4, showing participants choose X more both when the ‘Multiplier Condition’ holds and it does not and they can communicate. This implies that cheap-talk can enable players to coordinate on a Pareto dominant equilibrium, providing the probability of facing an “aggressive” type is sufficiently small.

As the data from Treatment 3 and 4 have similar distributions, and as Result 1 means there is no effect from the ‘Multiplier Condition’, this indicates that the change in frequency likely caused by communication alone. This could indicate that the decision to play Y is predominantly due to the participants being able to engage in cheap-talk.

Our results suggest the positive role of communication in reaching Pareto superior outcomes even in the presence of incomplete information while casting doubt on the ability of subjects to do higher order iterations.

There is a possibility for future extensions. We could investigate the effect of repeating the game by having multiple rounds of each game, for all treatments. We could also look at the impact of different types of communication, in particular whether fixed-form messages would lead to different outcomes. This would mean a player would only have the option of sending two messages, for example, "I WILL CHOOSE OPTION X" or "I WILL CHOOSE OPTION Y." Although previous experimental work suggests that the impact of communication varies with the type of messages, the results are yet far from conclusive (Oprea et al., 2014). In some designs, communication has been more structured and players make simultaneous or one-sided announcements (Crawford, 1997). These include Clark et al. (2001), Cooper et al. (1992) and Palfrey and Rosenthal (1991). In other designs, the communication's timing and content is unstructured, such as Forsythe et al. (1991), Valley et al. (1996) and Roth (1987). In studies of equilibrium selection, such as coordination games, simple forms of anonymous communication have been shown to be effective. Cooper et al. (1996) show simple pre-play fixed messages are able to cause players to choose the efficient, payoff-dominant equilibrium. On the other hand, Ostrom et al. (1992) find face-to-face communication causes an increase in cooperation in a public-goods game. It would be interesting to compare how our results vary.

2.8 Appendix A

Instructions without Communication (Treatment 1)

Preliminary Instructions

Welcome to the Birmingham Experimental Economics Laboratory. This is an experiment in the economics of decision-making. The funding for this research has been provided by various research foundations. Just for showing up, you have already earned a show up fee of £2.50. You can earn additional money depending on the decisions made by you and other participants. It is very important that you read the instructions carefully.

It is important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate your following of these rules.

We will first go over the instructions together. After we have read the instructions, you will have time to ask clarifying questions. We would like to stress that any decisions you make in this experiment are completely anonymous. Please do not touch the computer or its mouse until you are instructed to do so. Thank you.

The experiment consists of one round. You will be paid according to the number of points you have collected in that round with the following exchange rate:

1 point = £0.50p

You will be paid privately in cash at the end of the experiment.

Instructions

Type and Matching

In this experiment, you will be randomly assigned a type by the computer, either 'Type A', 'Type B' or 'Type C'. **Your type will remain the same throughout the whole experiment.** In total there are 20 participants in this room; 10% will be assigned the role of **Type A**, 40% will be assigned the role of **Type B** and the remaining 50% of the participants will be assigned the role of **Type C**. You will be informed of your type but not of the other participants' types.

At the beginning of the round, you will be matched with one other person, randomly selected from the participants in this room, to form a group of two. You will therefore be in a group with one other participant. You will never learn the identity of the other participant you are matched with, nor will the other participants be informed about your identity.

The Decision Stage

You and the other participant you are matched with have to **simultaneously select one** of two options, either OPTION X or OPTION Y. When you make your own choice, you do not know what the other participant's choice is. Your type and the decisions made by both of you will determine the points you each get, as shown in the following tables. The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type A:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	11, 11	16, 3
	OPTION Y	3, 16	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	11, 8.9	16, 3
	OPTION Y	3, 13.9	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	11, 5	16, 3
	OPTION Y	3, 10	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type B:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	8.9, 11	13.9, 3
	OPTION Y	3, 16	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	8.9, 8.9	13.9, 3
	OPTION Y	3, 13.9	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	8.9, 5	13.9, 3
	OPTION Y	3, 10	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type C:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5, 11	10, 3
	OPTION Y	3, 16	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5, 8.9	10, 3
	OPTION Y	3, 13.9	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

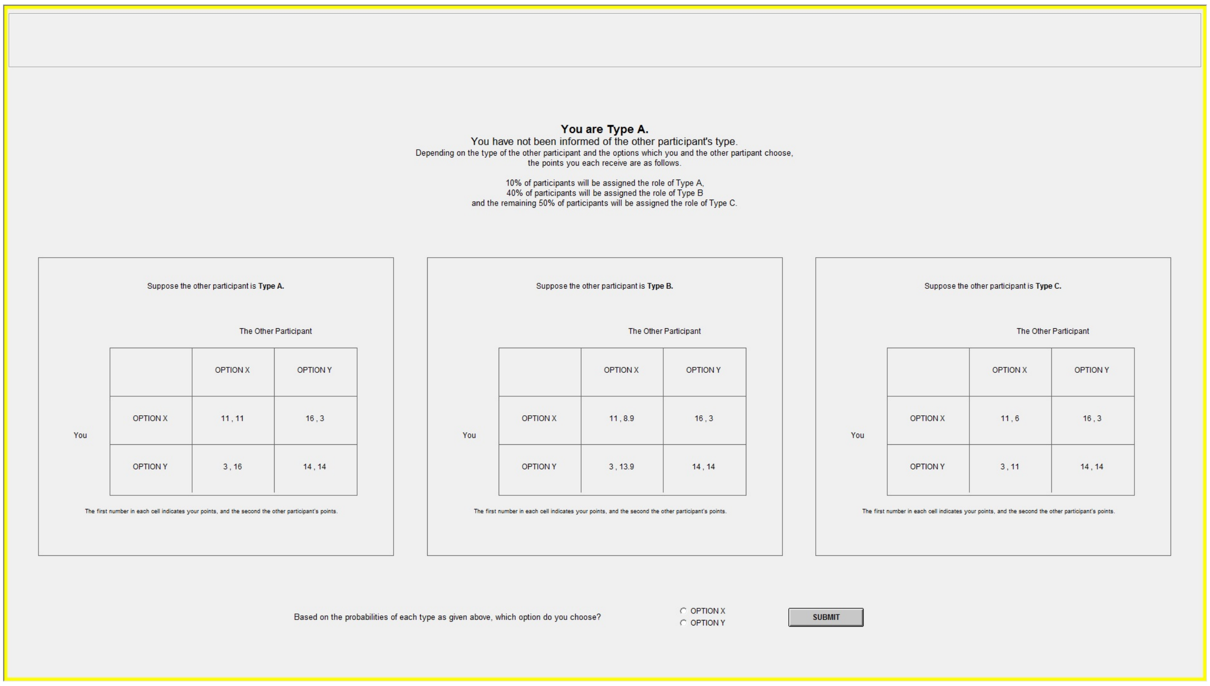
		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5, 5	10, 3
	OPTION Y	3, 10	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

For example, suppose that you are Type A and the other participant you are matched with is Type B. If you choose OPTION X and they choose OPTION Y, then you will earn 16 points and the other participant will earn 3 points.

The Decision Screen

You will enter your decisions on a screen that looks as follows. The example below is used for illustrative purposes. This screenshot corresponds to the case where you are Type A. In the actual experiment, you can be Type A, Type B or Type C as mentioned earlier in the instructions.



You are Type A.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B
and the remaining 50% of participants will be assigned the role of Type C.

Suppose the other participant is **Type A.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	11 , 11	16 , 3
	OPTION Y	3 , 16	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is **Type B.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	11 , 8.9	16 , 3
	OPTION Y	3 , 13.9	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is **Type C.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	11 , 6	16 , 3
	OPTION Y	3 , 11	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

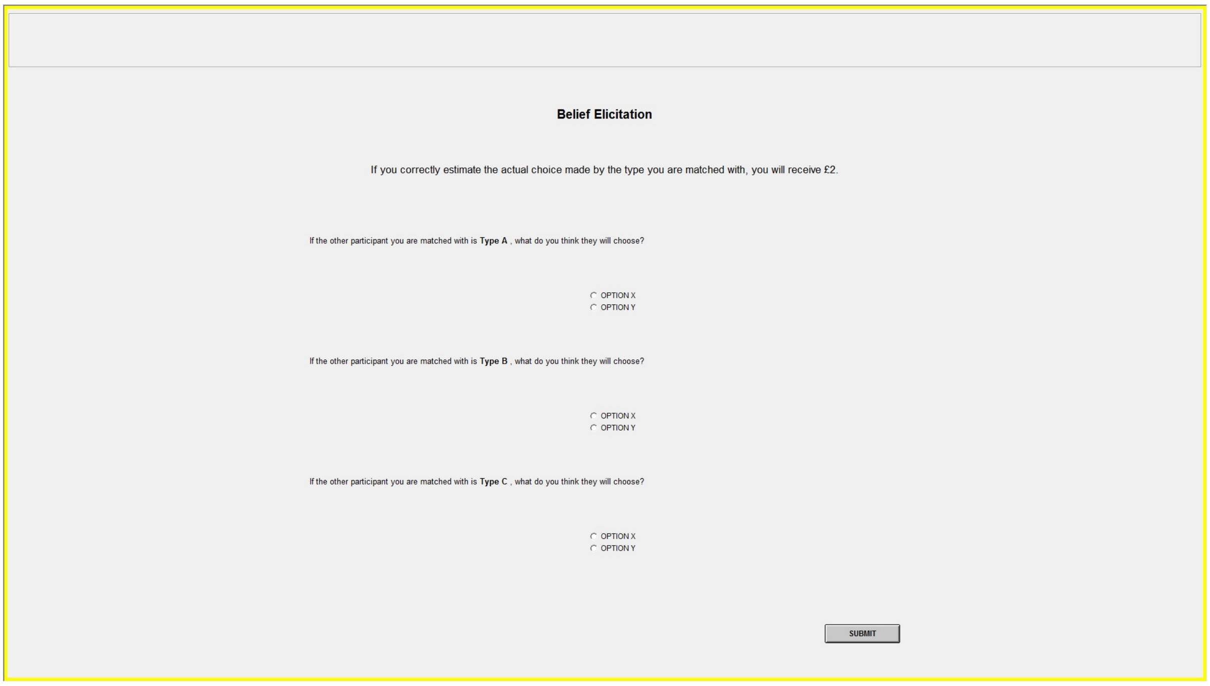
Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

As mentioned before, before you make your decision, you will be informed of your Type and the payoff tables for facing each type. You must select **one** of two options, either OPTION X or OPTION Y. To make your choice, you simply select the appropriate button and then click on the ‘Submit’ button by using your mouse. Once you have pressed the ‘Submit’ button, you can no longer revise your decision.

Belief Elicitation

On the next screen, you are asked to indicate which of the two options (OPTION X or OPTION Y) each type of the other person you could be matched with will choose. You will make your choices on a screen that looks as follows.



Belief Elicitation

If you correctly estimate the actual choice made by the type you are matched with, you will receive £2.

If the other participant you are matched with is **Type A** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If the other participant you are matched with is **Type B** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If the other participant you are matched with is **Type C** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If you correctly estimate the actual choice made by the participant you are matched with, you will receive £2.

Ending of the Experiment

When both you and the other participant you are matched with have made your decisions, you will see a final screen reporting your type and your choice, the other participant's type and their choice, the total points you have earned, the total points you have earned in £, your earnings from the belief elicitation, your show up fee, plus the corresponding final £ payment. You will be paid £0.50p for every point and £2 if you estimate the other participant's choice correctly, in addition to your £2.50 show-up fee.

Do you have any questions? Please do not ask any questions out loud. If you have any questions, please raise your hand and an experimenter will come to your desk.

Control Questions

To make sure everyone understands the instructions, please complete the questions below. In a couple of minutes someone will come to your desk to check your answers. Once everybody answers the following questions correctly, the experiment will start. (The decisions and earnings used for the questions below are simply for illustrative purposes. In the experiment decisions and earnings will depend on the actual choices of the participants.)

1. Will your type change or remain the same throughout the experiment? (Please circle the correct answer).

SAME / CHANGE

2. How many 'Type A' participants are there in this experiment?
…………

3. If you are a 'Type A' participant, how many points would you receive by choosing OPTION X when matched with a 'Type B' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

4. If you are a 'Type B' participant, how many points would you receive by choosing OPTION Y when matched with a 'Type C' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

5. If you are a 'Type C' participant, how many points would you receive by choosing OPTION X when matched with a 'Type A' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

2.9 Appendix B

Instructions with Communication (Treatment 3)

Preliminary Instructions

Welcome to the Birmingham Experimental Economics Laboratory. This is an experiment in the economics of decision-making. The funding for this research has been provided by various research foundations. Just for showing up, you have already earned a show up fee of £2.50. You can earn additional money depending on the decisions made by you and other participants. It is very important that you read the instructions carefully.

It is important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate your following of these rules.

We will first go over the instructions together. After we have read the instructions, you will have time to ask clarifying questions. We would like to stress that any decisions you make in this experiment are completely anonymous. Please do not touch the computer or its mouse until you are instructed to do so. Thank you.

The experiment consists of one round. You will be paid according to the number of points you have collected in that round with the following exchange rate:

1 point = £0.50p

You will be paid privately in cash at the end of the experiment.

Instructions

Type and Matching

In this experiment, you will be randomly assigned a type by the computer, either 'Type A', 'Type B' or 'Type C'. **Your type will remain the same throughout the whole experiment.** In total there are 20 participants in this room; 10% will be assigned the role of **Type A**, 40% will be assigned the role of **Type B** and the remaining 50% of the participants will be assigned the role of **Type C**. You will be informed of your type but not of the other participants' types.

At the beginning of the round, you will be matched with one other person, randomly selected from the participants in this room, to form a group of two. You will therefore be in a group with one other participant. You will never learn the identity of the other participant you are matched with, nor will the other participants be informed about your identity.

The round consists of **two stages:** i) the Chat stage and ii) the Decision stage.

i) The Chat Stage

You are given 180 seconds (3 minutes) in which you can use a chat program to write messages to the participant you are matched with.

When you use the chat program, you can type whatever you want in the lower box of the chat program, except for the following restrictions.

Restrictions on messages:

1. You **must not identify yourself** or send any information that could be used to identify you (for example, your name, contact details or seat in the room)
2. You **must not make any** threats, insults or use any obscene or offensive language.

If you violate these rules your payment will be forfeited.

Messages will be shared only between you and the participant you are matched with. You will not be able to see the messages of the other groups and the other groups will not see the messages in your chat.

Once 180 seconds have elapsed, we will proceed to the ‘Decision Stage’.

ii) The Decision Stage

You and the other participant you are matched with have to **simultaneously select one** of two options, either OPTION X or OPTION Y. When you make your own choice, you do not know what the other participant’s choice is. Your type and the decisions made by both of you will determine the points you each get, as shown in the following tables. The first number in each cell indicates your points, and the second the other participant’s points.

Payoff tables for Type A:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	11, 11	16, 3
	OPTION Y	3, 16	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	11, 8.9	16, 3
	OPTION Y	3, 13.9	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	11, 5	16, 3
	OPTION Y	3, 10	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type B:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	8.9, 11	13.9, 3
	OPTION Y	3, 16	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	8.9, 8.9	13.9, 3
	OPTION Y	3, 13.9	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	8.9, 5	13.9, 3
	OPTION Y	3, 10	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type C:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5, 11	10, 3
	OPTION Y	3, 16	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5, 8.9	10, 3
	OPTION Y	3, 13.9	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

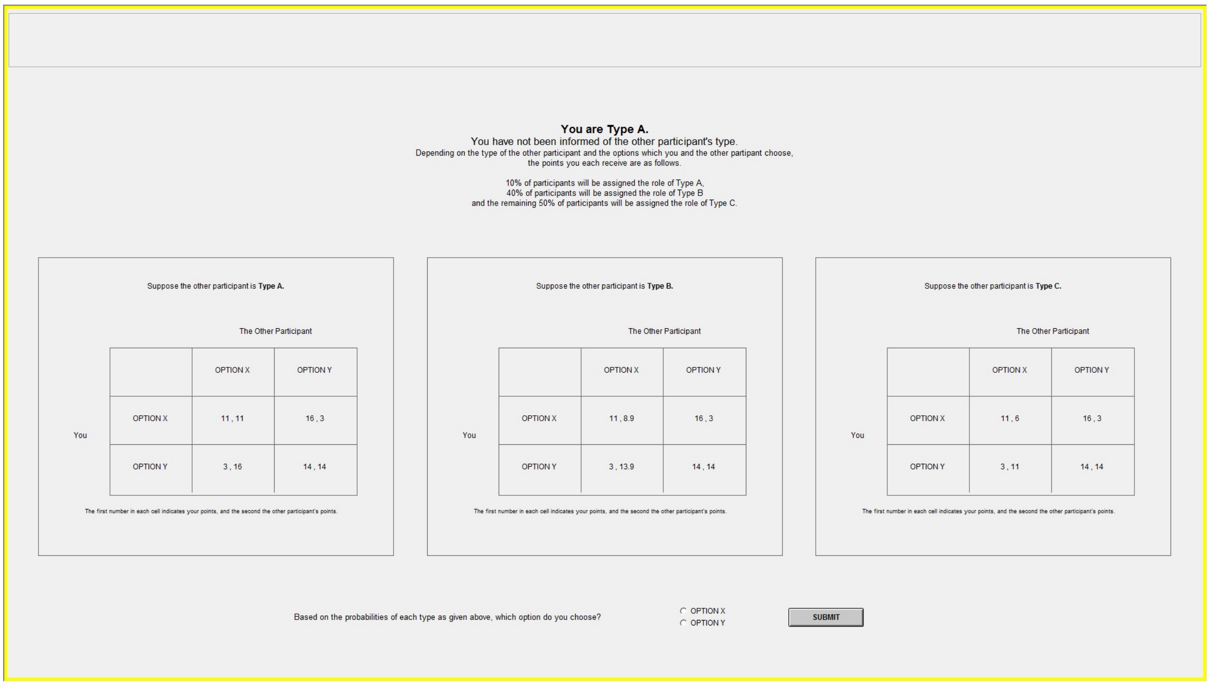
		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5, 5	10, 3
	OPTION Y	3, 10	14, 14

The first number in each cell indicates your points, and the second the other participant's points.

For example, suppose that you are Type A and the other participant you are matched with is Type B. If you choose OPTION X and they choose OPTION Y, then you will earn 16 points and the other participant will earn 3 points.

The Decision Screen

You will enter your decisions on a screen that looks as follows. The example below is used for illustrative purposes. This screenshot corresponds to the case where you are Type A. In the actual experiment, you can be Type A, Type B or Type C as mentioned earlier in the instructions.



You are Type A.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B
and the remaining 50% of participants will be assigned the role of Type C.

Suppose the other participant is **Type A.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	11 , 11	16 , 3
	OPTION Y	3 , 16	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is **Type B.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	11 , 8.9	16 , 3
	OPTION Y	3 , 13.9	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is **Type C.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	11 , 6	16 , 3
	OPTION Y	3 , 11	14 , 14

The first number in each cell indicates your points, and the second the other participant's points.

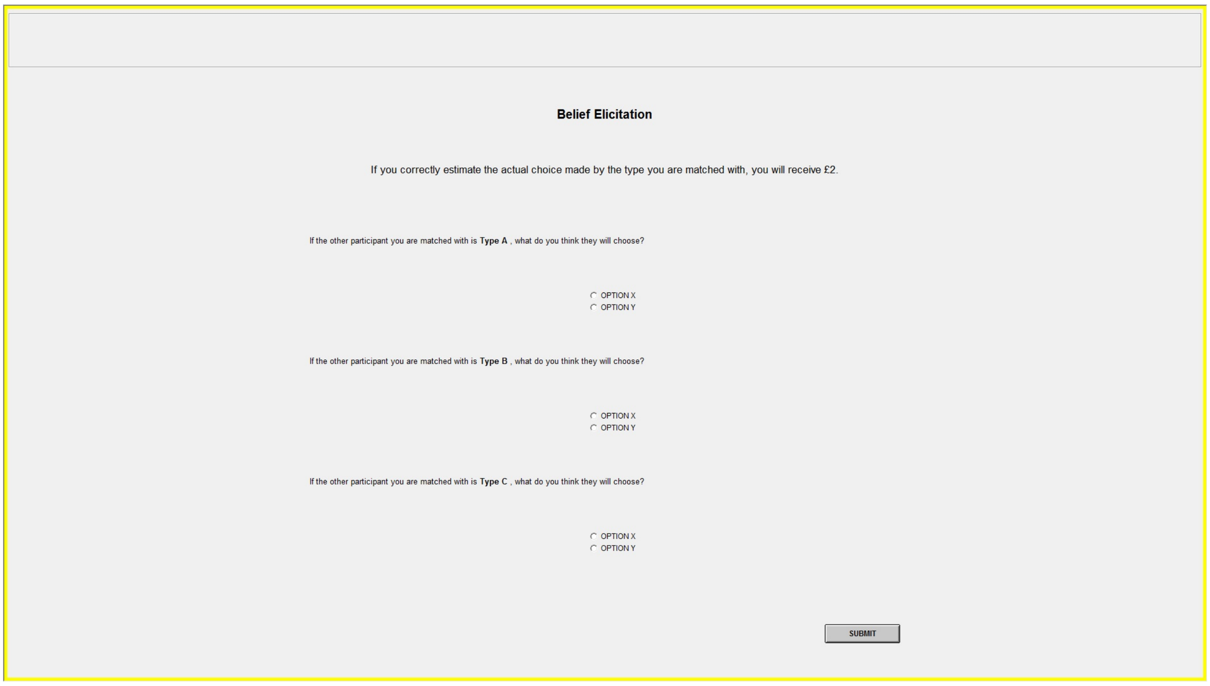
Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

As mentioned before, before you make your decision, you will be informed of your Type and the payoff tables for facing each type. You must select **one** of two options, either OPTION X or OPTION Y. To make your choice, you simply select the appropriate button and then click on the ‘Submit’ button by using your mouse. Once you have pressed the ‘Submit’ button, you can no longer revise your decision.

Belief Elicitation

On the next screen, you are asked to indicate which of the two options (OPTION X or OPTION Y) each type of the other person you could be matched with will choose. You will make your choices on a screen that looks as follows.



Belief Elicitation

If you correctly estimate the actual choice made by the type you are matched with, you will receive £2.

If the other participant you are matched with is **Type A** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If the other participant you are matched with is **Type B** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If the other participant you are matched with is **Type C** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If you correctly estimate the actual choice made by the participant you are matched with, you will receive £2.

Ending of the Experiment

When both you and the other participant you are matched with have made your decisions, you will see a final screen reporting your type and your choice, the other participant's type and their choice, the total points you have earned, the total points you have earned in £, your earnings from the belief elicitation, your show up fee, plus the corresponding final £ payment. You will be paid £0.50p for every point and £2 if you estimate the other participant's choice correctly, in addition to your £2.50 show-up fee.

Do you have any questions? Please do not ask any questions out loud. If you have any questions, please raise your hand and an experimenter will come to your desk.

Control Questions

To make sure everyone understands the instructions, please complete the questions below. In a couple of minutes someone will come to your desk to check your answers. Once everybody answers the following questions correctly, the experiment will start. (The decisions and earnings used for the questions below are simply for illustrative purposes. In the experiment decisions and earnings will depend on the actual choices of the participants.)

1. Will your type change or remain the same throughout the experiment? (Please circle the correct answer).

SAME / CHANGE

2. How many 'Type A' participants are there in this experiment?
…………

3. If you are a 'Type A' participant, how many points would you receive by choosing OPTION X when matched with a 'Type B' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

4. If you are a 'Type B' participant, how many points would you receive by choosing OPTION Y when matched with a 'Type C' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

5. If you are a 'Type C' participant, how many points would you receive by choosing OPTION X when matched with a 'Type A' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

2.10 Appendix C

Payoff Matrices in the Experiment when the 'Multiplier Condition' Does Not Hold

Payoff Tables for Type A

Type A

		Option X	Option Y
Type A	Option X	11, 11	16, 3
	Option Y	3, 16	14, 14

Type B

		Option X	Option Y
Type A	Option X	11, 8.9	16, 3
	Option Y	3, 13.9	14, 14

Type C

		Option X	Option Y
Type A	Option X	11, 5	16, 3
	Option Y	3, 10	14, 14

Payoff Tables for Type B

Type A

		Option X	Option Y
Type B	Option X	8.9, 11	13.9, 3
	Option Y	3, 16	14, 14

Type B

		Option X	Option Y
Type B	Option X	8.9, 8.9	13.9, 3
	Option Y	3, 13.9	14, 14

Type C

		Option X	Option Y
Type B	Option X	8.9, 5	13.9, 3
	Option Y	3, 10	14, 14

Payoff Tables for Type C

Type A

		Option X	Option Y
Type C	Option X	5, 11	10, 3
	Option Y	3, 16	14, 14

Type B

		Option X	Option Y
Type C	Option X	5, 8.9	10, 3
	Option Y	3, 13.9	14, 14

Type C

		Option X	Option Y
Type C	Option X	5, 5	10, 3
	Option Y	3, 10	14, 14

2.11 Appendix D

Payoff Matrices in the Experiment when the 'Multiplier Condition' Holds

Payoff Tables for Type A

		Type A	
		Option X	Option Y
Type A	Option X	11, 11	16, 3
	Option Y	3, 16	14, 14

		Type B	
		Option X	Option Y
Type A	Option X	11, 8.9	16, 3
	Option Y	3, 13.9	14, 14

		Type C	
		Option X	Option Y
Type A	Option X	11, 6	16, 3
	Option Y	3, 11	14, 14

Payoff Tables for Type B

Type A

		Option X	Option Y
Type B	Option X	8.9, 11	13.9, 3
	Option Y	3, 16	14, 14

Type B

		Option X	Option Y
Type B	Option X	8.9, 8.9	13.9, 3
	Option Y	3, 13.9	14, 14

Type C

		Option X	Option Y
Type B	Option X	8.9, 6	13.9, 3
	Option Y	3, 11	14, 14

Payoff Tables for Type C

Type A

		Option X	Option Y
Type C	Option X	6, 11	11, 3
	Option Y	3, 16	14, 14

Type B

		Option X	Option Y
Type C	Option X	6, 8.9	11, 3
	Option Y	3, 13.9	14, 14

Type C

		Option X	Option Y
Type C	Option X	6, 6	11, 3
	Option Y	3, 11	14, 14

2.12 Appendix E

Questionnaire

Note: This questionnaire was used for all experiments.

Questionnaire

Please answer each of the following questions as accurately as possible. Naturally your responses will be completely confidential. Your answers will be of immense value for our scientific investigation. Thank you in advance for your cooperation.

What is your gender?

○ Female
○ Male

How old are you?

[ ]

What is the degree programme you are registered for?

○ Natural Sciences
○ Engineering or Computer Science
○ Medical Science
○ Law
○ Humanities
○ Economics or Business
○ Social Sciences (other than Economics/Business)

What is your nationality?

[ ]

What is your marital status?

○ Single
○ Married
○ Separated

What is your religion?

○ Protestant
○ Catholic
○ Jewish
○ Muslim
○ Hindu
○ Other Religions
○ No religion

What is the highest level of education your father has received?

○ Primary school

○ Secondary school
○ Some university
○ Undergraduate degree
○ Postgraduate degree
○ Not known

What is the highest level of education your mother has received?

○ Primary school
○ Secondary school
○ Some university
○ Undergraduate degree
○ Postgraduate degree
○ Not known

Please indicate your political attitude in the following scale

Left • • • • • • • • • Right

How many participants of the experiment do you know by name?

[ ]

How did you find the instructions?

○ Very clear
○ Clear
○ Neutral
○ Unclear
○ Very unclear

How important was it for you to maximise your own income?

○ Very important
○ Important
○ Indifferent
○ Not important
○ Not important at all

Thank you very much for participating in the experiment!

2.13 Appendix F

Regression Output Table

No. of Obs.: 160; Dependent variable: *Choice* =1 if X and =0 ow.		
	Model 3	Model 4
Treatment2	0.08 (0.35)	-1.35 (1.45)
Treatment3	-1.45 *** (0.49)	-3.69 ** (1.61)
Treatment4	-1.40 *** (0.49)	-3.93 ** (1.61)
Type	-0.6 *** (0.23)	-1.15 *** (0.40)
T2	-	0.58 (0.56)
T3	-	0.95 (0.66)
T4	-	1.09 * (0.65)
Gender	-0.53 * (0.31)	-0.61 * (0.32)
Studies	-0.47 (0.37)	-0.52 (0.37)
Religion	0.80 ** (0.33)	0.79 ** (0.33)
Instructions	-0.07 (0.24)	-0.12 (0.25)
Belief	1.74 *** (0.36)	1.77 *** (0.37)
Constant	0.10 (0.71)	2.47 ** (1.12)
	Pseudo R^2: 0.4670	Pseudo R^2: 4852

Chapter 3

Provoking conflict: Do Hawkish recommendations make players more aggressive?

3.1 Introduction

We experimentally look at the effect of third party recommendations in a ‘conflict game’ adapted from Baliga and Sjostrom (2012) to see if such recommendations can increase the probability of conflict. Unlike our earlier set of experiments, which also looked at the role of communication by the players themselves, we consider the role of a third player sending a message that both players can observe. Baliga and Sjostrom (BS, 2012) look at third party messages and shows they can increase or decrease conflict depending on whether the actions of the two players are strategic complements or substitutes. Our experiments deal with the case of strategic complements but we will briefly describe both cases.

Similar to the previous chapter, the model is based on an augmented stag-hunt game with incomplete information. There are two countries (A and B) that choose D (peaceful strategy) or H (aggressive strategy). Types of players (“dominant strategy Doves”, “dominant strategy Hawks” or “moderate”) differ and are differentiated by their costs of aggression, which is private information. The normal form of the game is as below.

	Hawk (H)	Dove (D)
Hawk (H)	$-c_i$	μ-c_i
Dove (D)	-d	0

Figure 3.1: Payoff matrix of Player *i* in Baliga and Sjostrom’s (2012) model

BS (2012) allow actions to be Strategic Complements or Strategic Substitutes. Players can be “dominant strategy” Hawks or Doves in which case their best response is to play H and D respectively or they can be a “moderate”. A “moderate” is a player whose best response depends on his beliefs about the opponent’s action. In the Strategic Complements game, they would be a “Coordination” type; H’s Best Response is H and D’s Best Response is D. In a game where actions are Strategic Substitutes, “moderates” would be an “Opportunist” type; D’s best response is H and H’s best response is D.

BS show that under ‘mild’ assumptions on the distribution of types, there is a unique Communication-Free Equilibrium for both players. If their cost is below a threshold, they fight, and otherwise they do not.

The paper then looks at communication by third parties. A third party called E has perfect information about Player A’s true preferences (he knows Player A’s type) and wants to influence Player A’s action by signaling Player A’s preferences (type) to Player B. The model assumes Player E’s true preferences are commonly known, and Player E can only communicate by sending a publicly observed cheap-talk message before Player A and Player B simultaneously make their decisions. The model assumes Player E can only communicate, and do nothing else, in order to analyse the impact of third party talk on the manipulation of conflict.

There are two cases within their cheap-talk model. Case 1 is where Player E is a Hawk (E_H) and wants Player A to play H. Case 2 is where Player E is a Dove (E_D) and wants Player A to play D. In both cases, Player E always wants Player B to play D.

The paper finds a unique Communication Equilibrium; an equilibrium where Player E’s cheap-talk influences Player A’s and Player B’s decisions. As it is commonly known what E wants the two players to do, Player A will only listen to Player E’s message if he thinks it might influence Player B.

For the Communication Equilibrium, BS (2012) looks at which case (Case 1 with Player E_H or Case 2 Player E_D) creates effective cheap-talk in the Strategic

Complements game and the Strategic Substitutes game. Without loss of generality, they assume two messages m_1 (Hawkish) and m_0 (peaceful) can be sent.

In the Strategic Complements game, E_H will send m_1 only if A is more likely to play H as a result; they wish to indirectly influence Player A by publicly provoking Player B, m_1 will make Player B likely to play H as it signals information about player A's type. As actions are strategic complements, it makes Type A play H with a higher probability as well. Therefore, E_H can communicate effectively via a Hawkish message by increasing tension and inducing a fear spiral. In equilibrium, E_H sends m_1 when Player A is "weak moderate" (a "Coordination" type) i.e. Player A chooses H if provocation makes Player B choose H, and Player A chooses D without cheap-talk. Otherwise, E_H sends m_o i.e. in a situation when Player A's action does not depend on the message. This is because it increases the probability that Player B plays D. Player A's action will not depend on the message if their costs are below a threshold and therefore, they are "dominant strategy" types. It is worth noting that eliminating E_H makes **everyone** strictly better off because both want their opponents to play D.

What happens when the Recommender is $E_{D,}$ who will want both players to choose D: In the Strategic Complements game, he must always send m_o in equilibrium. However, always sending the same message is not informative, thus the outcome is equivalent to the unique Communication-Free Equilibrium and cheap-talk is ineffective.

However, in the Strategic Substitutes game, E_D will send m_1 only if it makes Player A more likely to play D, whilst m_1 will also make Player B more likely to play H. As with the Strategic Complements game, Player A is indirectly influenced by the message; when Player B becomes more Hawkish, Player A backs down and plays D. E_D can communicate effectively via a Hawkish message by reducing tension and causing Player A to back down. In equilibrium, E_D sends m_1 when A is a "tough moderate" (an "Opportunist" type) i.e. Player A chooses D if he fears provocation makes Player B choose H, (note Player A chooses H without cheap-talk). Otherwise, E_D sends m_o. As Player A becomes more Dovish, Player B becomes better off.

Thus, m_1, which makes Player B more likely to play H, now makes Player A choose D, in the Strategic Substitutes game.
Now consider E_H. He will always want to send m_0 as he always prefers Player B to play D and Player A to play H. Therefore, he must always send m_0. As in the Strategic Complements case, a 'constant message' conveys no information and cheap-talk is ineffective here too.

The provocative recommender could represent a number of real world phenomena. Firstly, agents with extreme agendas sometimes take provocative actions that inflame conflicts. Ariel Sharon's symbolic visit to the Temple Mount in September 2000 helped spark the Second Intifada and derailed the Israeli-Palestinian peace process (Hefetz and Bloom, 2006). ISI-sponsored militants seem to deliberately inflame the conflict between Pakistan and India, partly because India is seen as an enemy, but also because the conflict relieves the pressure on extremists supported by the ISI. For Pakistani and Indian leaders, a Hawkish stance may be the best response, given the (correct) belief that their opponent will become more aggressive. This could include ideologues, such as lobby groups or partisan media, signaling intentions of their leaders. An example of these is Abu Hamza al-Masri, who was extradited from the UK to the US to face terrorism charges, after being charged by British authorities for preaching Islamic fundamentalism and militant Islamism. In the USA, Steve Bannon has been accused of being a right-wing ideologue wanting to instigate war not only with the Islamic State and China but also internally within America, saying to the conservative Liberty Restoration Foundation "this is the fourth great crisis in American history. We had the Revolution. We had the Civil War. We had the Great Depression and World War II. This is the great Fourth Turning in American history". Many of these parties and individuals do actually lie. We isolate the situation where they have incentive to lie but they can't because the information may be verifiable.

A provocative recommender could represent also aggressive bargaining tactics. In 2000, Ehud Barak and Yasser Arafat had to decide whether to adopt a tough stance *H* or a conciliatory stance *D* in peace negotiations. Furthermore, it might represent choosing a Hawkish or Dovish agent who will take aggressive actions on the decision maker's behalf. The median voters in Israel and Palestine decide whether to support Likud or Kadima, or Hamas or Fatah, respectively.

Our experiments are based on the BS (2012) model but we ran our experiments only for the Strategic Complements game.

We ran a first set of experiments (referred to as ‘Experiment 1’) to test whether the introduction of a nonbinding, pre-play recommendation causes an increase in play of the aggressive strategy. Players did not choose an aggressive strategy with a recommendation. There could have been a number of reasons why this occurred therefore, we conducted a second set of experiments (referred to as ‘Experiment 2’) with modified payoffs. This is described in details Section 3.3. We now briefly describe the experiment.

The experiment consisted of a one-shot stag hunt game, adapted from Baliga and Sjostrom (2012) with a discrete number of types. There with two treatments in each set of experiments; we refer to them as Treatment 1, 2 in both experiments.

Table 3.1: Treatments of Experiment 1 and 2

		No Recommendation	Recommendation
Experiment 1 (First Set of Experiments)	Treatment 1	x	
	Treatment 2		x
Experiment 2 (Second Set of Experiments)	Treatment 1	x	
	Treatment 2		x

Treatment 1 serves as a baseline. Treatment 2 gives the recommendation effect in both experiments.

Subjects were only allowed to participate in one treatment and played with an anonymous opponent. The sessions did not last more than forty minutes each. There were a total of 160 participants, as each treatment included two sessions with 20 participants in each. Treating each game (pair of participants) as an independent observation, this gave 10 independent observations per session and a total of 20 independent observations per treatment. In Treatments 2, a recommendation was announced simultaneously to both players before the players submitted their decisions. Players submitted their decisions simultaneously and then their beliefs about their opponent's behaviour were elicited. The payoff matrices determined their payoffs, the outcomes were disclosed and both were paid, privately.

In the experiment, neutral framing was used to describe the two actions, H and D. In the Section 3.3 (the 'Experiment' section) and thereafter, we refer to them as 'Option X' and 'Option Y', respectively. The model suggests that without a recommendation; "dominant strategy Hawk" and "tough moderate" players will play X, whereas "dominant strategy Dove" and "weak moderate" players will play Y. When a recommendation is made, the model shows that there exists another equilibrium where all players will play X, except "dominant strategy Dove" players who choose Y. Thus, our first hypothesis is that when the recommendation is not made in Treatment 1, we expect more players to choose Y.

We make another hypothesis to verify how supportive our results are of the model. Theory states that the reason more players choose X after a recommendation, is because "weak moderate" players switch from playing Y to play X. Therefore, for our second hypothesis, we check the proportion of "weak moderate" players who play X in each treatment, and expect more to play X in Treatment 2 than in Treatment 1.

We present our analysis in two parts; Part 1 analyses Experiment 1 and Part 2 analyses Experiment 2. Within each part, we give a descriptive analysis; checking whether there are differences in the frequency of X between treatments and formally examining whether or not there are significant differences in the observed frequencies among treatments by using a non-parametric Wilcoxon ranksum test. Then, as the non-parametric statistical test does not control for other characteristics that may affect an individual's decision, we also use a multivariate regression to address our hypotheses. We use a probit model as the dependent variable is binary.

We found that while in some sessions with a recommendation, more players do choose an aggressive strategy compared to the treatment without, there is no significant difference in the frequency of X among Treatment 1 and Treatment 2; participants do not choose X more after a recommendation. This can be for a number of reasons, not least because players can be in a babbling equilibrium i.e. they are choosing to ignore the messages. It also could be that the symmetric payoffs of the Dovish strategy seem like a ‘focal point’ for all but the “dominant strategy” types. Further, this behaviour is consistent with players choosing the socially optimal outcome or being inequality averse.

The remainder of this paper is organized as follows. Section 3.2 explains the model and Section 3.3 describes the experimental design. In Section 3.4, we give our hypotheses. In Section 3.5, we give our results and discuss them and, in Section 3.6, we conclude.

3.2 Model

3.2.1 Baliga and Sjostrom’s Model Without Cheap-Talk

There are two decision makers, Player A and Player B, who simultaneously choose one of two actions, either H or D.

μ is the benefit from being more aggressive than your opponent and d is the cost of being caught out when your opponent is aggressive. Player i $\in$ {A, B} has a privately known cost, c_i, of taking action H. This is their type, which is independently drawn from the same distribution.

Player i can be one of three types; a “dominant strategy Hawk”, a “dominant strategy Dove” or a “Coordination” type (or an “Opportunist” type in the Strategic Substitutes game). For a “dominant strategy Hawk”, H is a dominant strategy and for a “dominant strategy Dove”, D is a dominant strategy. Player i is a “Coordination type if H is a best response to H and D is a best response to D.

A strategy for Player i is a function $\sigma_i : [\underline{c}, \overline{c}] \rightarrow \{H, D\}$ which specifies an action $\sigma_i(c_i) \in \{H, D\}$ for each cost type $c_i \in [\underline{c}, \overline{c}]$. Player i uses a cutoff strategy if there is a cutoff point x such that $\sigma_i(c_i) = H$ if and only if $c_i < x$. In the paper, all Bayesian Nash equilibria turn out to be in cutoff strategies and any such strategy can be identified with its cutoff point. Using this, the paper shows a unique Communication-Free

Equilibrium can be reached via iterated deletion of dominated types. In the Strategic Complements game, the fear of “dominant strategy Hawks” causes the “Coordination” type who are “almost dominant strategy Hawks” to play H, which causes the “almost-almost dominant strategy Hawks” to play H, etc. The cascade causes higher and higher types to choose H.

At the other end of the distribution when “dominant strategy Doves” choose D, this causes the “Opportunist” type who are “almost dominant strategy Doves” to play D, which causes the “almost-almost dominant strategy Doves” to choose D, etc. This cascade causes lower and lower types to choose D.

With sufficient uncertainty about the type they may be facing, the two cascades resolve the ambiguity about what “moderate” types will do. Thus, in equilibrium, all below a certain cost (type) choose H and those above choose D.

3.2.1.1 Parameters

The paper assumes $d > 0$ and $\mu > 0$ so that Player B’s aggression reduces Player A’s payoff. If it is a Strategic Complements game, then $d > \mu$. This means Player A’s incentive to choose H increases with probability that Player B chooses H. If the game is one of Strategic Substitutes, then $d < \mu$. This means that Player A’ s incentive to choose H decreases with probability that Player B chooses H.

A player is “dominant strategy Hawk” if $\mu \geq c_i$ and $d \geq c_i$, with at least one strict inequality. They are a “dominant strategy Dove” if $\mu \leq c_i$ and $d \leq c_i$, with at least one strict inequality. In the Strategic Complements game, they are a “Coordination” type if $\mu \leq c_i \leq d$, so action H’s best response is H and action D’s best response is D. They are an “almost dominant strategy Hawk” if their cost is slightly above μ. The player is an “Opportunist” Type if $d \leq c_i \leq \mu$, so action D’s best response is H and action H’s best response is D. They are an “almost dominant strategy Dove” if their cost is slightly below d.

Without cheap-talk, the unique Bayesian Nash Communication-Free Equilibrium is where Player i chooses H if $c_i \leq x$, where x is the "cut-off" cost. If c_i is above x, Player i will choose D.

3.2.2 Baliga and Sjostrom's Model With Cheap-Talk

To model the effects of communication, a third player, Player E, is introduced. Player E's payoff function is such that they are either known "dominant strategy Hawks" (denoted by E_H) or "dominant strategy Doves" (E_D). The model assumes c_E is common knowledge among all three players, c_A is known by Player A and Player E, whereas c_B is only known by Player B.

The paper gives two possibilities. In the first case, E_H wants Player A to choose H, and in the second case, E_D wants Player A to choose D. In both cases, Player E always wants Player B to choose D.

In the model, the time line is as follows. c_i, $i \in \{A, B\}$ is determined. Player A and Player E learn c_A, and Player B learns c_B. Then, Player E sends a publicly observed cheap-talk message, $m \in M$, where M is his message space. Finally, Player A and B simultaneously choose H or D.

Note: we can of course have a "babbling equilibrium" where the players ignore the message and play the unique Communication-Free Equilibrium. A Perfect Bayesian Equilibrium with effective cheap-talk is called a Communication Equilibrium. Cheap-talk is only effective if a positive measure of types choose different actions after the message, than they would have done in Communication-Free Equilibrium. In order for this to happen, Player E must reveal some information about A's type.

A strategy for Player E is a function $m : [\underline{c}, \overline{c}] \rightarrow M$ where $m(c_A)$ is the message sent by Player E when Player A's type is c_A. The BS paper shows how, without loss of generality, each player uses a "conditional" cutoff strategy. For any message m, there is a cutoff $c_B(m)$ such that if Player B hears message m, he chooses H (if and only if

$c_B \leq c_B(m)$). Therefore, without loss of generality, they can assume that M contains only two messages, $M = \{m_0, m_1\}$.

3.2.2.1 Strategic Complements

When E is E_H, a Communication Equilibrium exists.

If E is E_H, he knows m_1 will make Player A more likely to play H as it makes Player B play H. This happens even though Player E always prefers Player B to choose D.

Player B is more likely to choose H after m_1 than after m_0 as m_1 conveys information about A's type to B and makes him more likely to play H as m_1 comes only if A comes from a range that does not contain the "dominant strategy Doves". If Player A's action doesn't depend on the message, then E prefers sending m_0. If Player A's action does depend on the message, then A must be "Coordination" type who plays H when sent m_1 (because provocation makes Player B choose H) and D when sent m_0 (because Player B is less likely to choose H).

Since all types want their opponent to choose D, they are all harmed by the third party's cheap-talk. Eliminating E_H makes **everyone** strictly better off.

If E is E_D then he wants both to choose D, so he must always send m_0 in equilibrium. However, a constant message is not informative and the cheap-talk is ineffective.

3.2.2.2 Parameters

In order for the game to be Strategic Complements, $\underline{c} < \mu < d < \overline{c}$. If E is of cost type, $c_E < 0$, then they are E_H. This is because they enjoy a benefit $(- c_E) > 0$ if Player A is aggressive and chooses H. At most, they get 0 if Player A chooses D. Thus, E_H wants Player A to play H.

If $c_E > \mu + d$, then Player E is E_D. They get no more than $\mu - c_E$ if Player A chooses H, and they get no less than $-d > \mu - c_E$ if Player A chooses D. Thus, E_D wants Player A to play D. Both E_H and E_D always want Player B to choose D, so $\mu > 0$ and $d > 0$.

3.2.2.3 Strategic Substitutes

In this section, we briefly consider the game of Strategic Substitutes where $\underline{c} < d < \mu < \overline{c}$.

If E is E_D, a Communication Equilibrium exists. Since m_1 makes Player B more Hawkish, $(c_B(m_1) > c_B(m_0)$, then by Strategic Substitutes, it makes Player A more Dovish, $(c_A(m_1) < c_A(m_0)$.

E_D will send m_1 if and only if Player A is an "Opportunist" who will be induced by m_1 to switch to D. The model shows how by making it more likely that Player B will play H (unless they are a "dominant strategy Dove"), Player A anticipating this will play D.

E_H always wants Player A to choose H so they will always send m_0 and cheap-talk is ineffective.

3.3 Experiment

In the experiment, neutral framing was used to describe the two actions, H and D. From hereafter, we refer to them as 'Option X' and 'Option Y', respectively. We also use neutral framing to describe the types, as explained in the following paragraph.

The game is adapted from BS (2012) with a discrete number of types. There are four types with Type i having cost c_i. There are "dominant strategy Hawks" (Type A, with costs c_A), "dominant strategy Doves" (Type D, with costs c_D) and two "moderate" types with differing costs. The "moderate" types could be a "tough moderate" type (Type B, with costs c_B) or a "weak moderate" (Type C, with costs c_C) type, and may play X or Y, depending on their beliefs about the strategies about other types. X or Y is a dominant strategy for the "dominant strategy" Hawks or Doves.

'Experimenter demand effects' (Zizzo, 2010) are changes in behaviour by experimental subjects due to cues about what constitutes appropriate behaviour. They are a potential problem when they are positively correlated with the true experimental objectives' predictions. In our experiment, our neutral framework avoids this. We only adopt neutral language and don't use words such as 'opponent' or 'partner', nor words such as 'conflict' in our design.

Two sets of experiments were run and each set had two treatments; one with and one without a recommendation. The two sets differed somewhat in the payoffs but both were games with Strategic Complements. We call them Experiment 1 (with Treatment 1 and 2) and Experiment 2 (with Treatment 1 and 2), respectively.

In the experiment, by construction we have ruled out any incentive for the recommender to lie. The third party is a computer who plays the role of a provocateur. If they were to lie, they would just be behaving as a pacifist and the BS theory predicts that this would have no effect on a Strategic Complements game. The provocateur wants Player A to choose X and Player B to choose Y. If Player A's action doesn't depend on the message, then E prefers sending m_0 (increasing the likelihood of Player B choosing Y). If Player A's action does depend on the message, then A must be "Coordination" type who plays H when sent m_1 (because provocation makes Player B choose H) and D when sent m_0 (because Player B is less likely to choose H).

3.3.1 Parameters

Experiment 1: Treatment 1 and 2

We have set c_A equal to 4.5 for the "dominant strategy Hawk" types, c_B is 5 for "tough moderate" types, c_C is 6 for "weak moderate" types and c_D is 6.5 for "dominant strategy Dove" types. The disutility of being less advantaged, d, is 6 and the gain of being more advantaged, μ, is 5.

The probability of being an aggressive “dominant strategy Hawk” type is 0.1, the probability of being either “moderate” type is 0.4 and the probability of being a peaceful “dominant strategy Dove” type is 0.1 across both treatments. This approximates the BS (2012) environment where the number of dominant types is a small fraction of the total population.

	X	Y
X	$-c_i$	$5-c_i$
Y	-6	0

Figure 3.2: Payoff matrix of Player *i* in Experiment 1 with values of *d* and μ inserted

Experiment 2: Treatment 1 and 2

We modified the payoffs in the second set of experiments. We kept c_A equal to 4.5 for the “dominant strategy Hawk” types, but set c_B as 5.5 for “tough moderate” types, c_C as 6.5 for “weak moderate” types and c_D as 7.5 for “dominant strategy Dove” types. μ was kept as 5 but the disutility of being less advantaged, *d*, was changed to 7. This was to increase the loss in payoffs from playing Y when the opponent plays X.

We changed the probabilities of the types. The probability of being either “dominant strategy” types became 0.2 and the probability of being either “moderate” type was set to 0.3. The higher probability of facing a “dominant strategy” type made each “moderate” type ‘safer’ in aligning with the type closest to them.

Treatments 2 included a recommendation. This is explained in more detail in the next section.

3.3.2 Experimental Design

The design consists of a one-shot game with two treatment conditions in each experiment (See Table 3.1). In Treatment 1, there is no recommendation, whereas in Treatment 2, there is. Treatment 1 served as a baseline and Treatment 2 gave the effect of a recommendation. Before playing, each subject in a pair was given a recommendation in Treatments 2, which could be either X or Y. Following a recommendation was not binding. The recommendation mimicked what the Recommender in BS (2012) would have suggested to each type.

Subjects were only allowed to participate in one session (between-subjects design) and, therefore, only one treatment, with an anonymous opponent. We chose to conduct separate subjects per treatment, to rule out any dependence on what happened in the previous treatment. There were a total of 160 participants, as each treatment included two sessions with 20 participants in each. Treating each game as an independent observation, this gave 10 independent observations per session and a total of 20 independent observations per treatment.

With a prior probability (see Section 3.3.1), the subjects were assigned a type; Type A (the "dominant strategy Hawk" type in the model), Type B (the "tough moderate" type), Type C (the "weak moderate" type) and Type D (the "dominant strategy Dove" type). They were told their possible payoffs from each action, which depended on what their opponent played. The payoff structure of the game, including the percentage of types in the experiment, was known. However, the type of each player was private information. Thus, the players knew their own type but not the other player's type. The distribution of types was drawn from a computer programmed with a probability of being either one of the four.

In order to avoid the participants facing negative outcomes, the payoffs were scaled up by 9.5 points in Experiment 1 and 8.5 points in Experiment 2, to ensure positive payoffs.

	X	Y
X	$9.5\text{-}c_i$	$14.5 – c_i$
Y	3.5	9.5

Figure 3.3: Payoff matrix for Player i within Experiment 1 when scaled

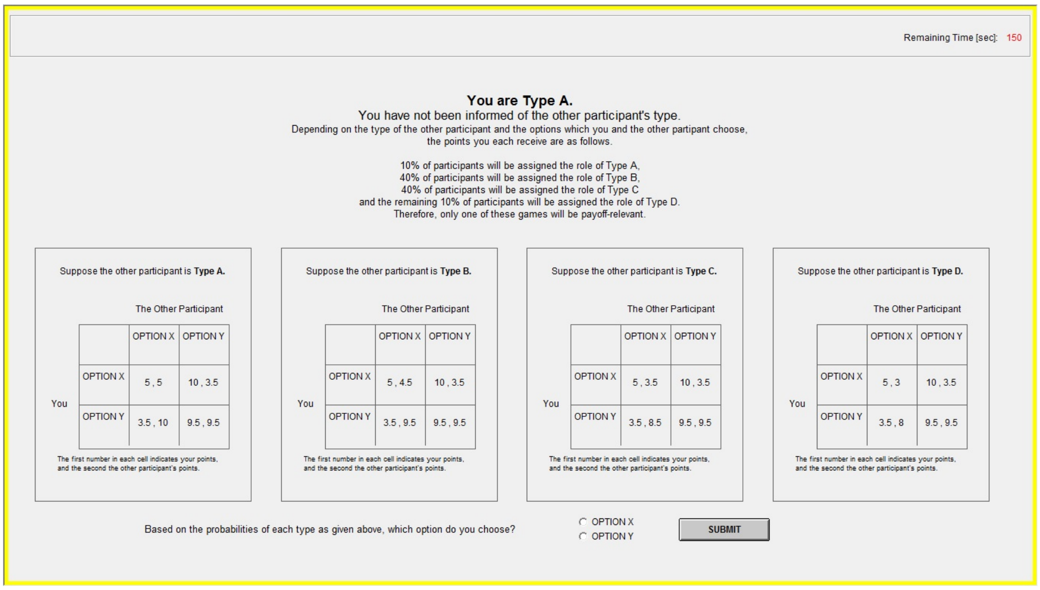
As in our previous chapter, firstly, players were given instructions[5] and asked to complete a quiz to make sure they had understood the information given. An experimenter read the instructions aloud and the experiment only began once all of the participants had answered the quiz's questions correctly. Then, they were assigned their type.

In the next stage, players were matched in groups of two. As the players' payoffs differed depending on what type they were exogenously assigned, we provided a number of payoff matrices to show all of the possible outcomes for the participants from facing each type. The sets of payoff matrices are shown in Figure 3.4, Figure 3.5, Figure 3.6 and Figure 3.7. We have provided the screenshots from the experiment[6].

---

[5] Full instructions are included in the Appendix.

[6] Payoff matrices for Experiment 1 and Experiment 2 are included in the Appendix.



Remaining Time [sec]: 150

You are Type A.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B,
40% of participants will be assigned the role of Type C
and the remaining 10% of participants will be assigned the role of Type D.
Therefore, only one of these games will be payoff-relevant.

Suppose the other participant is **Type A.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 5	10 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type B.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 4.5	10 , 3.5
	OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type C.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 3.5	10 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type D.**

The Other Participant

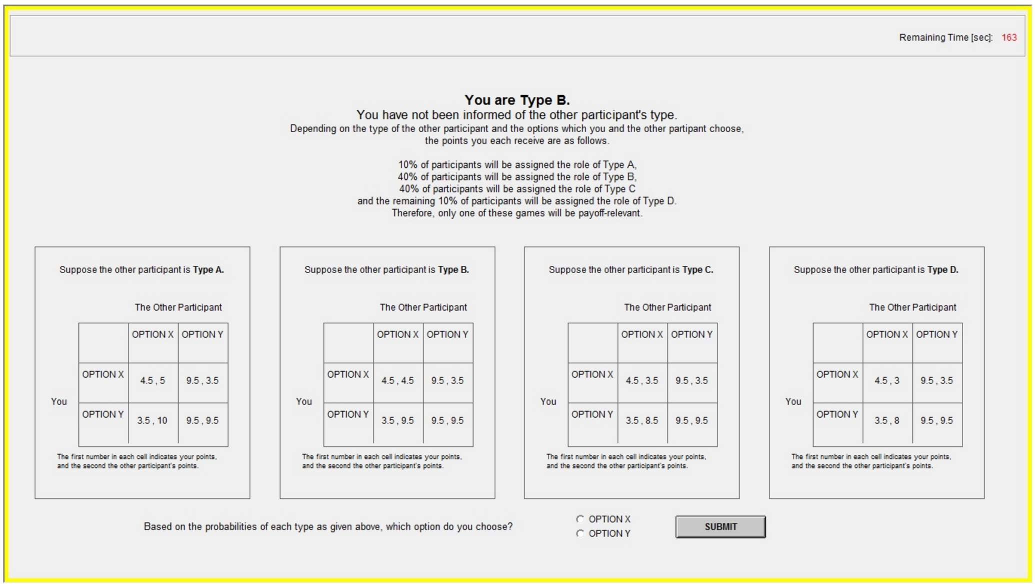
		OPTION X	OPTION Y
You	OPTION X	5 , 3	10 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

Figure 3.4: Payoff matrices for Type A players as it appeared in Experiment 1 (Treatment 1)



Remaining Time [sec]: 163

You are Type B.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B,
40% of participants will be assigned the role of Type C
and the remaining 10% of participants will be assigned the role of Type D.
Therefore, only one of these games will be payoff-relevant.

Suppose the other participant is **Type A.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	4.5 , 5	9.5 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type B.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	4.5 , 4.5	9.5 , 3.5
	OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type C.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	4.5 , 3.5	9.5 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type D.**

The Other Participant

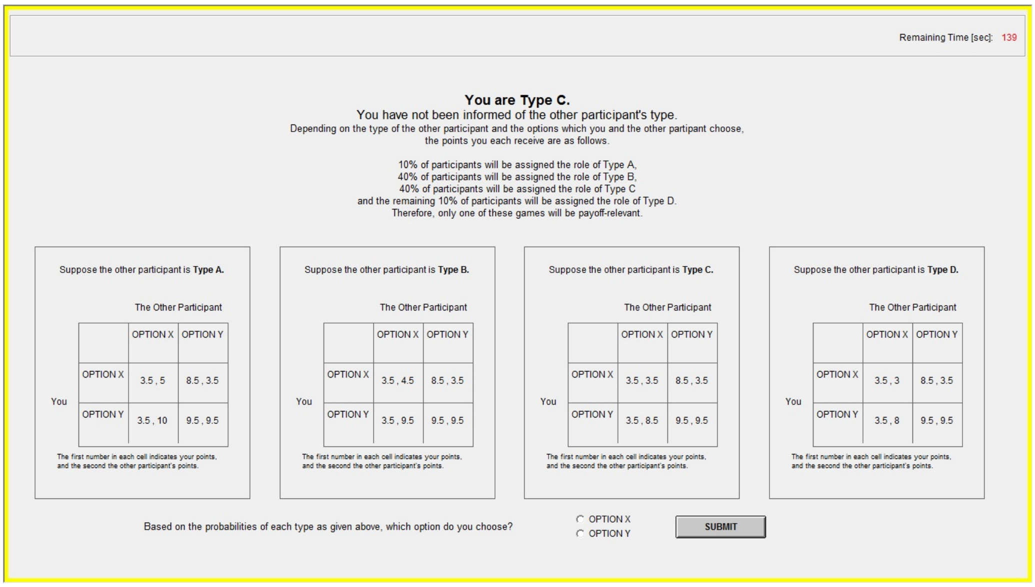
		OPTION X	OPTION Y
You	OPTION X	4.5 , 3	9.5 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

Figure 3.5: Payoff matrices for Type B players as it appeared in Experiment 1 (Treatment 1)



Remaining Time [sec]: 139

You are Type C.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B,
40% of participants will be assigned the role of Type C
and the remaining 10% of participants will be assigned the role of Type D.
Therefore, only one of these games will be payoff-relevant.

Suppose the other participant is **Type A.**

The Other Participant

You

	OPTION X	OPTION Y
OPTION X	3.5 , 5	8.5 , 3.5
OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type B.**

The Other Participant

You

	OPTION X	OPTION Y
OPTION X	3.5 , 4.5	8.5 , 3.5
OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type C.**

The Other Participant

You

	OPTION X	OPTION Y
OPTION X	3.5 , 3.5	8.5 , 3.5
OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type D.**

The Other Participant

You

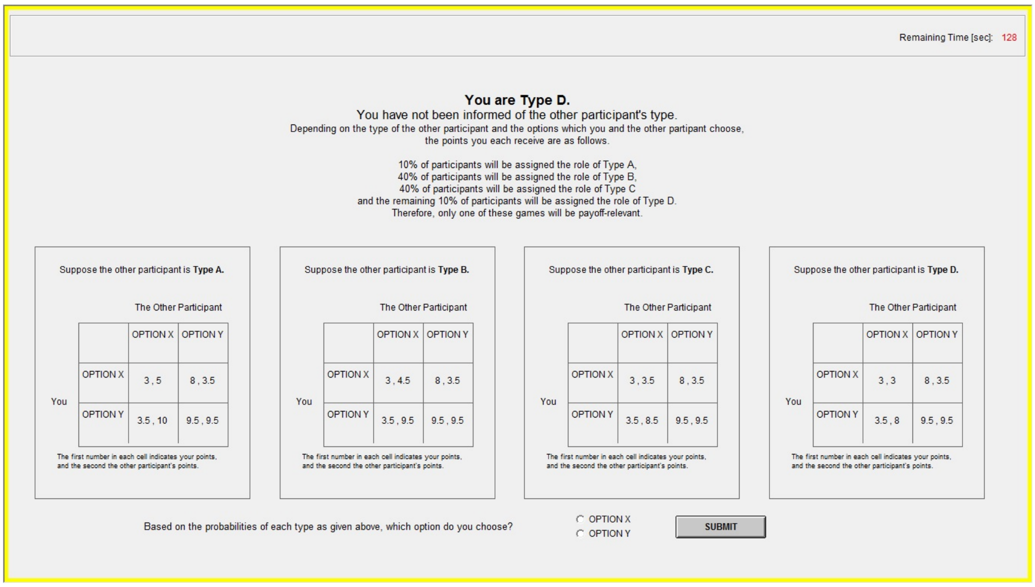
	OPTION X	OPTION Y
OPTION X	3.5 , 3	8.5 , 3.5
OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

Figure 3.6: Payoff matrices for Type C players as it appeared in Experiment 1 (Treatment 1)



Remaining Time [sec]: 128

You are Type D.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B,
40% of participants will be assigned the role of Type C
and the remaining 10% of participants will be assigned the role of Type D.
Therefore, only one of these games will be payoff-relevant.

Suppose the other participant is **Type A.**

The Other Participant

You

	OPTION X	OPTION Y
OPTION X	3 , 5	8 , 3.5
OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type B.**

The Other Participant

You

	OPTION X	OPTION Y
OPTION X	3 , 4.5	8 , 3.5
OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type C.**

The Other Participant

You

	OPTION X	OPTION Y
OPTION X	3 , 3.5	8 , 3.5
OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type D.**

The Other Participant

You

	OPTION X	OPTION Y
OPTION X	3 , 3	8 , 3.5
OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

Figure 3.7: Payoff matrices for Type D players as it appeared in Experiment 1 (Treatment 1)

In Treatment 2, a recommendation was publicly announced within each group (on-screen) simultaneously for 60 seconds, before participants submitted their decisions simultaneously. They were told that the recommendation was made by a Recommender, who knew the type of one participant in their group and who made the recommendation based on this knowledge. They were informed of whether the Recommender knew their type or the other participant's type. They were also told what the recommendations meant. i.e. if the Recommender knew the type was B or C, the recommendation to choose Option X made, whereas, if the Recommender knew the type was A or D, the recommendation to choose Option Y was made. They were reminded that they did not have to choose the option that had been recommended to them.

After this, beliefs about their opponent's behaviour were elicited. They were asked to indicate which option (X or Y) they thought each type they could be matched with, will choose. If they correctly estimated the actual choice made by their opponent, they received £2, thus there were incentives for correct elicitation of beliefs.

At the end of the game, the outcomes were disclosed to each person and everyone was paid privately. The final screen reported their type and their choice, their opponent's type and their choice, the total points they earned, the total points they earned in pounds, their earnings from the belief elicitation, their £2.50 show up fee, plus their corresponding final pound payment. Before they left the laboratory, participants were asked to complete a short questionnaire[7] providing information such as their gender and how clearly they thought the instructions had been.

3.3.3 Experimental Procedures

The experiment was programmed with z-Tree software (Fischbacher, 2007) and all subjects were recruited at the University of Birmingham, using ORSEE software (Greiner, 2015). The experiment was conducted at the Birmingham Experimental Economics Laboratory. In total, 160 subjects participated in the experiment with 40 subjects per treatment. We ran two sessions per treatment, therefore, there were 20 subjects in each session. At the end of each session, subjects were privately paid according to their total amount of experimental currency units (ECUs), using an exchange rate of £1 per ECU. Average earnings (excluding a show- up fee of £2.50) were £8.37. Sessions lasted 40 minutes, on average.

---

[7] For the questionnaire, see 2.12 Appendix E.

3.4 Hypotheses

In this section, we formally present our theoretical hypothesis. The model shows that without a recommendation; Type A and Type B players will play X, whereas Type C and Type D players will play Y. When a recommendation is made, the model shows that there exists another equilibrium where all players will play X, except Type D players who choose Y. This occurs as the Recommender conveys information by saying 'choose X' when the player whose type is known is not a "dominant strategy" type. As a result, the other player is aware that the 'firewall' of a "dominant strategy Dove" is not there and this may cause a switch in her moving from Y to X, however, that may also cause the other player to move from Y to X. Thus, when the recommendation is not made in Treatment 1, we expect more players to choose Y.

Table 3.2: The choices types will make, based on the model

	Treatment	
	T_1 (N=40)	T_2 (N=40)
Type	A, B → X C, D → Y	A, B, C → X D→ Y

Hypothesis 1
Null Hypothesis:
The frequency of X in Treatment 2 is at least the frequency of X in Treatment 1.
Alternate Hypothesis:
The frequency of X is lesser in Treatment 2 compared to Treatment 1.

In order to reject the null, the frequency of X must be lesser in Treatment 2 than in Treatment 1.

More specifically, we have another hypothesis to verify how supportive our results are of the model. BS (2012) show that more players choose X because Type C players switch from playing Y, without recommendation, in Treatment 1 to playing X in

Treatment 2. Therefore, we can check the proportion of Type C players who play X in each Treatment.

Hypothesis 2
Null Hypothesis:
The frequency of X within Type C players in Treatment 2 is at most the frequency of X in Treatment 1.
Alternate Hypothesis:
The frequency of X within Type C players is greater in Treatment 2 compared to Treatment 1.

In order to reject the null, the frequency of X must be greater in Treatment 2 than in Treatment 1.

We now give each experiment's results and interpret the findings.

3.5 Results

We present our analysis in two parts; Hypothesis 1 and Hypothesis 2. We analyse the data from both sets of experiments within each part. In both Experiments 1 and 2, we compare Treatments 1 and 2. We test the hypotheses by first giving a descriptive analysis; checking whether there are differences in the frequency of X between treatments and formally examining whether or not there are statistically significant differences in the observed frequencies among treatments by using a non-parametric Wilcoxon ranksum test. We also test for the Equality of Proportions. Then, we use multivariate probit regressions to see if the differences hold after we include a host of control variables which are likely to affect the dependent variable we consider.

3.5.1 Hypothesis 1

Based on the model by Baliga and Sjostrom, we can predict that the recommendation will affect whether all types prefer to play X. The paper states that without communication, in equilibrium, all below a certain cost (type) choose X and those above choose Y. In our experiment, we ensured that the “tough moderate” type (Type B) plays X and the “weak moderate” type (Type C) plays Y, whilst the remaining types follow their dominant strategy.

BS show that with communication, if a recommendation is made by a E_H type Player E, there is a unique Bayesian Nash Equilibrium in the Strategic Complements game. In this equilibrium, only Types A, B and C choose X.

To test this hypothesis, we consider the frequency of X within the two treatments as Treatment 1 is the baseline and Treatment 2 shows the effect of a recommendation. In order to reject the null, the frequency of X must be greater in Treatment 2 than in Treatment 1.

3.5.1.1 Experiment 1

We look at the observed frequencies of how often the individual’s chose X. Each treatment consisted of 80 subjects. Our data shows that there is an increase of three participants choosing X when the recommendation was made.

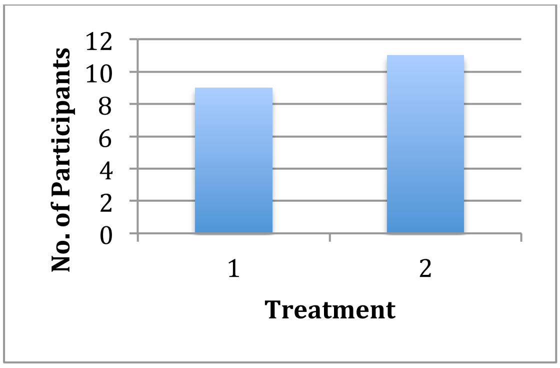


Figure 3.8: Frequency of X in Experiment 1

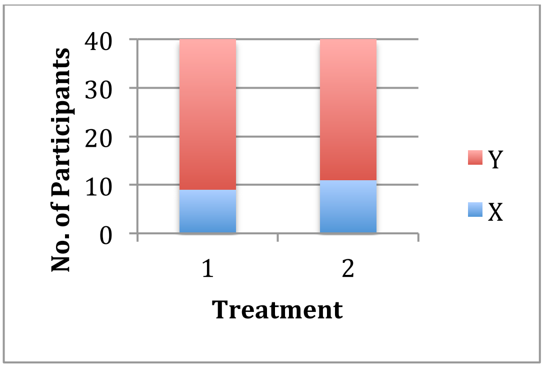


Figure 3.9: Frequency of X and Y in Experiment 1

We can also look at the proportion of X within the individual observations (pairs). An individual observation consists of a group of two subjects. We measure the proportion of X by showing the percentage of each group that chose X; so if both members chose X, the group is assigned 100. The frequency of the outcome (X,X), the number of groups in which both subjects chose X, increased from no groups (out of 20) to one group. The frequency of the outcome (Y,Y) decreased from 11 individual observations to 10.

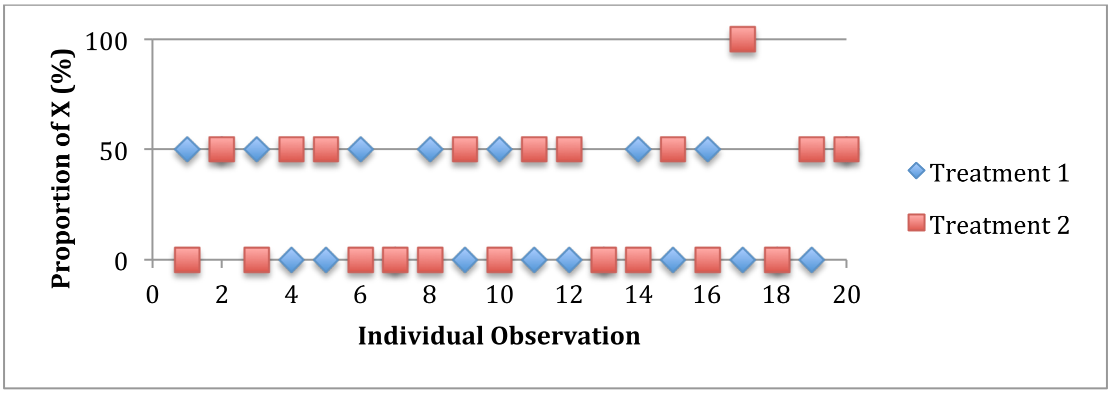


Figure 3.10: Proportion of X for each individual observation in Experiment 1

Figure 3.10 shows a scatter graph of every individual observation in both treatments. Observations 1 to 10 are from the first session and observations 11 to 20 are from the second. The stacked bar chart in Figure 3.11 depicts the lower number of (Y,Y) observations, more clearly.

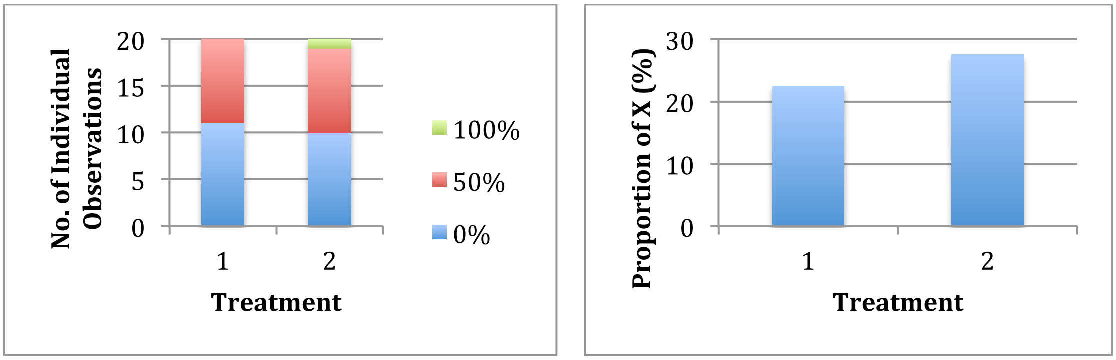


Figure 3.11: Proportion of X within individual observations in Experiment 1

Figure 3.12: Mean proportion of X within individual observations in Experiment 1

Considering the proportion of X in each pair, we present the means of Treatment 1 and 2 for our individual observations, in the table below.

Table 3.3: Mean proportion of X within individual observations in Experiment 1

	Treatment 1	**Treatment 2**
Mean	22.5	27.5

If all participants chose X, the mean would be 100. From Table 3.3, we observe that the mean proportion of pairs choosing X is 22.5 in Treatment 1, and 27.5 in Treatment 2. Therefore, Treatment 2 has a mean that is 22% higher, supporting our hypothesis. This indicates that there are differences in the frequency of X between treatments.

We formally examined whether the differences in the observed frequencies in our treatments are statistically significant, by using a non-parametric Wilcoxon ranksum test, and report the corresponding p-values from the test performed for a pairwise comparison. It tests the hypothesis that two independent samples (that is, unmatched data) are from populations with the same distribution. The p-value for differences in choosing X between Treatments 1 and 2 was 0.6538. This is statistically insignificant and we cannot reject that the two populations have similar distributions.

We also conduct a test for the equality of proportions. This tests whether one population's proportion equals another population's proportion. It tests the equality of two proportions against the alternative that they are not equal. The z-value is -0.52, which means we reject the null and the proportions are not different.

We ran a series of regressions with the data from our first set of experiments but all of these were statistically insignificant too. The results of these regressions are presented in the Table 3.4.

For the regression, the dependent variable *Choice* is binary; it can have only two possible outcomes, which we will denote as 1 and 0. *Choice* takes a value of 1 if the individual chose X and a value of 0 if the individual chose Y.

In our first regression, we include the following independent variables:

1. *Treatment2*: a binary variable that takes value 1 when the data belongs to the second treatment.
2. *Type*: that takes a value of 1 if the individual is of Type A, a value of 2 if the individual is of Type B, a value of 3 if the individual is of Type C and a value of 4 if the individual is of Type D.
3. *Understanding*: a binary variable that takes value 1 when their belief elicitation showed they had understood the model. In Treatment 1, if they correctly identified the strategies for all types, we say they have an understanding. In Treatment 2, if they thought Type As and Type Bs would choose X and Type Ds were likely to play Y, we say they had an understanding. We accepted either X or Y for Type C.
4. *Gender*: a binary variable that takes value 1 when the participant is female.
5. *Economics*: a binary variable that takes value 1 when the participant has studied Economics previously, and may therefore be familiar with game theory.
6. *Religion*: a binary variable that takes value 1 when the participant reports to have any religious inclination.
7. *Instruction*s*:* that takes a value of 0 if the individual found the instructions 'Very Unclear', a value of 1 if the individual found them 'Unclear', a value of 2 if the individual found them 'Neutral', a value of 3 if the individual found them 'Clear' and a value of 4 if they found the instructions 'Very Clear'. No participants in this experiment reported that they had found the instructions 'Unclear' or 'Very Unclear'.

We have 20 observations from each treatment, giving a total of 80 overall. We present the coefficients for these variables from the probit regression with robust standard errors clustered on independent matching groups.

Table 3.4: Probit regression results for Experiment 1

No. of Obs.: 80; Dependent variable: *Choice* =1 if X and =0 ow.				
	Model 1	**Model 2**	**Model 3**	**Model 4**
Treatment2	0.17 (0.34)	0.13 (0.34)	0.16 (0.36)	0.16 (0.36)
Type	-0.90*** (0.25)	-0.86*** (0.25)	-1.00*** (0.27)	-0.98*** (0.28)
Understanding		0.61* (0.35)		0.59 (0.36)
Gender			0.16 (0.39)	0.07 (0.40)
Econ			0.52 (0.35)	0.48 (0.36)
Religion			0.18 (0.38)	0.23 (0.39)
Instructions			0.32 (0.27)	0.36 (0.28)
Constant	1.32** (0.59)	1.03* (0.61)	0.33 (0.89)	-0.02 (0.93)
	Pseudo R^2: 0.1875	Pseudo R^2: 0.2212	Pseudo R^2: 0.2308	Pseudo R^2: 0.2603

There are limited ways in which we can interpret the individual regression coefficients. A positive coefficient means that an increase in the predictor leads to an increase in the predicted probability. We use the probability the z test statistic would be observed under the null hypothesis that a particular predictor's regression coefficient is zero, given that the rest of the predictors are in the model. For a given alpha level, it determines whether or not the null hypothesis can be rejected. If it is less than alpha, then the null hypothesis can be rejected and the parameter estimate is considered statistically significant at that alpha level. We use, following standard notations, * to denote statistical significance at the 10% level, ** at the 5% level and *** at the 1% level.

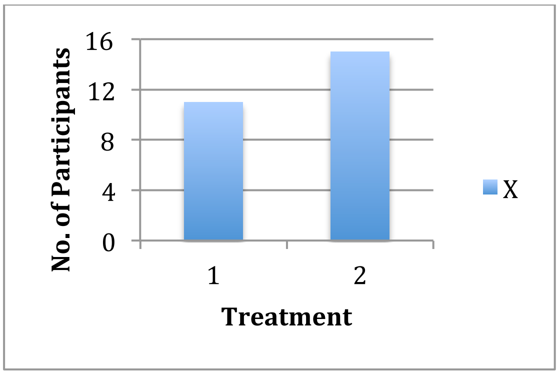
Type has a negative coefficient and is significant at a 1% level, showing us that Type As were the most likely to choose X and Type D's were the least likely, as we expected. Understanding is significant at a 10% level. This means that the more participants seemed to understand the experiment, the more likely they were to play X. While the frequency of X shows more players do seem to choose an aggressive strategy with a recommendation compared to the treatment without, the *Treatment2* variable shows the difference is not statistically significant at conventional levels of significance. The regression analysis confirms our previous results. This may occur for a variety of reasons. Players may simply be playing a 'babbling' equilibrium and ignoring recommendations. It is also possible that the payoff pair (9.5, 9.5) from the non-aggressive strategy may be appearing focal and may look like a 'safe strategy' with implicit co-ordination between players.

We also run regressions using interactive terms. Most interactive terms showed no additional significance of other variables. By interacting *Econ* and *Instructions*, we get a statistically significant (at a 10% significant level) result with a positive coefficient. Also, interacting *Religion* and *Money* (how important a participant thought earning money within the experiment was), *Econ* became significant at 10% level.

We conducted a second set of experiments, whose results we will now give in Section 3.5.1.2. In these experiments, we modified to the payoffs to reduce the likelihood of implicit coordination by making it more costly for them to play Y if there is chance that the other player will play X. Thus, in Experiment 2, a player was penalized greater for playing Y more, when their opponent chooses X. Additionally, by increasing the probability of "dominant strategy" types it provided a better 'pull' for Types C and D to play X and Y, thus creating a bigger incentive to deviate when the Recommender signals that the opponent is not a "dominant strategy Dove".

3.5.1.2 Experiment 2

We look at the observed frequencies of how often the individuals chose X. Both treatments also consisted of 80 subjects and there is an increase of four participants choosing X when the recommendation was made, in comparison to three in the Experiment 1.



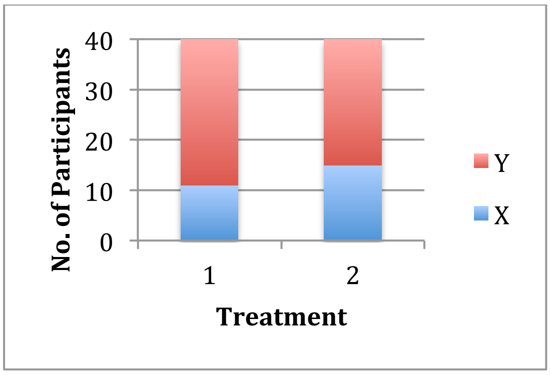


Figure 3.13: Frequency of X in Experiment 2

Figure 3.14: Frequency of X and Y in Experiment 2

The frequency of the outcome (X,X), the number of groups in which both subjects chose X, increased from one group (out of 20) to two groups. The frequency of the outcome (Y,Y) decreased from 10 individual observations to 7. These results support our first hypothesis slightly more than those from Experiment 1, where there was a decrease of one group.

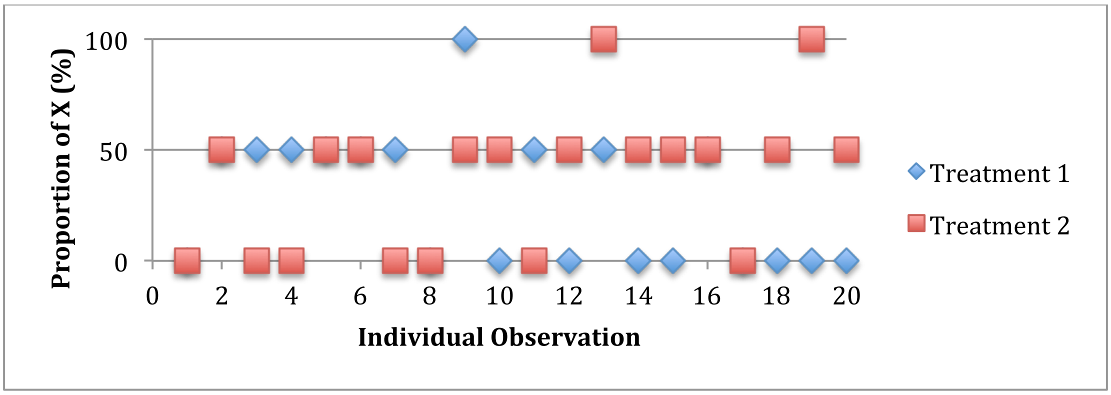
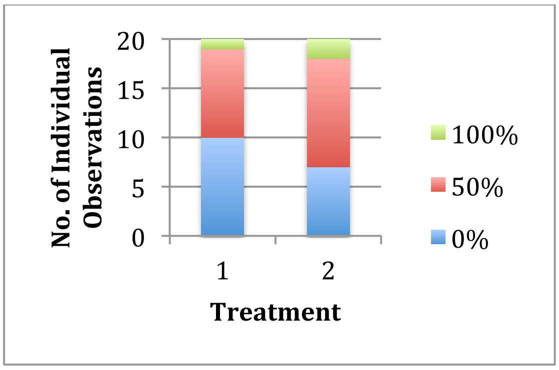


Figure 3.15: Proportion of X for each individual observation in Experiment 2



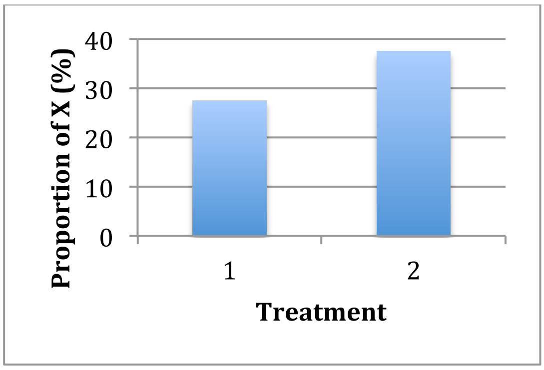


Figure 3.16: Proportion of X within individual observations in Experiment 2

Figure 3.17: Mean proportion of X within individual observations in Experiment 2

In Figure 3.15, Observations 1 to 10 are from the first session and Observations 11 to 20 are from the second session. Figure 3.16 shows more of a decline in the frequency of the outcome (Y,Y) than in Experiment 1, indicating that Experiment 2 may have results that better fit the hypothesis. The difference in the mean proportions of choosing X with the pairs doubled from the previous experiment, from a difference of 5% within an observation to a difference of 10%.

Table 3.5: Mean proportion of X within individual observations in Experiment 2

	Treatment 1	**Treatment 2**
Mean	27.5	37.5

Despite the data looking more promising, the p-value from the Wilcoxon ranksum test was 0.3105 and, therefore, still statistically insignificant. We cannot reject the null and both treatments do not have similar distributions. The z-value from the test for the equality of proportions is -0.95, which means we reject the null and the proportions are not different either.

The results from conducting the probit regressions are shown in Table 3.6. As in Experiment 1, *Type* and *Understanding* are significant. The *Religion* variable is significant here. The negative coefficient suggests participants in this experiment who did not identify themselves with a religion, behaved less aggressively. *Instructions* is statistically significant and has a negative coefficient too, suggesting those participants who understood played less aggressively too. Including interactive terms did not make any other variables significant.

Table 3.6: Probit regression results for Experiment 2

No. of Obs.: 80; Dependent variable: *Choice* =1 if X and =0 ow.				
	Model 1	**Model 2**	**Model 3**	**Model 4**
Treatment2	0.30 (0.30)	0.07 (0.33)	0.13 (0.34)	-0.14 (0.37)
Type	-0.37** (0.16)	-0.35** (0.17)	-0.40** (0.17)	-0.36* (0.19)
Understanding		1.22*** (0.36)		1.15*** (0.40)
Gender			-0.59 (0.38)	-0.36 (0.40)
Econ			0.21 (0.33)	0.18 (0.35)
Religion			-0.84** (0.34)	-0.60 (0.37)
Instructions			-0.62** (0.28)	-0.68** (0.30)
Constant	0.27 (0.42)	-0.02 (0.46)	2.55** (1.01)	2.23** (1.05)
	Pseudo R^2: 0.0677	Pseudo R^2: 0.1899	Pseudo R^2: 0.1800	Pseudo R^2: 0.2682

Result 1 There is no significant difference in the frequency of X between Treatment 1 and Treatment 2; participants do not choose X more or less when a recommendation is made by a provocateur.

3.5.2 Hypothesis 2

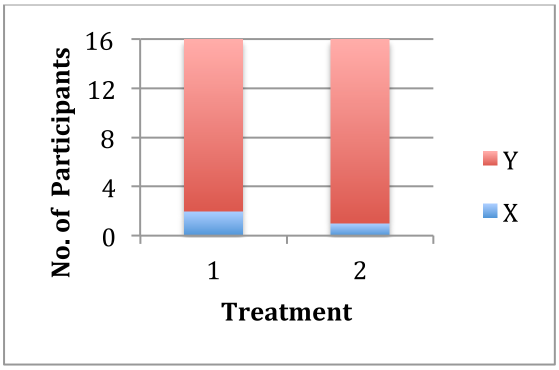
Beyond Hypothesis 1, we can make another, more specific, hypothesis regarding Type C players. The theory states that without communication, in equilibrium, the "weak moderate" type will play Y. This is due to a cascade in the Strategic Complements game, whereby the "dominant strategy Doves" playing Y causes the "almost dominant strategy Doves" to choose Y, causing lower and lower types to choose Y. In our experiment, we ensured that the "weak moderate" type was above a cutoff cost, which meant that in the model, they would play Y in the Communication-Free Equilibrium.

BS show that with communication, if a recommendation is made by a E_H type Player E, in the unique Bayesian Nash Equilibrium in the Strategic Complements game, then Type C players switch from playing Y to playing X.

To test this hypothesis, we consider the frequency of X within Type C players for the two treatments. Once again, Treatment 1 is the baseline and Treatment 2 shows the effect of a recommendation. In order to reject the null, the frequency of X must be greater in Treatment 2 than in Treatment 1.

3.5.2.1 Experiment 1

We check the frequency of X within Type C participants, as well as the frequency of X within the individual observations (matched pairs) that include Type C participants. In Figure 3.18, the number of Type Cs choosing X falls from 2 participants to 1, unlike what theory states. The frequency of X within the individual observations remained the same for the (Y,Y) outcome, at 9 groups. In this experiment, it does not seem that our data supports BS' model.



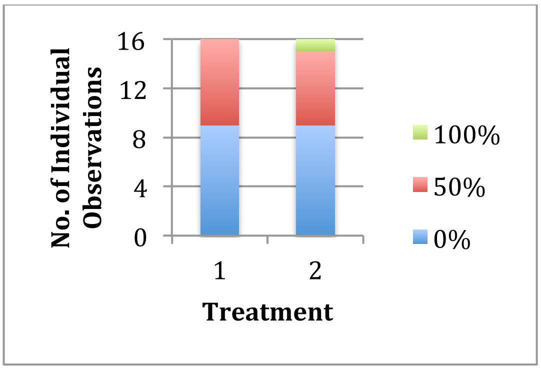


Figure 3.18: The frequency of X and Y for Type C participants in Experiment 1

Figure 3.19: Proportion of X within individual observations that include Type C participants in Experiment 1

The mean proportions of X within the individual observations that included Type C participants are shown in Table 3.7. Although there is a 3.1 increase in the mean, the Wilcoxon ranksum test's p-value of 0.5506 shows that the difference in the distributions is statistically insignificant. We also conducted a Wilcoxon ranksum test using Type C players alone, rather than the individual observations that included Type C players. This reported a p-value of 1.0000, which is also statistically insignificant. The z-value of the test for the equality of proportions is -0.36 for data with only Type C players and 1.47 for data including the pairs. In both tests, we reject the null so the proportions are not different too. Therefore, we cannot conclude that the recommendation is making Type C participants play X less frequently either.

Table 3.7: Mean proportion of X within individual observations that include Type 3 participants in Experiment 1

	Treatment 1	**Treatment 2**
Mean	21.9	25

We conducted a series of probit regressions with the data only including Type C players; however, none of the variables appeared statistically significant. We ran a further set of regressions using the individual observations that included Type C players, and the results are given in the following table where *Type* and

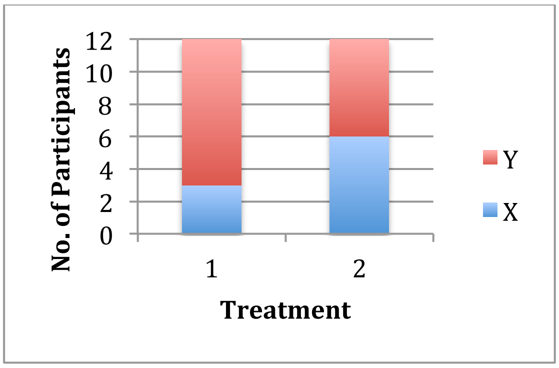
Understanding are statistically significant. Including interactive terms did not make any other variables significant.

Table 3.8: Probit regressions for Experiment 1 using only individual observations that include Type C players

No. of Obs.: 64; Dependent variable: *Choice* =1 if X and =0 ow.				
	Model 1	**Model 2**	**Model 3**	**Model 4**
Treatment2	0.09 (0.37)	0.04 (0.38)	0.12 (0.39)	0.09 (0.39)
Type	-0.81*** (0.26)	-0.76*** (0.27)	-0.90*** (0.29)	-0.82*** (0.30)
Understanding		0.63* (0.38)		0.58 (0.40)
Gender			0.70 (0.43)	-0.05 (0.44)
Econ			0.32 (0.39)	0.22 (0.41)
Religion			-0.03 (0.42)	0.01 (0.42)
Instructions			0.31 (0.32)	0.28 (0.32)
Constant	1.05* (0.64)	0.72 (0.67)	0.31 (1.02)	0.09 (1.06)
	Pseudo R^2: 0.1473	Pseudo R^2: 0.1866	Pseudo R^2: 0.1714	Pseudo R^2: 0.2008

3.5.2.2 Experiment 2

The frequency of X within Type C participants increased in Experiment 2; it doubled from 3 participants to 6. This is a greater change from the results we found in Experiment 1. The proportion of X within the individual observations that included Type C participants also depicts this reduction in the (Y,Y) outcome.



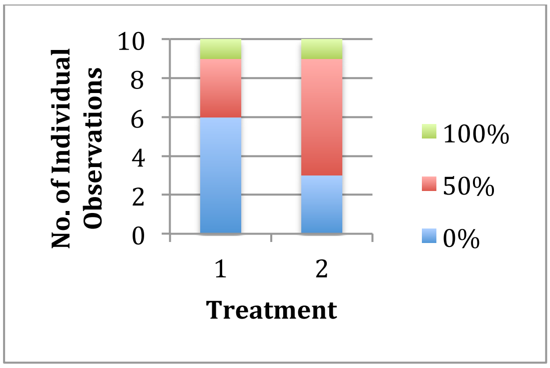


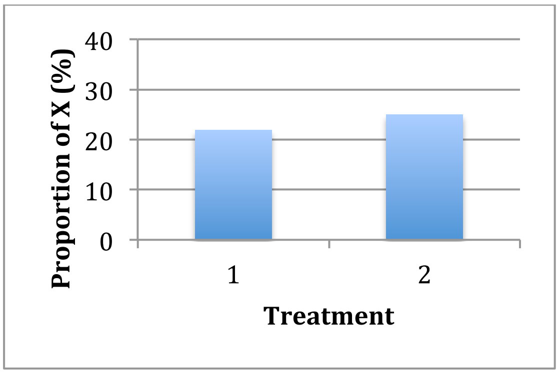
Figure 3.20: The frequency of X and Y for Type C participants in Experiment 2

Figure 3.21: Proportion of X within individual observations that include Type C participants in Experiment 2

The mean proportion of X within the individual observations that include Type C participants is shown in Table 3.9. This increase in 15 is greater than the 3.1 increase in Experiment 1. The difference can be seen more clearly when comparing Figure 3.22 and 3.23.

Table 3.9: Mean proportion of X within individual observations that include Type 3 participants in Experiment 2

	Treatment 1	**Treatment 2**
Mean	25	40



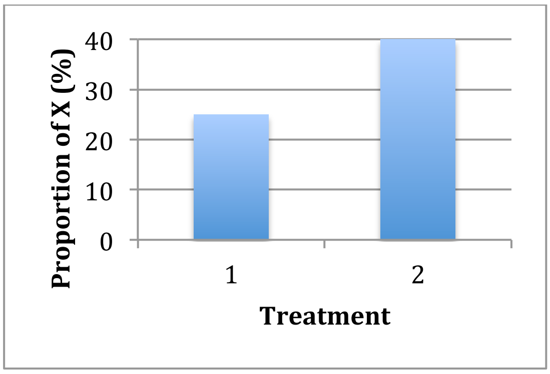


Figure 3.22: Mean proportion of X within individual observations that include Type 3 participants in Experiment 1

Figure 3.23: Mean proportion of X within individual observations that include Type 3 participants in Experiment 2

The Wilcoxon ranksum test's p-value of 0.2594 showed that the difference in the distributions is statistically insignificant. The z-value of the test for the equality of proportions is -1.82 and we reject the null again. Using Type C players alone, it reported a p-value of 0.6602, which is also statistically insignificant, however, the z value from the test for equality of proportions in -3.54. Here, we cannot reject the null so the proportions are different.

Table 3.10: Probit regressions for Experiment 2 using only individual observations that include Type C players

No. of Obs.: 40; Dependent variable: *Choice* =1 if X and =0 ow.				
	Model 1	**Model 2**	**Model 3**	**Model 4**
Treatment2	0.42 (0.42)	0.50 (0.43)	-0.09 (0.52)	-0.04 (0.53)
Type	-0.07 (0.23)	0.06 (0.25)	-0.14 (0.30)	-0.08 (0.32)
Understanding		0.89 (0.57)		0.46 (0.66)
Gender			-1.61** (0.66)	-1.50** (0.67)
Econ			0.23 (0.49)	0.21 0.50)
Religion			-0.39 (0.56)	-0.33 (0.56)
Instructions			-1.45** (0.59)	-1.40** (0.60)
Constant	-0.49 (0.66)	-1.03 (0.78)	4.61** (2.32)	4.14* (2.42)
	Pseudo R^2: 0.0224	Pseudo R^2: 0.0715	Pseudo R^2: 0.2677	Pseudo R^2: 0.2773

Conducting probit regressions with the data only including Type C players, also showed no variables were statistically significant. When running regressions using the individual observations that included Type C players, the *Instructions* variable was statistically significant. *Gender* also became significant, suggesting that females were more likely to play Y. The results of these regressions are given in Table 3.10.

We also conduct a regression of Model 1 with a *belief* variable. This is a binary variable that takes value 1 when the participant believed their opponent would play X. When participants elicited their beliefs, we checked if they thought the type that they were actually matched with, would play X. *Belief* was statistically significant at a 1% level with a positive coefficient, which means if a participant believed their opponent was playing X, they were more likely to play X too.

We run regressions using interactive terms. Interacting *Religion* and *Money* (how important a participant thought earning money within the experiment was) made several variables statistically significant. *Gender, Religion*, *Instructions* and *Money* became statistically significant at a 5% level with negative coefficients. *Money* suggests that if participants cared about earning money, they were more likely to play Y, supporting the possibility that the outcome for (Y,Y) may have appeared focal and "safe".

Result 2 There is no significant difference in the frequency of X within Type C players, between Treatment 1 and Treatment 2 across both experiments. Type C participants do not choose X more or less when a recommendation is made by a provocateur.

3.5.3 The Recommender's Knowledge (Treatment 2 in Experiment 2 only)

As the frequency of X within Type C participants increased in Experiment 2, we looked at the data for which Type C players received recommendations on the knowledge of their type and compared them to those who received recommendations based on their opponent's type. This could enable us to gain a deeper insight into the decisions made by the participants. The data showed that, overall, participants did seem to follow the recommendation. When given a recommendation based on their type (and they were recommended to choose Option X), four followed and two did not. When the recommendation was not based on their type, if they were recommended X then 50% followed (2 out of 4) and if they were recommended Y then 100% followed (2 out of 2).

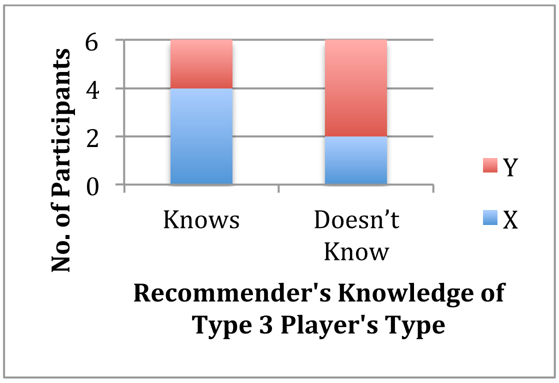


Figure 3.24: The frequency of X and Y for Type C individuals by Recommender's knowledge

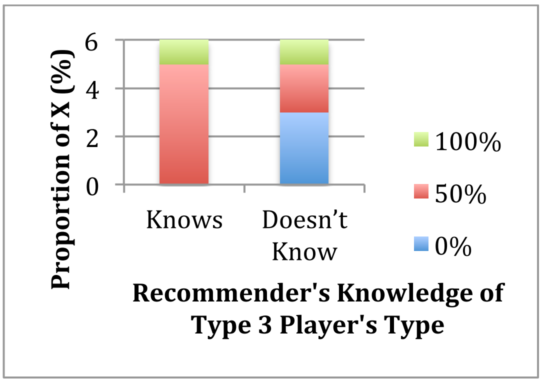


Figure 3.25: The proportion of X within individual observations that include Type C participants by Recommender's knowledge

Furthermore, the frequency of the outcome (Y,Y) became 0 when the recommendation was based on the Type C participant's type. The Wilcoxon ranksum test still reported a statistically insignificant p-value of 0.3515, as did the test when using data solely from Type C players (p-value of 0.1380). The results of the test for equality of proportions was significant for both sets of data (z value is -5.96 for pairs and 5.66 for individuals) and we cannot reject the null so the proportions are different. However, the change is sign suggests these results may be unreliable.

3.6 Conclusion

We conducted a set of experiments based on BS (2012). We wanted to see if a Recommender could cause a change in behaviour of players. In particular, we wanted to see if that could cause a change from a 'peaceful' strategy to an 'aggressive' strategy. Our experiments do not support the hypotheses that players play more aggressively with the recommendation. However, for the Experiment 2, when we conduct an ‘equality of proportions’ test, we do find a higher portion of Type C players switching from Y to X. The fact that recommendations do not cause a change is of course perfectly consistent with theory, as even with a recommendation, there exist a ‘babbling’ equilibrium where all players ignore the recommendation. Further, at least for Experiment 1, the payoff differential from playing Y (rather than X) when the opponent plays X is not great and the payoffs from playing (Y,Y) look attractive for all but the “dominant strategy Hawk” type. Even when the payoff differential is increased, the change across the two treatments is statistically insignificant. There could be several reasons for this, players may be behaving in a way consistent with maximizing sum of payoffs (Englemann and Strobel, 2004), they may be inequity averse (Fehr and Schmidt, 1999) or they may have (pro-)social preferences (Kahneman, Knetsch and Thaler, 1986).

Maximising their sum of payoffs is choosing the outcome that gives the highest total payoff between both players. The sum of both players’ payoffs from YY is always 19 ECU and this is greater than the sum of any other combination of actions, for every type. For example, if a Type A plays X and their opponent (who is also a Type A) plays Y, then the sum of their payoffs is only 13.5 ECU. Inequity aversion is the disutility arising from differences between one’s own payoff and others’ payoffs. Unless your partner is of the same type and plays the same action, YY is the fairest outcome with the least difference (0 because 9.5-9.5=0) between the players’ payoffs, regardless of which types are in your group. (Pro-)social preferences is voluntary behaviour intended to benefit another, possibly motivated by empathy - the concern about the welfare and rights of others - or adherence to one's perceived system of fairness. Unless their partner is a Type A participant, the outcome for YY always yields the highest payoff (9.5) for their partner. Even for Type A participants, the highest payoff they can receive is only 0.5 ECU greater at 10 ECU.

While one would not like to draw wide reaching general conclusions from this, this experiment suggests that ‘provocation’ may not always be successful and if the gains from not engaging in conflict is sufficient, it may be stable enough to withstand provocation.

3.7 Appendix A

Instructions without Communication (Experiment 1)

Preliminary Instructions

Welcome to the Birmingham Experimental Economics Laboratory. This is an experiment in the economics of decision-making. The funding for this research has been provided by various research foundations. Just for showing up, you have already earned a show up fee of £2.50. You can earn additional money depending on the decisions made by you and other participants. It is very important that you read the instructions carefully.

It is important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate your following of these rules.

We will first go over the instructions together. After we have read the instructions, you will have time to ask clarifying questions. We would like to stress that any decisions you make in this experiment are completely anonymous. Please do not touch the computer or its mouse until you are instructed to do so. Thank you.

The experiment consists of one round. You will be paid according to the number of points you have collected in that round with the following exchange rate:

1 point = £1

You will be paid privately in cash at the end of the experiment.

Instructions

Type and Matching

In this experiment, you will be randomly assigned a type by the computer, either 'Type A', 'Type B', 'Type C' or 'Type D'. **Your type will remain the same throughout the whole experiment.** In total there are 20 participants in this room; 10% will be assigned the role of **Type A**, 40% will be assigned the role of **Type B**, 40% of the participants will be assigned the role of **Type C** and the remaining 10% of the participants will be assigned the role of **Type D**. You will be informed of your type but not of the other participants' types.

At the beginning of the round, you will be matched with one other person, randomly selected from the participants in this room, to form a group of two. You will therefore be in a group with one other participant. You will never learn the identity of the other participant you are matched with, nor will the other participants be informed about your identity.

The Decision Stage

You and the other participant you are matched with have to **simultaneously select one** of two options, either OPTION X or OPTION Y. When you make your own choice, you do not know what the other participant's choice is. Your type and the decisions made by both of you will determine the points you each get, as shown in the following tables. The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type A:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 5	10 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 4.5	10 , 3.5
	OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 3.5	10 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type D:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 3	10 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type B:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	4.5 , 5	9.5 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	4.5 , 4.5	9.5 , 3.5
	OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	4.5 , 3.5	9.5 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type D:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	4.5 , 3	9.5 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type C:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3.5 , 5	8.5 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3.5 , 4.5	8.5 , 3.5
	OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3.5 , 3.5	8.5 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type D:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3.5 , 3	8.5 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type D:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3 , 5	8 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3 , 4.5	8 , 3.5
	OPTION Y	3.5 , 9.5	9.5, 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3 , 3.5	8 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

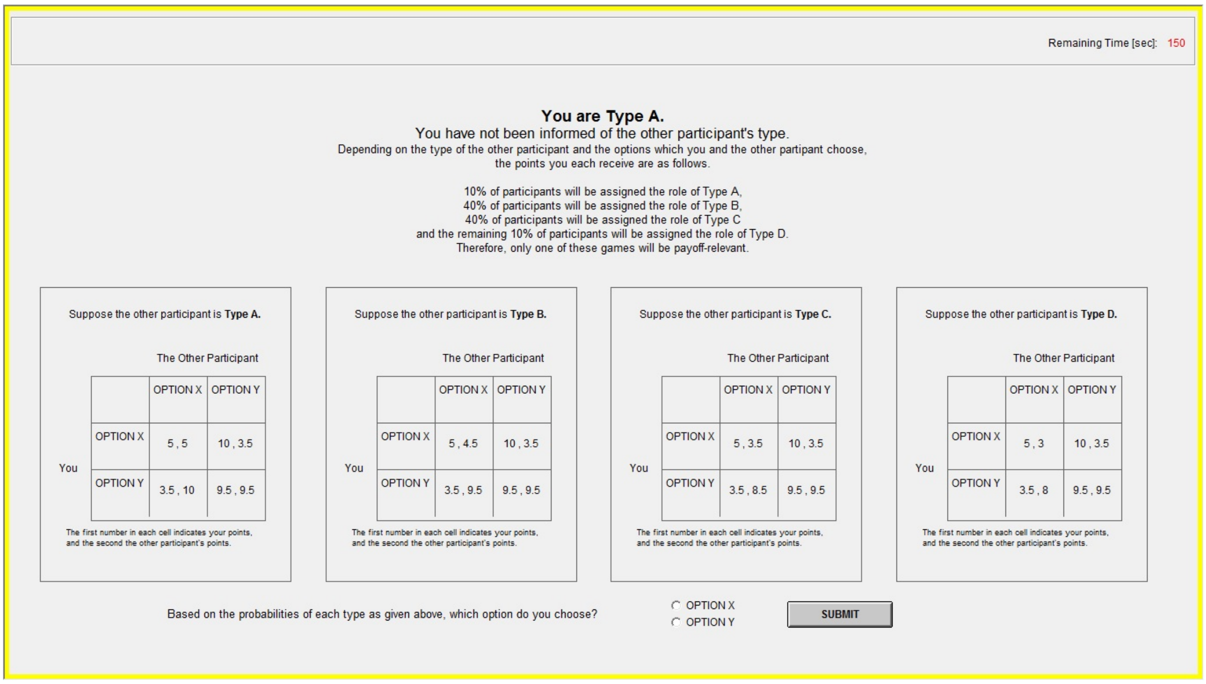
Suppose the other participant is Type D:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3 , 3	8 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

For example, suppose that you are Type A and the other participant you are matched with is Type B. If you choose OPTION X and they choose OPTION Y, then you will earn 10 points and the other participant will earn 3.5 points.

The Decision Screen
You will enter your decisions on a screen that looks as follows. The example below is used for illustrative purposes. This screenshot corresponds to the case where you are Type A. In the actual experiment, you can be Type A, Type B, Type C or Type D as mentioned earlier in the instructions.



Remaining Time [sec]: 150

You are Type A.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B,
40% of participants will be assigned the role of Type C
and the remaining 10% of participants will be assigned the role of Type D.
Therefore, only one of these games will be payoff-relevant.

Suppose the other participant is **Type A.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 5	10 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type B.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 4.5	10 , 3.5
	OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type C.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 3.5	10 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type D.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 3	10 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

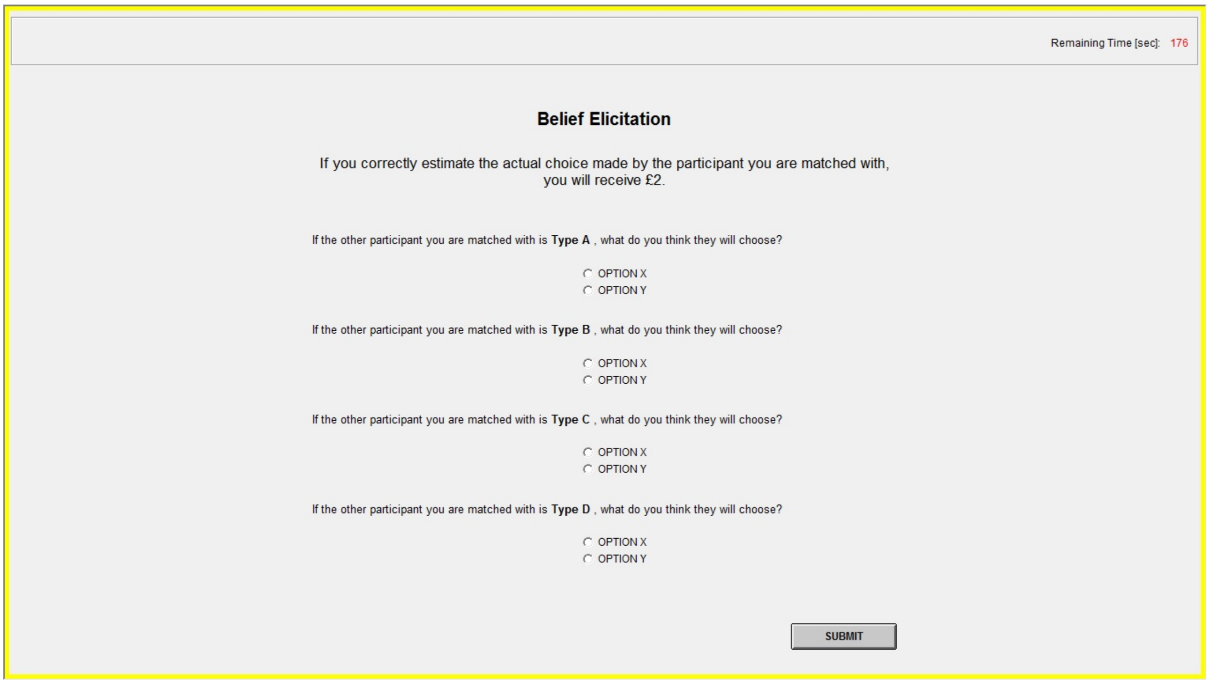
The first number in each cell indicates your points,
and the second the other participant's points.

Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

As mentioned before, before you make your decision, you will be informed of your Type and the payoff tables for facing each type. You must select **one** of two options, either OPTION X or OPTION Y. To make your choice, you simply select the appropriate button and then click on the ‘Submit’ button by using your mouse. Once you have pressed the ‘Submit’ button, you can no longer revise your decision.

Belief Elicitation
On the next screen, you are asked to indicate which of the two options (OPTION X or OPTION Y) each type of the other person you could be matched with will choose. You will make your choices on a screen that looks as follows.



Remaining Time [sec]: 176

Belief Elicitation

If you correctly estimate the actual choice made by the participant you are matched with, you will receive £2.

If the other participant you are matched with is **Type A** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If the other participant you are matched with is **Type B** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If the other participant you are matched with is **Type C** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If the other participant you are matched with is **Type D** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If you correctly estimate the actual choice made by the participant you are matched with, you will receive £2.

Ending of the Experiment

When both you and the other participant you are matched with have made your decisions, you will see a final screen reporting your type and your choice, the other participant's type and their choice, the total points you have earned, the total points you have earned in £, your earnings from the belief elicitation, your show up fee, plus the corresponding final £ payment. You will be paid £1 for every point and £2 if you estimate the other participant's choice correctly, in addition to your £2.50 show-up fee.

Do you have any questions? Please do not ask any questions out loud. If you have any questions, please raise your hand and an experimenter will come to your desk.

Control Questions

To make sure everyone understands the instructions, please complete the questions below. In a couple of minutes someone will come to your desk to check your answers. Once everybody answers the following questions correctly, the experiment will start. (The decisions and earnings used for the questions below are simply for illustrative purposes. In the experiment decisions and earnings will depend on the actual choices of the participants.)

1. Will your type change or remain the same throughout the experiment? (Please circle the correct answer).

SAME / CHANGE

2. How many 'Type A' participants are there in this experiment?
…………

3. If you are a 'Type A' participant, how many points would you receive by choosing OPTION X when matched with a 'Type D' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

4. If you are a 'Type B' participant, how many points would you receive by choosing OPTION Y when matched with a 'Type C' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

5. If you are a 'Type C' participant, how many points would you receive by choosing OPTION X when matched with a 'Type A' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

3.8 Appendix B

Instructions with Communication (Experiment 1)

Preliminary Instructions

Welcome to the Birmingham Experimental Economics Laboratory. This is an experiment in the economics of decision-making. The funding for this research has been provided by various research foundations. Just for showing up, you have already earned a show up fee of £2.50. You can earn additional money depending on the decisions made by you and other participants. It is very important that you read the instructions carefully.

It is important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate your following of these rules.

We will first go over the instructions together. After we have read the instructions, you will have time to ask clarifying questions. We would like to stress that any decisions you make in this experiment are completely anonymous. Please do not touch the computer or its mouse until you are instructed to do so. Thank you.

The experiment consists of one round. You will be paid according to the number of points you have collected in that round with the following exchange rate:

1 point = £1

You will be paid privately in cash at the end of the experiment.

Instructions

Type and Matching

In this experiment, you will be randomly assigned a type by the computer, either 'Type A', 'Type B', 'Type C' or 'Type D'. **Your type will remain the same throughout the whole experiment.** In total there are 20 participants in this room; 10% will be assigned the role of **Type A**, 40% will be assigned the role of **Type B**, 40% of the participants will be assigned the role of **Type C** and the remaining 10% of the participants will be assigned the role of **Type D**. You will be informed of your type but not of the other participants' types.

At the beginning of the round, you will be matched with one other person, randomly selected from the participants in this room, to form a group of two. You will therefore be in a group with one other participant. You will never learn the identity of the other participant you are matched with, nor will the other participants be informed about your identity.

The round consists of **two stages:** i) the Recommendation stage and ii) the Decision stage.

i) The Recommendation Stage

On your screen, a message shall appear for 60 seconds, recommending you to choose OPTION X or OPTION Y.

The Recommendation is made by a Recommender. The Recommender knows the type of **one** participant in your group and it makes the Recommendation based on this knowledge.

You will be informed of whether the Recommender knows **your** type or the **other participant in your group's** type. If the Recommender knows the type is B or C, the Recommendation to 'choose OPTION X' is made. If it knows the type is A or D, the Recommendation to 'choose OPTION Y' is made. Both participants in your group can see the Recommendation.

If you are advised to choose OPTION X, the following message will appear on your screen:

The recommendation is to choose OPTION X.

If you are advised to choose OPTION Y, the following message will appear on your screen:

The recommendation is to choose OPTION Y.

You **do not** have to choose the option that has been recommended in this stage.

Once 60 seconds have elapsed, we will proceed to the 'Decision stage'.

ii) The Decision Stage
You and the other participant you are matched with have to **simultaneously select one** of two options, either OPTION X or OPTION Y. When you make your own choice, you do not know what the other participant's choice is. Your type and the decisions made by both of you will determine the points you each get, as shown in the following tables. The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type A:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 5	10 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 4.5	10 , 3.5
	OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 3.5	10 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type D:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	5 , 3	10 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type B:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	4.5 , 5	9.5 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	4.5 , 4.5	9.5 , 3.5
	OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	4.5 , 3.5	9.5 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type D:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	4.5 , 3	9.5 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type C:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3.5 , 5	8.5 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3.5 , 4.5	8.5 , 3.5
	OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3.5 , 3.5	8.5 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type D:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3.5 , 3	8.5 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Payoff tables for Type D:

Suppose the other participant is Type A:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3 , 5	8 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type B:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3 , 4.5	8 , 3.5
	OPTION Y	3.5 , 9.5	9.5, 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type C:

		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3 , 3.5	8 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

Suppose the other participant is Type D:

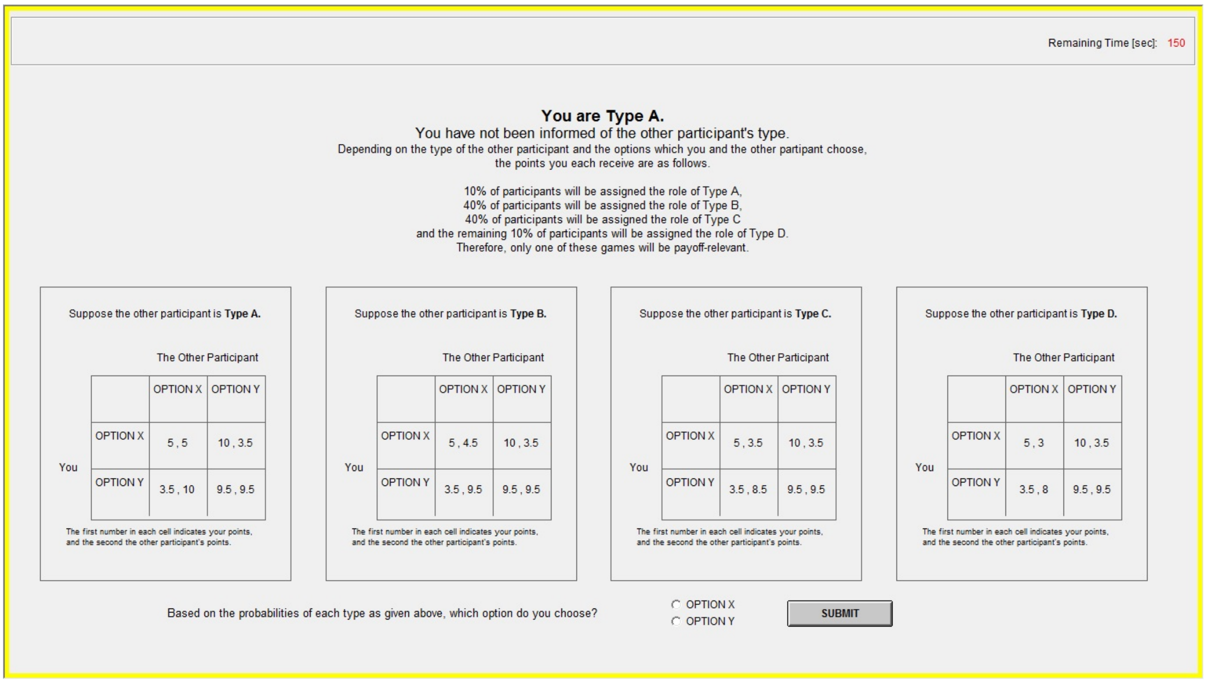
		The Other Participant	
		OPTION X	OPTION Y
You	OPTION X	3 , 3	8 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points, and the second the other participant's points.

For example, suppose that you are Type A and the other participant you are matched with is Type B. If you choose OPTION X and they choose OPTION Y, then you will earn 10 points and the other participant will earn 3.5 points.

The Decision Screen

You will enter your decisions on a screen that looks as follows. The example below is used for illustrative purposes. This screenshot corresponds to the case where you are Type A. In the actual experiment, you can be Type A, Type B, Type C or Type D as mentioned earlier in the instructions.



Remaining Time [sec]: 150

You are Type A.
You have not been informed of the other participant's type.
Depending on the type of the other participant and the options which you and the other partipant choose,
the points you each receive are as follows.

10% of participants will be assigned the role of Type A,
40% of participants will be assigned the role of Type B,
40% of participants will be assigned the role of Type C
and the remaining 10% of participants will be assigned the role of Type D.
Therefore, only one of these games will be payoff-relevant.

Suppose the other participant is **Type A.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 5	10 , 3.5
	OPTION Y	3.5 , 10	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type B.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 4.5	10 , 3.5
	OPTION Y	3.5 , 9.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type C.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 3.5	10 , 3.5
	OPTION Y	3.5 , 8.5	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

Suppose the other participant is **Type D.**

The Other Participant

		OPTION X	OPTION Y
You	OPTION X	5 , 3	10 , 3.5
	OPTION Y	3.5 , 8	9.5 , 9.5

The first number in each cell indicates your points,
and the second the other participant's points.

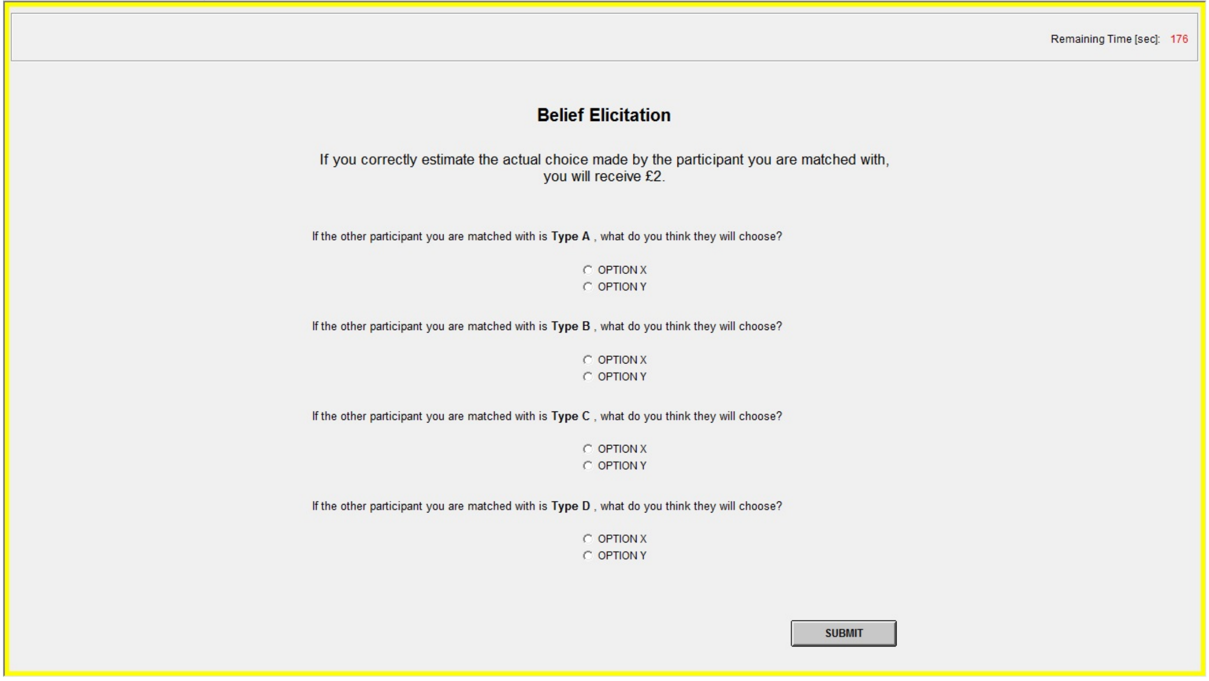
Based on the probabilities of each type as given above, which option do you choose?

○ OPTION X
○ OPTION Y

As mentioned before, before you make your decision, you will be informed of your Type and the payoff tables for facing each type. You must select **one** of two options, either OPTION X or OPTION Y. To make your choice, you simply select the appropriate button and then click on the ‘Submit’ button by using your mouse. Once you have pressed the ‘Submit’ button, you can no longer revise your decision.

Belief Elicitation

On the next screen, you are asked to indicate which of the two options (OPTION X or OPTION Y) each type of the other person you could be matched with will choose. You will make your choices on a screen that looks as follows.



Remaining Time [sec]: 176

Belief Elicitation

If you correctly estimate the actual choice made by the participant you are matched with,
you will receive £2.

If the other participant you are matched with is **Type A** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If the other participant you are matched with is **Type B** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If the other participant you are matched with is **Type C** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If the other participant you are matched with is **Type D** , what do you think they will choose?

○ OPTION X
○ OPTION Y

If you correctly estimate the actual choice made by the participant you are matched with, you will receive £2.

Ending of the Experiment

When both you and the other participant you are matched with have made your decisions, you will see a final screen reporting your type and your choice, the other participant's type and their choice, the total points you have earned, the total points you have earned in £, your earnings from the belief elicitation, your show up fee, plus the corresponding final £ payment. You will be paid £1 for every point and £2 if you estimate the other participant's choice correctly, in addition to your £2.50 show-up fee.
Do you have any questions? Please do not ask any questions out loud. If you have any questions, please raise your hand and an experimenter will come to your desk.

Control Questions

To make sure everyone understands the instructions, please complete the questions below. In a couple of minutes someone will come to your desk to check your answers. Once everybody answers the following questions correctly, the experiment will start. (The decisions and earnings used for the questions below are simply for illustrative purposes. In the experiment decisions and earnings will depend on the actual choices of the participants.)

1. Will your type change or remain the same throughout the experiment? (Please circle the correct answer).
SAME / CHANGE

2. How many 'Type A' participants are there in this experiment?
…………

3. If you are a 'Type A' participant, how many points would you receive by choosing OPTION X when matched with a 'Type D' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

4. If you are a 'Type B' participant, how many points would you receive by choosing OPTION Y when matched with a 'Type C' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

5. If you are a 'Type C' participant, how many points would you receive by choosing OPTION X when matched with a 'Type A' participant and…
    a.) they choose Option X? …………
    b.) they choose Option Y? …………

3.9 Appendix C

Payoff Matrices in Experiment 1

Payoff Tables for Type A

Type A

		Option X	Option Y
Type A	Option X	5, 5	10, 3.5
	Option Y	3.5, 10	9.5, 9.5

Type B

		Option X	Option Y
Type A	Option X	5, 4.5	10, 3.5
	Option Y	3.5, 9.5	9.5, 9.5

Type C

		Option X	Option Y
Type A	Option X	5, 3.5	10, 3.5
	Option Y	3.5, 8.5	9.5, 9.5

Type D

		Option X	Option Y
Type A	Option X	5, 3	10, 3.5
	Option Y	3.5, 8	9.5, 9.5

Payoff Tables for Type B

		Type A	
		Option X	Option Y
Type B	Option X	4.5, 5	9.5, 3.5
	Option Y	3.5, 10	9.5, 9.5

		Type B	
		Option X	Option Y
Type B	Option X	4.5, 4.5	9.5, 3.5
	Option Y	3.5, 9.5	9.5, 9.5

		Type C	
		Option X	Option Y
Type B	Option X	4.5, 3.5	9.5, 3.5
	Option Y	3.5, 8.5	9.5, 9.5

		Type D	
		Option X	Option Y
Type B	Option X	4.5, 3	9.5, 3.5
	Option Y	3.5, 8	9.5, 9.5

Payoff Tables for Type C

		Type A	
		Option X	Option Y
Type C	Option X	3.5, 5	8.5, 3.5
	Option Y	3.5, 10	9.5, 9.5

		Type B	
		Option X	Option Y
Type C	Option X	3.5, 4.5	8.5, 3.5
	Option Y	3.5, 9.5	9.5, 9.5

		Type C	
		Option X	Option Y
Type C	Option X	3.5, 3.5	8.5, 3.5
	Option Y	3.5, 8.5	9.5, 9.5

		Type D	
		Option X	Option Y
Type C	Option X	3.5, 3	8.5, 3.5
	Option Y	3.5, 8	9.5, 9.5

Payoff Tables for Type D

Type A

		Option X	Option Y
Type D	Option X	3, 5	8, 3.5
	Option Y	3.5, 10	9.5, 9.5

Type B

		Option X	Option Y
Type D	Option X	3, 4.5	8, 3.5
	Option Y	3.5, 9.5	9.5, 9.5

Type C

		Option X	Option Y
Type D	Option X	3, 3.5	8, 3.5
	Option Y	3.5, 8.5	9.5, 9.5

Type D

		Option X	Option Y
Type D	Option X	3, 3	8, 3.5
	Option Y	3.5, 8	9.5, 9.5

3.10 Appendix D

Payoff Matrices in Experiment 2

Payoff Tables for Type A

Type A

		Option X	Option Y
Type A	Option X	4, 4	9, 1.5
	Option Y	1.5, 9	8.5, 8.5

Type B

		Option X	Option Y
Type A	Option X	4, 3	9, 1.5
	Option Y	1.5, 8	8.5, 8.5

Type C

		Option X	Option Y
Type A	Option X	4, 2	9, 1.5
	Option Y	1.5, 7	8.5, 8.5

Type D

		Option X	Option Y
Type A	Option X	4, 1	9, 1.5
	Option Y	1.5, 6	8.5, 8.5

Payoff Tables for Type B

		Type A	
		Option X	Option Y
Type B	Option X	3, 4	8, 1.5
	Option Y	1.5, 9	8.5, 8.5

		Type B	
		Option X	Option Y
Type B	Option X	3, 3	8, 1.5
	Option Y	1.5, 8	8.5, 8.5

		Type C	
		Option X	Option Y
Type B	Option X	3, 2	8, 1.5
	Option Y	1.5, 7	8.5, 8.5

		Type D	
		Option X	Option Y
Type B	Option X	3,1	8, 1.5
	Option Y	1.5, 6	8.5, 8.5

Payoff Tables for Type C

		Type A	
		Option X	Option Y
Type C	Option X	2, 4	7, 1.5
	Option Y	1.5, 9	8.5, 8.5

		Type B	
		Option X	Option Y
Type C	Option X	2, 3	7, 1.5
	Option Y	1.5, 8	8.5, 8.5

		Type C	
		Option X	Option Y
Type C	Option X	2, 2	7, 1.5
	Option Y	1.5, 7	8.5, 8.5

		Type D	
		Option X	Option Y
Type C	Option X	2, 1	7, 1.5
	Option Y	1.5, 6	8.5, 8.5

Payoff Tables for Type D

		Type A	
		Option X	Option Y
Type D	Option X	1, 4	6, 1.5
	Option Y	1.5, 9	8.5, 8.5

		Type B	
		Option X	Option Y
Type D	Option X	1, 3	6, 1.5
	Option Y	1.5, 8	8.5, 8.5

		Type C	
		Option X	Option Y
Type D	Option X	1, 2	6, 1.5
	Option Y	1.5, 7	8.5, 8.5

		Type D	
		Option X	Option Y
Type D	Option X	1, 1	6, 1.5
	Option Y	1.5, 6	8.5, 8.5

Conclusion

This conclusion discusses the external validity of this thesis. It also includes the limitations of our study, a reminder of its policy implications and possible extensions to Chapter 2 and Chapter 3.

In Chapter 2, we investigated the impact that communication had on deescalating conflict. Consistent with the role of communication seen in our experiment, Reiter (1995) finds that leaders can prevent conflict by communicating if they understand the spiraling logic. The paper looks at a war in which a player attacks to forestall what they believe is an imminent attack on themselves, and in fact, the main empirical finding is that preemptive wars almost never happen.

There are a few real-world examples of how fear has motivated an attack, for example, the Arab-Israeli War in 1967 and China's intervention in the Korean War. Although, as with many studies on conflict, the causes are often inconclusive with many other motivations for war being possible. It is, therefore, difficult to provide empirical studies to compare our findings with.

We investigate the impact that communication had on escalating conflict in Chapter 3. Walsh (2010) reported on how a provocative attack by a third party can trigger conflict. He stated that after a terrorist attack in Mumbai, Indian government officials identified the perpetrators as the ISI who were not controlled by Pakistan's civilian government.

Not only is it often difficult to identify the identity of provocateurs (whether they are third party or not) and their motives, it is also difficult to determine whether actions are truly strategic substitutes or complements. During the Cold War, there was much uncertainty on whether toughness would make the Soviet Union more or less aggressive. These are further reasons as to why it is difficult to find empirical studies for comparison.

Potential limitations of this research may be the sample size, particularly for Chapter 3. Although the number of individual observations is larger than many other laboratory experiments, it is possible that a greater number of people may provide a more robust test of the differences across treatments. It would also be useful to conduct more treatments in both chapters, to distinguish whether communication by itself is responsible for our results, or whether the nature of the communication matters. It would be interesting to see the impact of one-way or fixed-form communication in Chapter 2 and, in Chapter 3, it would be interesting to investigate the effect of bi-lateral communication between the Recommender and the players, for example.

Both chapters better enable us to analyse if information that would be used to reduce or increase the likelihood of conflict, particularly in the case of an arm's race, can be generated via communication. In terms of policy implications, Chapter 2 may provide insight into whether formal agreements can be self-enforcing. Communication was successful in our experiment, suggesting that non-binding treaties may actually be effective. Chapter 3 could potentially contribute to tackling extremism. An improved understanding of extremist behaviour will help us to approach extremists in a variety of situations. Agents with extreme agendas (ideologues, such as lobby groups) sometimes take provocative actions to signal the intention of their leaders and these actions are able to inflame conflicts. A provocative recommender could also employ aggressive bargaining tactics during negotiations or they may be a person elected by voters, who want them to make aggressive actions on their behalf. Our research suggests that these signals/tactics/actions may not be the actual root of the escalation and the inflammation may be attributed to other factors such as those mentioned in Chapter 1.

There are a number of extensions for our research. In Chapter 2, we could investigate the impact of one-sided communication, repeated interaction with multiple rounds and the differences that may occur when incorporating an Endowment effect. Players have been shown to behave differently when facing a reduction of their earnings (Kahneman, Knetsch and Thaler, 1990), which may actually be more similar to the cost they would be affected by if the situations we created within our experiments had existed in real life. In Chapter 3, it would be interesting to see the effect of a third-party pacifist. Furthermore, we could run the experiments but with the payoffs such that it becomes a Strategic Substitutes game to see if a peaceful ideologue can bring a reduction in Hawkish actions.

References

Abbink, K. and S. Pezzini (2005). “Determinants of Revolt: Evidence from Survey and Laboratory Data”. Working paper, University of Nottingham.

Abbink, K, D. Masclet and D. Mirza (2008). “Inequality and Riots – Experimental Evidence”. Working paper, University of East Anglia.

Abbink, K. and J. Brandts (2009). “Political Autonomy and Independence: Theory and Experimental Evidence”. Working paper, University of Amsterdam.

Abbink, K. (2011). “Trust on the Brink of Armageddon: The First-Strike Game”. Working paper, University of East Anglia.

Aumann, R. J. (1990). “Nash Equilibria are Not Self-Enforcing”, in J. J. Gabszewicz, J. F. Richard and L. A. Wolsey (eds.) Economic Decision-Making: Games, Econometrics and Optimization, Amsterdam, Elsevier.

Aumann, R. J. and L. S. Shapley (1994). "Long-Term Competition—A Game-Theoretic Analysis", Essays in Game Theory, pp. 1.

Aumann, R. J. and S. Hart (2003). “Long Cheap Talk”, Econometrica, Vol. 71, pp. 1619–1660.

Austen-Smith, D. (1990). “Information Transmission in Debate”, American Journal of Political Science, Vol. 34, pp. 124–152.

Axelrod, R. (2006). “The Evolution of Cooperation”. Revised Edition, Basic Books.

Baliga, S. and S. Morris (2002). “Coordination, Spillovers and Cheap-Talk”, Journal of Economic Theory, Vol. 105, pp. 450–468.

Baliga, S. and T. Sjostrom (2004). “Arms Races and Negotiations”. Review of Economic Studies, Vol. 71, pp. 351–369.

Baliga, S., D. Lucca and T. Sjostrom (2011). “Domestic Political Survival and International Conflict: Is Democracy Good for Peace?”, The Review of Economic Studies, Vol. 78(2), pp. 458–486.

Baliga, S. and T. Sjostrom (2012). "The Strategy of Manipulating Conflict," American Economic Review, American Economic Association, Vol. 102(6), pp. 2897–2922.

Bandyopadhyay, S. and M. Oak (2010). “Conflict and Leadership: Why is There a Hawkish Drift in Politics?”. Working paper, University of Birmingham.

Banks, J. and R. Calvert (1992). “A Battle-of-the-Sexes Game with Incomplete Information”, Games and Economic Behavior, Vol. 4, pp. 347–372.

Bates, R., D. L. Epstein, J. Goldstone, I. Kristensen and S. O’Halloran (2006). “Democratic Transitions”, American Journal of Political Science, Vol. 50(3), pp. 551–569.

Bausch, A. W. (2013). “Democracy and Conflict: An Experiment”. Working paper, New York University.

Bellows, J. and E. Miguel (2008). "War and Local Collective Action in Sierra Leone". UC Berkeley.

Bennett, D. S. (2006). “Towards a Continuous Specification of the Democracy-Autocracy Connection”, International Studies Quarterly, Vol. 50, pp. 513–537.

Benoit, J. P. and V. Krishna (1985). "Finitely Repeated Games", Econometrica, Vol. 53(4), pp. 905.

Besley, T. J. and M. Ghatak (2010). “Property Rights and Economic Development “ in D. Rodrik and M. Rosenzweig (ed.s) Handbook of Development Economics V, North Holland.

Besley, T. J. and T. Persson. (2008). "The incidence of civil war: theory and evidence". Working paper, London, London School of Economics.

Blattman, C. and E. Miguel (2010). “Civil War”, Journal of Economic Literature, Vol. 48, pp. 3–57.

Bornstein, G., D. Budescu and S. Zamir, (1997). “Cooperation in Intergroup, N-Person, and Two-Person Games of Chicken”, Journal of Conflict Resolution, Vol. 41(3), pp. 384–406.

Boyer, Y. and J. Lindley-French (2012). “The Oxford Handbook of War”. Oxford, Oxford University Press.

Bueno de Mesquita, B. and R. M. Silverson (1995). “War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability”, American Political Science Review, Vol. 89, pp. 841–855.

Bueno de Mesquita, B., J. D. Morrow, R. M. Silverson and A. Smith (1999). “An Institutional Explanation of the Democratic Peace”, American Political Science Review, Vol. 93, pp. 791–807.

Cabrales, A., R. Nagel and R. Armenter (2007). “Equilibrium selection through incomplete information in coordination games: an experimental study”, Experimental Economics, Vol. 10(3), pp. 221–234.

Carecon (2011). Lecture 5: The Determinants of Conflict. [Online]. Available at: <http://carecon.org.uk/UCT/Lecture%205%20Determinants%20of%20Conflict%20notes.pdf> [Accessed 18 March 2014].

Carlsson, H. and E. van Damme (1993). “Global Games and Equilibrium Selection”, Econometrica, Vol. 61, pp. 989–1018.

Carter, J. R. and C. H. Anderton (2001). “An experimental test of a predator‐prey model of appropriation”, Journal of Economic Behaviour and Organization, Vol. 45, pp. 83–97.

Cason, T. N. and V. L. Mui (2007). “Communication and Coordination in the Laboratory Collective Resistance Game”, Experimental Economics, Vol. 10, pp. 251–267.

Cason, T. N. and V. L. Mui (2009). “Coordinating Collective Resistance through Communication and Repeated Interaction”. Working paper, Monash University.

Charness, G. (2000). “Self-Serving Cheap Talk: A Test of Aumann’s Conjecture”, Games and Economic Behaviour, Vol. 33(2), pp. 177–194.

Clark, K., S. Kay and M. Sefton (2001). “When Are Nash Equilibria Self-Enforcing? An Experimental Analysis”, International Journal of Game Theory, Vol. 29, pp. 495–515.

Coleman, E. A. and M. C. Lopez (2010). “Reconstructing Cooperation from Civil Conflict: Experimental Evidence from Colombian Development Policy”. Working paper, Florida State University.

Collier, P. (2001). “Implications of Ethnic Diversity”, Economic Policy, Vol. 32(16), pp. 127–166.

Colombier, N., D. Masclet, D. Mirza, and C. Montmarquette (2009). “Global Security Policies Against Terrorism and the Free Riding Problem: An Experimental Approach”. Working paper, University of Rennes.

Cooper, R., D. DeJong, R. Forsythe, and T. Ross (1992). “Communication in coordination games”, Quarterly Journal of Economics, Vol. 107, pp. 739–771.

Cooper, R., D. DeJong, R. Forsythe and T. Ross (1996). “Cooperation without reputation: experimental evidence from prisoner’s dilemma games”, Games of Economic Behavior, Vol. 12, pp. 187–318.

Cooper, R. (1999). “Coordination Games: Complementarities and Macroeconomics”. Cambridge, Cambridge University Press.

Cornand, C. (2006). “Speculative attacks and informational structure: an experimental study”, Review of International Economics, Vol. 14, pp. 797–817.

Cramer, C. (2003). “Does inequality cause conflict?”, Journal of International Development, Vol. 15(4), pp. 397.

Crawford, V. and J. Sobel (1982). “Strategic Information Transmission”, Econometrica, Vol. 50, pp. 1431–1452.

Crawford, V. (1998). “A Survey of Experiments on Communication via Cheap Talk”, Journal of Economic Theory, Vol. 78, pp. 286–298.

Dal Bó, E. and P. Dal Bó (2004). "Workers, Warriors, and Criminals: Social Confict in General Equilibrium". Unpublished paper, Stanford University and NBER.

Downs, G. and D. M. Rocke (1994). “Conflict, Agency and Gambling for Resurrection: The Principal-Agent Model Goes to War”, American Journal of Political Science, Vol. 38, pp. 362–380.

Duffy, J. and M. Kim (2005). “Anarchy in the laboratory (and the role of the state)”, Journal of Economic Behavior and Organization, Vol. 56, pp. 297–329.

Duffy, R. D., B. A. Allen and B. J. Dik (2010). “The presence of a calling and academic satisfaction: Examining potential mediators”, Journal of Vocational Behavior, Vol. 79 (1), pp. 74–80.

Durham, Y., J. Hirshleifer and V. L. Smith (1998). “Do the Rich Get Richer and the Poor Poorer? Experimental Tests of a Model of Power”, American Economic Review, Vol. 88, pp. 970–983.

Eidelson, R. J. and J. I. Eidelson (2003). "Dangerous ideas: Five beliefs that propel groups toward conflict", American Psychologist, Vol. 58(3), pp. 182–192.

Engelmann, D. and M. Strobel (2004). "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments", The American Economic Review, Vol. 94(4), pp. 857–869.

Epstein, D., R. Bates, J. Goldstone, I. Kristensen and S. O'Halloran (2006). "Democratic transitions", American Journal of Political Science, Vol. 50(3), pp. 551–569.

Esteban, J. and D. Ray (1994). "On the Measurement of Polarization", Econometrica, Vol. 62(4), pp. 819–851.

Esteban, J. and D. Ray (2008). "On the Salience of Ethnic Conflict", American Economic Review, Vol. 98, pp. 2185–2202.

Farrell, J. and M. Rabin (1996). "Cheap talk", Journal of Economic Perspectives, Vol. 10, pp. 103–118.

Fearon, J. (1994). "Domestic Political Audiences and the Escalation of International Disputes", American Political Science Review, Vol. 88, pp. 577–592.

Fearon, J. and D. Laitin (2003). "Ethnicity, insurgency, and civil war", American Political Science Review, Vol. 97(1), pp. 75–90.

Fehr, E. and K. M. Schmidt (1999). "A Theory of Fairness, Competition and Cooperation", The Quarterly Journal of Economics, Vol. 114 (3), pp.817–868.

Fischbacher, U. (2007). "z-Tree: Zurich toolbox for ready-made economic experiments", Experimental Economics, Vol. 10, pp. 171–178.

Forges, F. (1990). "Equilibria with Communication in a Job Market Example", The Quarterly Journal of Economics, Vol. 105, pp. 375–398.

Forsythe, R., J. Kennan, and B. Sopher (1991). “An experimental analysis of strikes in bargaining games with one-sided private information”, American Economic Review, Vol. 81, pp. 253–270.

Friedman, J. (1971). "A non-cooperative equilibrium for supergames", Review of Economic Studies, Vol. 38(1), pp. 1–12.

Fudenberg, D. and E. Maskin (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information", Econometrica, Vol. 54(3), pp. 533.

Garfinkel, M. R. and S. Skaperdas (2012). “The Oxford Handbook of the Economics of Peace and Conflict”. New York, Oxford University Press USA.

Ghosh, S., P. Postl and J. Roy (2011). “Cheap Talk by Partisan Media and Referendum Outcomes”, mimeographed, Department of Economics, University of Birmingham.

Greiner, B. (2015). “Subject pool recruitment procedures: organizing experiments with ORSEE”, Journal of the Economic Science Association, Vol. 1, pp. 114–125.

Grossman, H. I. and M. Kim (1995). “Swords or plowshares? A theory of the security of claims to property”, Journal of Political Economy, Vol. 103, pp. 1275–1288.

Guth, W. and B. Kalkofen, (1989). “Unique solutions for strategic games – Equilibrium selection based on resistance avoidance”, Lecture Notes in Economics and Mathematical Systems, No. 328.

Guth, W. (1992). “Equilibrium Selection by Unilateral Deviation Stability”, Rational Interaction, Essays in Honor of John C. Harsanyi, R. Selten (ed.), Heidelberg: Springer-Verlag, pp. 161–189.

Halpern, J. Y. (1986) “Reasoning about Knowledge: An Overview,” in Reasoning about Knowledge, J. Y. Halpern (ed.), Morgan Kaufmann, pp. 1–18.

Harsanyi, J. C. (1973). “Games with Randomly Disturbed Payoffs: A New Rationale for Mixed Strategy Equilibrium Points,” International Journal of Game Theory, Vol. 2, pp. 1–23.

Harsanyi, J. C. and R. Selten (1988). “A General Theory of Equilibrium Selection in Games”. Cambridge MA, MIT Press.

Heinemann, F., R. Nagel and P. Ockenfels (2004). “The theory of global games on test: experimental analysis of coordination games with public and private information”, Econometrica, Vol. 72(5), pp.1583–1599.

Hess, G. and A. Orphanides (1995). “War Politics: An Economic, Rational-Voter Framework”, American Economic Review, Vol. 85, pp. 828–846.

Hirshleifer, J. (1991). “The paradox of power”, Economics and Politics, Vol. 3, pp. 177–200.

Hoff, K. and J. E. Stiglitz (2001). “Modern Economic Theory and Development”, Frontiers of Development Economics: The Future in Perspective, G. Meier and J. Stiglitz (eds.), Oxford, Oxford University Press, pp. 389–459.

Huddy, L., R. Jervis and D. O. Sears (2003). “The Oxford Handbook of Political Psychology”. New York, Oxford University Press USA.

Hume, D. (1739). “A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning Into Moral Subjects”, London, Thomas Longman. Edited by L.A. Selby-Bigge, Oxford, Clarendon Press, 1888, 2nd edition, 1978.

Italiots (2012). Vision of Humanity: Peace’s Index. [Online]. Available at: <http://italiots.wordpress.com/2012/06/17/> [Accessed 18 March 2014].

Jackson, M. and M. Morelli (2005). “Political Bias and War”, mimeographed, Caltech and OSU.

Janis, I. L. and L. Mann (1977). “Decision-making: a psychological analysis of conflict, choice, and commitment”. New York, Free Press.

Jervis, R. (1976). “Perception and Misperception in International Politics”. Princeton N.J., Princeton University Press.

Jervis, R. (1978). “Cooperation under the Security Dilemma”, World Politics, Vol. 30(2), pp. 167–214.

Kagel, J. H. and A. E. Roth (1997). “The Handbook of Experimental Economics”. Princeton N.J., Princeton University Press.

Kahneman, D., J. L. Knetsch and R. H. Thaler, (1986). “Fairness as a Constraint on Profit Seeking: Entitlements in the Market”, The American Economic Review, Vol. 76(4), pp. 728–741.

Kahneman, D., J. L. Knetsch and R. H. Thaler (1990). "Experimental Tests of the Endowment Effect and the Coase Theorem", Journal of Political Economy, Vol. 98(6), pp. 1325–1348.

King, G., R.O. Keohane and S.Verba (1994). “Designing Social Inquiry: Scientific Inference in Qualitative Research”. Princeton N.J., Princeton University Press.

King, G., J. Honaker, A. Joseph, and K. Scheve (2001). “Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation”, American Political Science Review, Vol. 95(1), pp. 49–69.

Krishna, V. and J. Morgan (2004). “A Model of Expertise”, The Quarterly Journal of Economics, Vol. 116, pp. 747–775.

Kydd, A. (1997). “Game Theory and the Spiral Model”, World Politics, Vol. 49(3), pp. 371–400.

Leeds, B. A. (1999). “Domestic Political Institutions, Credible Commitments, and International Cooperation”, American Journal of Political Science, Vol. 43(4), pp. 979–1002.

Levy, G. (2004). “Anti-herding and Strategic Consultation”, European Economic Review, Vol. 48(3), pp. 503-525.

Levy, G. and R. Razin (2004). “It takes Two: An Explanation for the Democratic Peace”, Journal of the European Economic Association, Vol. 2, pp. 1–29.

Mansfield, E. and J. Snyder (2005). “Electing to Fight: Why Emerging Democracies go to War”. Cambridge MA, M.I.T. Press.

Matthews, S. and A. Postlewaite (1989). “Pre-play Communication in Two-person Sealed-bid Double Auctions”, Journal of Economic Theory, Vol. 48, pp. 238–263.

McBride, M. and S. Skaperdas (2009). “Conflict, Settlement and the Shadow of the Future”. Working paper, University of California Irvine.

McDaniel, T. M. (2011). “Coordination in games with incomplete information: experimental results,” International Game Theory Review, Vol. 13(4), pp. 461–474.

Mehta, J., C. Stanner and R. Sugden (1994). “Focal points in pure coordination games: an experimental investigation”, Theory and Decision, Vol. 36, pp.163–185.

Midlarsky, M.I. and K. Roberts (1985). “Class, state, and revolution in Central America: Nicaragua and El Salvador compared”, Journal of Conflict Resolution, Vol. 29, pp. 163–193.

Mitra, A. and D, Ray (2014). “Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India”, Journal of Political Economy, Vol. 122(4), pp. 719–765.

Morris, S. and H. Shin (2002). “Heterogeneity and Uniqueness in Interaction Games”, mimeographed, Yale University.

Morris, S. and H. Shin (2003). “Global Games: Theory and Applications” in M. Dewatripont, L. Hansen and S. Turnovsky (eds.) Advances in Economics and Econometrics: Proceedings of the 8th World Congress of the Econometric Society, Cambridge, Cambridge University Press.

Mukand, S. and S. Majumdar (2004). “Policy Gambles”, The American Economic Review, Vol. 94(4), pp. 1207–22.

Muller, E.N. and M.A. Seligson (1987). “Inequality and insurgency”, American Political Science Review, Vol. 81(2), pp. 425–452.

Munster, J. (2007). “Simultaneous inter- and intra- group conflicts”, Economic Theory, Vol. 32, pp. 333.

O’Neill, B. (1999). “Honor, Symbols and War”. Ann Arbor, University of Michigan Press.

Opreaa, R., G. Charnessa and D. Friedman (2014). “Continuous time and communication in a public-goods experiment”, Journal of Economic Behavior & Organization, Vol. 108, pp. 212–223.

Ostrom, E., J. Walker, and R. Gardner (1992). “Covenants with and without a sword: self-governance is possible”, American Political Science Review, Vol. 86(2), pp. 404–417.

Ottaviani, M. and P. Sorensen (2006). “Professional Advice”, Journal of Economic Theory, Vol. 126(1), pp. 120–142.

Palfrey, T. and H. Rosenthal (1991). “Testing for effects of cheap talk in a public goods game with private information”, Games of Economic Behavior, Vol. 3, pp.183–220.

Peceny, M., C. Beer and S. Sanchez-Terry (2004). “Dictatorial Peace?”, American Political Science Review, Vol. 96(1), pp. 15–26.

Rahim, M. A. (2010). “Managing Conflict in Organizations”. Transaction Publishers, pp. 17.

Ray, I and C. Ganguly (2006). “Cheap Talk: Basic Models and New Developments”, mimeographed, Department of Economics, University of Birmingham.

Ray, D., J. M. Baland, and O. Dagnelie (2007). “Inequality and Inefficiency in Joint Projects”, The Economic Journal, Vol. 117(522), pp. 922–935.

Reiter, D. (1995). “Exploding the Powder Keg Myth: Pre-emptive Wars Almost Never Happen”, International Security, Vol. 20, pp. 5–34.

Roth A. (1987). “Bargaining phenomena and bargaining theory”, in “Laboratory Experimentation in Economics: Six Points of View” (A. Roth, Ed.), Cambridge, Cambridge University Press.

Rubinstein, A. (1979). "Equilibrium in supergames with the overtaking criterion", Journal of Economic Theory, Vol. 21(1).

Rubinstein, A. (1989). “The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge”, American Economic Review, Vol. 79, pp. 385–391.

Rummel, R. J. (1977). “Understanding Conflict and War, Volume 3: Conflict in Perspective”. Beverly Hills, California, Sage Publications.

Schelling, T. C. (1960). “The Strategy of Conflict”. Cambridge MA, Harvard University Press.

Schmidt, D., R. Shupp, J. M. Walker, and E. Ostrom (2003). “Playing Safe in Coordination Games: The Role of Risk Dominance, Payoff Dominance, Social History, and Reputation”, Games and Economic Behavior, Vol. 42, pp. 281–299.

Smith, A. C., D. B. Skarbek, and B. J. Wilson (2009). “Anarchy, Groups, and Conflict: An Experiment on the Emergence of Protective Associations”. Working paper, George Mason University.

Straub, P. G. (1995). “Risk Dominance and Coordination Failures in Static Games”, The Quarterly Review of Economics and Finance, Vol. 35(4), pp. 339–363.

Sugden, R. (1995). “A Theory of Focal Points”, The Economic Journal, Vol. 105(430), pp. 533-550.

Tangeras, T. (2008). “Democracy, Autocracy and the Likelihood of International Conflict”, mimeographed, Stockholm.

Valley, K., L. Thompson, R. Gibbons, and M. Bazerman (1996). “Is Talk Really Cheap? Outperforming Equilibrium Models of Communication in Bargaining Games”. Working Paper pp. 96–022, Harvard Business School.

Van Huyck, J. B., R. Battalio, and R. Beil (1990). “Tacit Coordination Games, Strategic Uncertainty, and Coordination Failures”, American Economic Review, Vol. 80, pp. 234–248.

Walsh, D. (2010). “WikiLeaks Cables: Pakistan Opposition ‘Tipped Off’ Mumbai Terror Group.” The Guardian, December 1.

Weingast, B. (1997). “The political foundations of democracy and the rule of law”, American Political Science Review, Vol. 91, pp. 245–263.

Whitt, S. and R. K. Wilson (2007). “The Dictator Game, Fairness and Ethnicity in Post-war Bosnia”, American Journal of Political Science, Vol. 51, pp. 655–668.

Zizzo, D. J. (2010). “Experimenter demand effects in economic experiments”, Experimental Economics, Vol. 13, pp. 75.