# Bargaining Impasse: The Role of Asymmetric Outside Options, Self-serving Biases and Participation Requirements

Inauguraldissertation
zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München

2005

vorgelegt von

Simone Kohnz



| | |
|---|---|
| Referent: | Prof. Dr. Klaus Schmidt |
| Korreferent: | Prof. Dr. Steffen Huck |
| Promotionsabschlussberatung: | 08. Februar 2006 |

## Acknowledgements

While preparing this thesis, I have received much support - through encouragement and advice - from many colleagues at the Munich Graduate School of Economics, the Economics Department, and in particular the Seminar for Economic Theory, at the Ludwig-Maximilians-Universität München, and the Economics Department at University College London.

First of all, I want to thank my supervisor Professor Klaus Schmidt. Whenever I needed advice, he provided me with invaluable suggestions and comments. This thesis benefited enormously from his tremendous experience and knowledge. I would also like to thank Professor Steffen Huck who supervised me during my visit to University College London and accepted to be my second examiner.

Numerous other people have helped improve the work on this thesis. I am especially grateful to Melanie Lührmann who has provided me with detailed, critical, and very helpful comments on all parts of this thesis and who has been a great friend during these, sometimes rather difficult, times. Many thanks are also due to Chantelle Bramley, Armin Falk, Alfred Garloff, Florian Herold, Hannah Hörisch, Katri Mikkonen, Natalie Timan, Katharina Sailer, Tobias Seidel, Matthias Weiss and Joachim Winter for many fruitful discussions on various versions of different chapters of this thesis and for even more coffee and ice-cream breaks.

Financial support from the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged.

Last, but by no means least, I want to thank my parents for their great support during these (many) years.

Simone Kohnz

# Contents

# List of Tables

# List of Figures

# Preface

Bargaining is important on all levels of social interaction from the small quarrels among friends and family to the important negotiations between states. The costs of its impasse can be substantial, consider for example, the amounts spent privately and publicly on civil litigation or the costs of strike and lockout. Understanding why and when bargaining fails is thus one of the major concerns in social sciences. This dissertation analyses bargaining from two very distinct angles. The first two chapters deal with strategic behaviour in *bilateral* bargaining situations. There, I analyse bargaining breakdown when two individuals bargain over the division of a fixed surplus. I depart from the neo-classical assumption of pure self-interest and provide a theoretical framework and empirical evidence for a self-serving bias in fairness perceptions. Contrary, the third chapter investigates, under standard behavioural assumptions, how the success of *multinational* bargaining over the provision of a global public good is affected by a compulsory minimum number of participating countries. It characterises the optimal minimum number depending on the model parameters and presents some comparative statics results.

Each economic model and therewith its predictive power is based on specific behavioural assumptions. For decades, economists used to postulate that agents behave in a purely self-interested way, taking only their own material well-being into consideration. This seems a very restrictive assumption and economists have not always been thinking like this. In "The Theory of Moral Sentiments" (1759, chapter 1), Adam Smith acknowledges that

> "How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it."

Economists recently started to reassess the fundamental behavioural assumption of pure self-interest and its implications. At the forefront were experimental economists who questioned whether the predictions of the neo-classical assumption of pure self-interest would be compatible with behaviour in laboratory experiments. In the last two, three decades, experimental economists have conducted a vast variety of experimental studies. Not surprisingly, these studies find that a large fraction of the agents do not behave as predicted by self-interest. For an extensive summary on the experimental findings, see chapter 2 in Camerer (2003).

These findings inspired microeconomic theorists to find behavioural assumptions that can explain the observed patterns in a rigorous and clear manner. By now, various approaches exist that model the experimental evidence. All of these models embed social comparison processes in preferences and are thus commonly referred to as other-regarding preferences. A comprehensive survey of the literature can be found in Fehr and Schmidt (2003). These new behavioural models see many economic phenomena from a different angle and offer solutions to longlasting puzzles in a novel and intriguingly intuitive way. For example, in the field of contract theory, it is difficult to explain why many contracts seem to be deliberately left incomplete. The theory of incentives suggests that it is best to attach the remuneration for a specific task closely to observable measures of effort. Concerns for fairness and reciprocity might render the standard incentive contract less attractive and improve the workings of incomplete contracts relying on the goodwill of each party, see Fehr, Klein, and Schmidt (2001). Other-regarding preferences also allowed economists to take a fresh look at other decision situations like incentives in teams, ownership structures or the provision of public goods.

In this light, the first chapter proposes a new approach to bilateral bargaining breakdown. It examines the connection of bargaining breakdown and self-serving biases in fairness perceptions. There is strong empirical evidence that people exhibit self-serving biases concerning their fairness judgements in bargaining situations with asymmetric outside options. Moreover, the psychological literature suggests that such a self-serving bias can be a driving force of bargaining impasse. I provide a theoretical framework for analysing the behaviour of self-servingly biased agents in simple bargaining situations. I build on the notion of inequity aversion and extend it to incorporate self-serving biases due to asymmetric outside options. To take account of agents' ignorance concerning their biases, I distinguish between sophisticated and naive agents. That is, those agents who understand their bias and those who do not. Then, I apply this framework to analyse the behaviour of naive and sophisticated biased agents in ultimatum and investment games. For ultimatum bargaining with complete information, I find that bargaining can only break down if biased proposers are not aware of their self-serving bias. In the incomplete information case, the propensity of bargaining breakdown is higher with naive than with sophisticated agents. In the investment game, a self-serving bias leads to reduced participation, but naiveté about the bias increases participation.

Behavioural assumptions drive the predictions of any economic model. It is thus of great importance to verify our underlying assumptions. One way of doing this is to examine behaviour of individuals in a decision situation in the laboratory, thus eliminating most of the exogenous variation to the situation. Comparing the accuracy of the predictions of different behavioural assumptions gives an indication which assumption fits the decision situation better. The second chapter thus explores whether the predictions generated for self-servingly biased agents fit experimental data better than other behavioural assumptions. To this aim, I run a standard ultimatum experiment with varying asymmetric outside options. The focus of the second chapter is thus on bilateral bargaining breakdown from the experimental perspective. I find that offers and minimal acceptable offers increase/ decrease with an increase in the outside option of the responder/ proposer. The evolution of offers and minimal acceptable offers is not compatible with the predictions generated by the assumption of purely self-interested agents. However, it is consistent with inequity averse agents with and without a self-serving bias. Furthermore, I find that rejection rates increase with the introduction of asymmetry. Neither pure self-interest, nor inequity aversion can explain this increase. However, as shown in the first chapter, a self-serving bias in the assessment of fairness can induce bargaining breakdown in the presence of asymmetries. The evidence found is in line with the theoretical predictions from chapter one. Moreover, I find that the propensity to offer/ demand the equal split of the pie decreases rapidly with the introduction of a small asymmetry. This holds in particular, when the proposer is endowed with the larger outside option implying that the power to propose is bundled with the advantage in terms of outside options. It further supports the hypothesis that agents change their perception of the fair offer or demand. Apart from the main analysis, I find evidence that subjects might get biased by the nature of the previous decision situations. This is an interesting effect as it suggests that subjects do not regard each situation independently, but that in particular situations they are influenced by previous decision circumstances.

In the third chapter, I switch perspective. Instead of examining bargaining between individuals, the focus is now on outcomes when many countries negotiate together. International agreements for the provision of global public goods gather more and more political and economic importance. Throughout the years the number of international agreements of any sort has been increasing. In many of the agreements, the provision of some public good is determined, mostly related to biodiversity and other environmental issues, human rights and the rules of warfare and arms control. Among the most prominent agreements of the late 20th century are the Kyoto Protocol regulating emissions of carbondioxides and the installation of an international criminal court. The provision of global public goods as regulated in such agreements is of great importance to present and future generations around the world. However, there is no mechanism to oblige sovereign states to participate and comply with the agreements. In case of the Kyoto Protocol, it has long been doubtful whether sufficiently many countries would join the agreement to render it legally binding for all participating countries. Only in 2004, seven years after the Conference of Kyoto, did Russia promise its ratification restoring the protocol to life.

Hence, it is important to understand the features of the bargaining process that hinder or enhance the chances of a successful agreement. The third chapter analyses the role of ratification quotas in multilateral agreements on the provision of a transnational public good. The setting is applied to the example of emission reduction. The general idea is the trade-off that the higher is the quota, the lower is the level of emissions in case the agreement comes into force, but the higher is also the risk of failure. I propose a three stage international bargaining game where countries first determine a ratification quota. Then, they decide whether to ratify and finally they decide on the emission of a global pollutant. With countries rather than individuals being the participants of the bargaining process, I adopt the behavioural assumption of pure self-interest. In a setting with incomplete information, two country types and a binary contribution to the provision, I examine the differences between simultaneous and sequential ratification. When the benefits from emission of both types are smaller than the social costs, the outcome in the simultaneous case is essentially identical to the sequential case. The optimal quota is 100% and achieves the first best. With the high type's benefits exceeding the social costs, I find that the optimal quota is as small as possible, if ratification is simultaneous. In the sequential ratification case, I cannot determine the optimal quota. However, I find that the aggregate expected surplus decreases with respect to the simultaneous case.

Each of the following chapters is a self-contained paper with its own introduction and appendix. This implies that each chapter can be read independently of the other two.

# Chapter 1

# Self-Serving Biases in Bargaining: Explaining Impasse

## 1.1 Introduction

There is a large body of experimental literature, both in psychology and economics, that finds self-serving biases in judgements of fairness. This literature suggests that self-serving biases are a driving force of bargaining impasse. It is evident that bargaining is important on all levels of social interaction from the small quarrels among friends and family to the big negotiations between states. The costs of its impasse can be substantial, consider for example, the amounts spent privately and publicly on civil litigation or the costs of strike and lockout. Understanding why bargaining fails in some cases is thus one of the major concerns in social sciences. A self-serving bias settles itself in a notion of fairness that, mostly unconsciously, tends to favour an agent. It is intuitive that in a situation where agents have different notions of fairness and moreover, are not aware of these differences, bargaining might fail. In the economics literature, there has been no attempt, so far, to model self-serving biases theoretically and to explore its impact on bargaining breakdown. This paper tries to do this by extending the notion of inequity aversion in the presence of asymmetric outside options and applying it to ultimatum bargaining and investment games.

For a self-serving bias to occur, the psychological literature suggests that "there needs to be some form of asymmetry in how the negotiation environment is viewed", Babcock and Loewenstein (1997, p. 119). In real life, one hardly finds a perfectly symmetric negotiation environment. In particular, most situations are characterised by asymmetric outside options. These occur, for example, in wage bargaining where the employer might have the choice between several different candidates whereas the employee's outside option is unemployment. Yet similarly, asymmetric outside options are present when countries negotiate emission targets of a global pollutant, those damaging effects vary across countries. Furthermore, there is a close link between outside options and individual wealth levels. Asymmetry in terms of outside options plays a prominent role and is maybe the most natural case triggering self-serving biases. This paper therefore focuses on self-serving biases induced by asymmetric outside options.

Economists usually analyse bargaining games with the neo-classical assumption of purely self-interested agents. However, experimental evidence suggests that a large fraction of agents do not behave as classical economic theory predicts. Simple set-ups such as dictator, ultimatum or investment games, suggest that subjects compare their payoff with the other participants' payoffs. For an extensive summary on the experimental findings, see chapter 2 in Camerer (2003). There exist various approaches to model the experimental evidence. All of these models embed social comparison processes in preferences. A comprehensive survey of the literature can be found in Fehr and Schmidt (2003). Here, I follow the approach of inequity aversion by Fehr and Schmidt (1999) where agents dislike income inequity.[1] Comparing monetary payoffs, agents base their judgement as to whether an outcome is considered as equitable on a reference allocation. Fehr and Schmidt argue that in a symmetric setting a natural reference outcome is one which attributes the same monetary payoff to all agents (*Equal Split*). With the introduction of asymmetric outside options this reasoning is no longer applicable. The *Equal Split* is just one among many other possible reference allocations like, for example, *Split the Difference* which advocates an equal split of the entire cake minus the sum of outside options. On which of the various reference allocations an agent is likely to base her fairness judgement is an empirical question. Yet, a self-serving bias would imply that with asymmetric outside options agents adopt a fairness perception that favours them in monetary terms.

The extension I propose allows inequity averse agents to base their decision on reference allocations different from the *Equal Split*. I render the reference allocation of the agents linearly dependent on the difference in outside options between two agents. The strength with which this difference influences the reference point can vary across agents. It serves as a measure of the extent to which the fairness perception favours the agent. According to Dahl and Ransom (1999, p. 703), agents that are self-servingly biased "...subconsciously alter their fundamental views about what is fair in a way that benefits their interests". Hence, a self-serving bias is characterised by two features: First, it settles itself in a notion of fairness that tends to favour the agent, i.e. that leaves the agent with a relatively big monetary payoff. Second, agents are not aware of their self-serving biases. I separate these two features of self-serving biases to analyse the influence of each component separately. An agent is *biased*, if she has a reference allocation that attributes a larger allotment to her than the reference allocation of her partner agent. To capture the second feature of a self-serving bias, namely that people are ignorant about the bias, I distinguish between *sophisticated* agents

[1] Inequity aversion provides a simple and sparse representation of the comparison processes that nevertheless captures a lot of the experimental findings.

who understand that their fairness notion favours themselves, and *naive* agents who have no such understanding.

Within this extended framework of inequity aversion, I analyse ultimatum bargaining and stylised investment games. These simple bargaining games deliver the ingredients to more sophisticated negotiation environments. It is thus interesting to understand in a first step how self-serving biases work in these simple settings. In an ultimatum game, a proposer and a responder bargain over the division of a fixed pie. The proposer announces a division which the responder can accept or reject. If he accepts, the pie is divided according to the proposed rule. If he rejects, each player gets an outside option, known to both agents. With purely self-interested agents, as well as with standard inequity averse agents, there will be no bargaining breakdown. With the mere introduction of differing evaluations of what allocation is fair, this does not change. As long as the proposer knows the fairness perception of the other agent, she prefers to offer a share that the responder is willing to accept rather than to get her outside option. Accordingly, as long as the biased agents are aware of their bias, agents reach an agreement. If instead the proposer is biased and naive, then there are circumstances where the bargain breaks down. The reasoning is straightforward. The respondent is willing to accept any offer that is above a certain threshold. The threshold level depends on the fairness perception of the respondent. Sophisticated and biased proposers predict the threshold correctly, while some naive and biased proposers underestimate it. Therefore, whenever a naive proposer offers the underestimated threshold level in equilibrium, the bargain fails.

In a simple version of the investment game, one agent decides whether she wants to invest her outside option and participate in a dictator game. If she decides to participate, the other agent then determines how to split a fixed surplus; otherwise both receive their outside option. Dictators, whose reference allocation favours themselves more, tend to give less to the recipient. If their fairness perception is sufficiently extreme, the first agent refuses to participate. However, if this agent is naive, her participation becomes more likely. In many employment contexts, agents are at least temporarily locked into a relationship where one agent cannot influence how a task or burden is shared between herself and her superior. The superior can take decisions that could be perceived as very unfair by the other employee. If the employee would have foreseen the unfair behaviour of the superior, she might have abstained from the job beforehand and might have taken up her outside option.

Related to the present paper is Konow (2000). He presents a model that incorporates, in addition to standard material utility, a genuine value of fairness intertwined with an incentive to change beliefs about the fairness concept. He postulates that there is an objective fairness concept from which agents voluntary deviate to favour themselves. In contrast, the present paper takes the belief about the fairness concept as given and analyses how this belief induces bargaining breakdown. Another related paper by **?** suggests that self-servingly biased agents have entered a tough state of mind vis-a-vis someone else. **?** characterise a class of bargaining mechanisms under which a population evolves that exhibits some moderate degree of toughness. They identify the underlying trade-off that toughness decreases the average probability of a bargain, but improves the terms of trade. In contrast, this paper examines how toughness influences behaviour on each bargaining stage. Finally, Frohlich, Oppenheimer, and Kurki (2004) extend equity aversion in a way similar to our extension. They introduce the concept of "just deserts" in the context of dictator games with preceding production. There, agents suffer when their inputs to the surplus are larger/ smaller than their final shares. The dictator faces a trade-off between material payoffs, equality and just deserts. However, their extension of inequity aversion differs substantially in that they assume that different norms are conflicting with each other. In their model, agents trade-off disutility from inequality with disutility from a deviation to just deserts. Whereas the present model postulates that agents adhere to one norm which depends on the context of the situation.

In the next section, I propose an extension of inequity aversion that incorporates heterogeneity in fairness perceptions and self-serving biases. The framework is first applied to ultimatum bargaining games in section 1.3 and then to investment games with asymmetric outside options in section 1.4. Section 1.5 discusses experimental evidence. Section 1.6 concludes and suggests further paths of research.

## 1.2 An extension of inequity averse preferences

Inequity averse agents compare their monetary payoff with the payoff of members of a specific reference group. Within this reference group they dislike outcomes that they perceive as unequal or unfair. That is, they derive negative utility of a deviation from their reference allocation. The reference allocation of an agent with respect to another agent is defined by the pair of payoffs that she considers to be equal or fair. The utility of an agent depends on the reference allocation as well as the reference group. Both these determinants are considered exogenous in the model of Fehr and Schmidt. They argue that in an experiment all participants form the reference group. Furthermore, they postulate that in symmetric situations a natural reference allocation is one in which each agent gets the same payoff, the *Equal Split*.[2] Other consequentialist models like Bolton and Ockenfels (2000) and Charness and Rabin (2002) also postulate the *Equal Split* as reference allocation. Once asymmetry is introduced, there is no reason to believe that the *Equal Split* is the natural reference allocation. The asymmetry may lead to various reference allocations.[3]

[2] A study by Hennig-Schmidt (2002) finds that in symmetric ultimatum games the only allocation that is perceived as fair by both agents is the allocation were each agent gets an equal amount. She conducted a video experiment where groups of individuals decided about a distributional task. During the group discussion preceding the decision, in the symmetric ultimatum game only the *Equal Split* was mentioned as a fair outcome.

[3] An experiment by Messick and Sentis (1979) divided subjects into two groups. One group was told that they should imagine they had worked 7 hours and were to receive a certain amount of money for that.

In simple bargaining situations such as the ultimatum game, the mere allocation of roles introduces asymmetry that can induce a self-serving bias.[4] More powerful sources of asymmetry in bargaining environments, however, are asymmetric payoff possibilities or differing outside options for the agents. The present paper models self-serving biases due to asymmetric outside options. Different outside options are an important source of asymmetry. They are present, for example, in situations where an employer and a worker bargain over the worker's wage. But there are a lot of other day-to-day examples where people with different outside options have to decide about the distribution of a surplus. Furthermore, there is a resemblance between outside options and individual wealth levels in terms of their impact on fairness judgements. In their fairness statements, agents can be and often are guided by considerations concerning the difference in individual wealth levels. Even though individual wealth levels are not "destroyed" if the parties successfully bargain with each other, relative wealth levels might nevertheless determine the reference allocation in the bargaining situation. In this sense, part of the analysis can be transferred to self-serving biases induced by different wealth levels. Apart from this apparent omnipresence, asymmetric outside options are relatively easy to capture. First, it is an easily observable characteristic of the bargaining situation. Second, it can be measured quantitatively. Last, outside options can be altered in experimental set-ups and thus the predictions of the theory should be testable.

An easy and straightforward way of incorporating outside options into fairness considerations is to render the reference allocation linearly dependent on the difference in outside options. The reference allocation then has to obey

$$x_i - x_j = \gamma_i \left(\omega_i - \omega_j\right) \quad \forall i \neq j \tag{1.1}$$

where $x_i$ represents the monetary payoff of agent $i$, $\omega_i$ her outside option and $\gamma_i$ measures the extent to which the reference allocation favours the agent. This representation has the property that whenever we consider agents in a symmetric environment, the reference allocation is the *Equal Split*, independent of $\gamma_i$.

Suppose two agents $i, j$ can jointly generate a fixed surplus, which I normalise to 1 and which is strictly larger than the sum of the outside options $1 > \omega_i + \omega_j$. The reference

Subjects of the other group were told to imagine they had worked for 10 hours on the same task. All subjects were asked to state the fair payment for the ones that had worked for 10 hours. There were two prominent concepts of fairness, one that induced the same hourly wage and one that induced the same overall payment. Among the group of subjects who was told to have worked 7 hours the fraction of subject regarding the same total payment as fair was significantly larger than the fraction in the second group. Another experiment by Babcock, Loewenstein, Issacharoff, and Camerer (1995) allocated the roles of prosecutor and defendant in a juridical case to different individuals. They find that parties with the same information about the case come to different conclusions about what settlement is fair depending on their allocated roles. These are two examples of experiments finding several comcepts of fairness in asymmetric situations and furthermore, a self-serving bias in the assessment of the fair outcome.

[4] Proposers view themselves in a relatively more powerful role and therefore believe that they deserve more than their opponents. The respondents in contrast think that the distribution of roles should not affect the division of the cake, see Hoffman, McCabe, Shachat, and Smith (1994).

allocation of agent $i$ is uniquely determined by equation (1.1) and the restriction that $x_j = 1 - x_i$. Denote the pair of payoffs that solves these equations by

$$\left(x_i^f(\gamma_i), x_j^f(\gamma_i)\right) = \left(\frac{1 + \gamma_i(\omega_i - \omega_j)}{2}, \frac{1 - \gamma_i(\omega_i - \omega_j)}{2}\right),$$

where the superscript $f$ stands for fair. There are several outstanding reference allocations. The most prominent being the allocation where both agents receive equal monetary payoffs (*Equal Split*). This would imply a fairness parameter $\gamma_i$ of zero. A reference allocation that splits the difference between the surplus both agents can jointly generate and the sum of the outside options $\omega_i + \omega_j$ (*Split the Difference*) implies a parameter $\gamma_i$ of one. Furthermore, a parameter of $\gamma_i = \frac{1}{\omega_i + \omega_j}$ represents a reference allocation that divides the entire cake proportionate to the agents' outside options (*Proportional Split*). Still, one could think of any other value of $\gamma_i$ constituting a reference allocation.[5] In the case where agent $i$ has the larger outside option, $\omega_i > \omega_j$, the parameter range of $\gamma_i$ can be reduced to $\left[-\frac{1}{\omega_i - \omega_j}, \frac{1}{\omega_i - \omega_j}\right]$. The upper value signifies a reference point where agents consider it fair that agent $i$ gets the entire pie and the lower value where agent $j$ gets everything.[6] Incorporating this approach in the representation of inequity aversion yields preferences of the form

$$\begin{aligned} u_i(x) &= x_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_j - x_i - \gamma_i(\omega_j - \omega_i), 0\} \\ &\quad - \beta_i \frac{1}{n-1} \sum_{j \neq i} \max\{x_i - x_j - \gamma_i(\omega_i - \omega_j), 0\} \end{aligned}$$

with $\alpha_i \geq \beta_i \geq 0$ and $\beta_i < 1$. The utility parameters $\alpha_i$ resp. $\beta_i$ measure the loss for agent $i$ resulting from a deviation to her disadvantage resp. advantage from her reference point.

The reference allocation of a self-servingly biased person attributes a relatively big monetary allotment to herself. Moreover, it is often the case that a biased person believes her reference allocation to be impartial. I split up the notion of self-serving biasedness into these two components: (i) the bias itself and (ii) the belief about the bias.

**Definition 1** *An agent $i$ is **self-servingly biased** with respect to another agent $j$ if a higher monetary payoff is attributed to herself by her own reference allocation than by the reference allocation of agent $j$, i.e. $x_i^f(\gamma_i) > x_i^f(\gamma_j)$.*

[5] In experiments by Hennig-Schmidt (2002) where subjects played in groups and were allowed to discuss their decisions, *Equal Split*, *Split the Difference* and *Proportional Split* have been frequently characterised as fair allocations. A few other divisions of the cake were also paraphrased as fair, but much less often.

[6] Suppose an agent considers it to be fair that she gets the entire surplus. One could ask in which ways this agent is different from an agent who is purely self-interested. Contrary to the pure self-interest agent, the inequity averse agent engages in social comparison processes, regardless of the fact that she considers it fair to receive the entire pie. Thus, she nearly always suffers from disadvantageous inequity aversion. As a consequence, behavioural predictions are different, in general, from the predictions derived for the self-interest agent.

This implies that the agent with the relatively large outside option is self-servingly biased if she has a relatively large fairness parameter $\gamma_i$. Conversely, an agent with the relatively small outside option is self-servingly biased if she has a relatively small $\gamma_i$. Consider for example the two specific reference allocations of *Equal Split* and *Split the Difference*. Agents are self-servingly biased if the agent with the relatively large outside option regards *Split the Difference* as a fair outcome, while the agent with the relatively small outside option considers the *Equal Split* fair. This relativistic view of biasedness might sound unfamiliar. One might argue that whenever an agent considers it to be equitable that she gets the entire surplus herself, she is self-servingly biased. However, biasedness requires a point of comparison. There is no such exogenous "objective" comparison available in the context of bilateral bargaining. Therefore, biasedness is defined here in comparison to the reference allocation of the other agent. To rule out cases where an agent allocates less to herself than the opponent does, I restrict the parameter range such that $\gamma_j \in \left[-\frac{1}{\omega_i - \omega_j}, \gamma_i\right]$ for $\omega_i \geq \omega_j$. The agent with the relatively small outside option is thus bound to have a smaller fairness parameter than her opponent.

I distinguish between those agents who are aware of differing fairness notions among individuals and those who are not. In analogy to O'Donoghue and Rabin (1999), I call an agent naive who thinks her reference allocation is impartial. A naive agent assumes therefore that the other agent has the same fairness parameter as herself. In contrast, a sophisticated agent knows that her reference allocation differs from the one of her opponents. Moreover, she knows the exact fairness parameter of the other agent.[7] Denote the belief of agent $i$ about the fairness parameter of agent $j$ by $\widehat{\gamma}_{ij}$.

**Definition 2** *Agent $i$ is* ***naive*** *if she believes that agent $j$'s fairness parameter is the same as hers, that is $\widehat{\gamma}_{ij} = \gamma_i$. Agent $i$ is* ***sophisticated*** *if her belief about agent $j$'s fairness parameter is correct, that is $\widehat{\gamma}_{ij} = \gamma_j$.*

In the presence of naive agents the solution concepts of subgame perfection and Bayesian perfection become problematic as beliefs might not be correct in equilibrium. I therefore employ the concept of "perception perfect strategies" introduced by O'Donoghue and Rabin (2001) in the context of hyperbolic discounting. This concept merely requires that agents choose an action that maximises their payoff according to their beliefs. But it does not require, as the concept of subgame perfection or Bayesian perfection, that agents' beliefs are correct in equilibrium.[8] Denote with $U_i\left(s_i\left(\gamma_i, \widehat{\gamma}_{ij}\right)\right)$ the (expected) utility of agent $i$ resulting from the strategy $s_i \in A_i$ where $A_i$ signifies the strategy space for agent $i$.

[7]If we allow sophisticated agents to be uncertain about the exact value of the fairness parameter of the other agent, we get partial sophistication. The case with perfect sophisticates and perfect naives can be regarded as a benchmark.

[8]In the dynamic context of time inconsistency, the solution concept of perception-perfection requires that beliefs be dynamically consistent. This implies that agents believe their future actions to be optimal in any period and that action plans do not change across periods.

**Definition 3** *The strategy* $s_i^{pp}\left(\gamma_i,\widehat{\gamma}_{ij}\right)$ *is perception-perfect for a* $\left(\gamma_i,\widehat{\gamma}_{ij}\right)$*-agent if and only if* $s_i^{pp}\left(\gamma_i,\widehat{\gamma}_{ij}\right)\in\arg\max_{s_i}U_i\left(s_i\left(\gamma_i,\widehat{\gamma}_{ij}\right)\right)$.

The belief of a sophisticated agent is correct. Therefore, the perception perfect equilibrium coincides with the subgame perfect equilibrium resp. the Bayesian perfect equilibrium.

In the next sections, I analyse the behaviour of self-servingly biased inequity averse agents in ultimatum bargaining as well as simple investment games. In particular, I focus on the behaviour of self-servingly biased agents who are naive.

## 1.3 Ultimatum game

In an ultimatum game, a proposer and a responder bargain over the division of a fixed surplus of one.[9] The proposer ($P$) announces a division of the surplus $(1-s,s)$ where $s$ denotes the share offered to the responder. The responder ($R$) in turn accepts or rejects the proposal. If he accepts, then the surplus is divided according to the proposed rule. If he rejects, each player gets her or his outside option denoted by $\omega_i\geq 0$ for $i=P,R$. Both agents know the outside options of either player.

In the subgame perfect equilibrium under the assumption of purely self-interested agents, the proposer offers a division of the surplus of $(1-\omega_R,\omega_R)$ which is accepted by the respondent. Contrary, with inequity averse agents, the equilibrium offer depends upon the characteristics of the utility functions of the proposer as well as the responder. For the case of no outside options, Fehr and Schmidt (1999) characterise the equilibrium of the ultimatum bargaining with inequity averse agents. With complete information concerning the utility parameters $\alpha_R$ and $\beta_R$, the equilibrium offer of an inequity averse proposer depends on the extent to which she suffers from advantageous inequity, that is situations in which she gets more than the responder. Proposers that suffer heavily from inequity to their advantage offer a relatively large share to the responder. However, they never go as far as to offer him more than half the pie. Conversely, proposers that do not suffer much from advantageous inequity, find it profitable to offer a share as small as possible such that the responder is just willing to accept. In equilibrium, proposers offer

$$s\begin{cases}=\frac{1}{2} & if\ \beta_P>\frac{1}{2}\\ \in\left[\underline{s},\frac{1}{2}\right] & if\ \beta_P=\frac{1}{2}\\ =\underline{s} & if\ \beta_P<\frac{1}{2}\end{cases}$$

where $\underline{s}=\frac{\alpha_R}{1+2\alpha_R}$ represents the minimum share the respondent is willing to accept. The model with inequity aversion predicts that the proposed shares for the responder are positive even if the outside option for the responder is zero.

[9] In what follows, I denote the first player as female and the second player as male.

With the introduction of asymmetric outside options, it might occur that the minimum offer the responder is willing to accept $\underline{s}$ exceeds the equal share of $\frac{1}{2}$. This occurs if the value of the outside option to the responder exceeds the equal share of a half. In this case, the proposer offers the minimum offer irrespective of her inequity aversion. In contrast to the case with symmetric outside options, asymmetry can therefore induce that the equilibrium outcome no longer depends on the inequity aversion of the proposer. The equilibrium offer of the proposer in the presence of positive, (a)symmetric outside options is

$$s \begin{cases} = \max\{\underline{s}, \frac{1}{2}\} & if\ \beta_P > \frac{1}{2} \\ \in \left[\underline{s}, \max\{\underline{s}, \frac{1}{2}\}\right] & if\ \beta_P = \frac{1}{2} \\ = \underline{s} & if\ \beta_P < \frac{1}{2} \end{cases}$$

where $\underline{s} = \frac{\alpha_R + \omega_R - \alpha_R \max\{\omega_P - \omega_R, 0\} - \beta_R \max\{\omega_R - \omega_P, 0\}}{1 + 2\alpha_R}$.[10] An increase in the outside option for the responder increases the minimum share he is willing to accept. However, an increase in the outside option of the proposer might decrease or increase the minimum acceptable share depending on the difference in outside options of the proposer and the responder. Note that in case of symmetric outside options, the minimum share simplifies to $\underline{s} = \frac{\alpha_R + \omega_R}{1 + 2\alpha_R} \leq \frac{1}{2}$ as $\omega_R = \omega_P \leq \frac{1}{2}$.

The focus of this paper is to study the behaviour and potential bargaining breakdown in equilibrium, when the reference allocations of agents differ from the *Equal Split*. In particular, I am interested in the impact of heterogeneity in reference allocations and of ignorance concerning this heterogeneity. Before analysing the equilibrium of the general case, I explain the workings of a self-serving bias with the help of a simple example in the next section.

### 1.3.1 An example

Suppose the proposer has got no outside option and the responder's outside option is positive, $\omega_R > \omega_P = 0$. Consider the two conflicting reference allocations of *Equal Split* and *Split the Difference*. A biased proposer believes that the *Equal Split* $\left(\frac{1}{2}, \frac{1}{2}\right)$ is fair, while a biased responder adopts *Split the Difference* $\left(\frac{1-\omega_R}{2}, \frac{1+\omega_R}{2}\right)$ as reference allocation. The fairness parameter of the proposer resp. the responder is $\gamma_P = 0$ resp. $\gamma_R = 1$. The mere introduction of a self-serving bias in reference allocations does not result in a breakdown of the bargaining. A sophisticated biased proposer is always willing to divide the pie such that the respondent is at least as well off as with his outside option. The efficiency gain resulting from the bargain is large enough to compensate for deviations from the reference allocation.

To see this, I compute the maximum share the proposer is willing to offer (MTO - *Maximum Tolerable Offer,* denoted by $\overline{s}$) and the minimum share the responder is willing to accept

[10] In section 1.3.2, we show that, even if the minimum share is larger than the equal share, the proposer always prefers to offer a share the responder is willing to accept rather than staying with her outside option.

(MAO - *Minimum Acceptable Offer,* denoted by $\underline{s}$). These shares render the proposer resp. the responder indifferent between their outside option and the division of the pie. In our example, the value of the outside option to the responder is $u_R(0, \omega_R) = \omega_R$ and a division $(1-s, s)$ of the cake which is disadvantageous to him, i.e. $s \leq \frac{1+\omega_R}{2}$, results in a value of $u_R(1-s, s) = s - \alpha_R(1-2s+\omega_R)$. The responder's MAO is thus $\underline{s}_{\gamma_R=1} = \frac{(1+\alpha_R)\omega_R+\alpha_R}{1+2\alpha_R}$.

The proposer values the outside option with $u_P(0, \omega_R) = -\alpha_P \omega_R$. She derives a utility of $u_P(1-s, s) = 1-s-\alpha_P(2s-1)$ of a disadvantageous division $(1-s, s)$ of the pie, with $s \geq \frac{1}{2}$. Hence she is better off with a division of the pie as long as the share for the respondent does not exceed the MTO of $\overline{s}_{\gamma_P=0} = \frac{1+\alpha_P(1+\omega_R)}{1+2\alpha_P}$. The MTO $\overline{s}_{\gamma_P=0}$ is strictly bigger than the MAO $\underline{s}_{\gamma_R=1}$. The bargain therefore never fails to take place. The reason being that agents also dislike inequity when they stay with their outside option. The proposer thus suffers from inequity aversion in case of the breakdown of the bargain. This increases the share she is maximally willing to give to the responder. In section 1.3.2, I show that this holds in general.

If, however, the proposer is biased and naive about the bias, then the bargain is likely to fail. The naive and biased proposer thinks that the responder shares the same reference allocation with $\gamma_P = 0$. She employs this fairness parameter to compute the MAO. Hence, she believes the MAO to be the same as in the standard case with simple inequity aversion $\underline{s}_{\widehat{\gamma}_{PR}=0} = \frac{(1-\beta_R)\omega_R+\alpha_R}{1+2\alpha_R}$. This level is strictly smaller than the actual MAO, i.e. $\underline{s}_{\gamma_R=1} > \underline{s}_{\widehat{\gamma}_{PR}=0}$. If the proposer's sufferance from advantageous inequity is sufficiently small, i.e. $\beta_P < \frac{1}{2}$, then, in equilibrium, the proposer is going to propose the smallest share to the responder. Therefore, she proposes a share that is below the minimum share the responder is willing to accept and the bargain fails.

The next section extends this result to more general notions of fairness and derives the equilibrium for the case of incomplete information concerning the utility parameters $\alpha_R$ and $\beta_R$.

### 1.3.2 General case

The introduction of asymmetric outside options has several implication for the equilibrium of the ultimatum game with inequity averse agents. On the one hand, asymmetry in outside options can increase the MAO such that it exceeds the fair share of the pie. On the other hand, the asymmetry might lead to a self-serving bias.

In the framework of Fehr and Schmidt, we have already seen that in some cases, namely when the utility of the outside option to the responder is larger than the utility of the fair share, the MAO exceeds the fair share. In these cases, the proposer simply offers the MAO regardless of her level of sufferance due to advantageous inequity aversion. For the general case of heterogenous reference allocations, Lemma 1 shows that the responder's MAO is larger than the share the proposer considers to be fair for the responder if and only if the utility the responder receives from the outside option constellation is larger than the utility derived from the fair share. Denote the fair allocation of agent $i$ depending on the fairness parameter $\gamma_i$ by $\left(1-s^f(\gamma_i), s^f(\gamma_i)\right) = \left(\frac{1+\gamma_i(\omega_P-\omega_R)}{2}, \frac{1-\gamma_i(\omega_P-\omega_R)}{2}\right)$ and the MAO depending on the fairness parameter $\gamma_R$ by $\underline{s}(\gamma_R)$.

**Lemma 1** *The MAO of the responder is larger than the fair share of the proposer,* $\underline{s}(\gamma_R) > s^f(\gamma_P)$*, if and only if* $u_R(\omega_R, \omega_P) > u_R\left(s^f(\gamma_P), 1-s^f(\gamma_P)\right)$.

**Proof.** The MAO $\underline{s}(\gamma_R)$ is a disadvantageous share for the responder such that he is indifferent between the outside option and that share. It is thus determined by $u_R(\omega_R, \omega_P) = u_R(\underline{s}(\gamma_R), 1-\underline{s}(\gamma_R))$. The responder's utility of a share $s$ that is to his disadvantage is given by $u_R(s, 1-s) = s - \alpha_R(1-2s-\gamma_R(\omega_P-\omega_R))$, which is strictly increasing in the share $s$. The fair share $s^f(\gamma_P)$ that the proposer attributes to the responder is weakly disadvantageous to the responder. If proposer and responder share the same reference allocation, then the fair share is not disadvantageous. Otherwise, if agents are biased, the fair share of the proposer by definition attributes less to the responder than the fair share of the responder, hence it is disadvantageous. Therefore,

$$\begin{aligned} & u_R(\omega_R, \omega_P) - u_R\left(s^f(\gamma_P), 1-s^f(\gamma_P)\right) > 0 \\ \Leftrightarrow\ & \left(\underline{s}(\gamma_R) - s^f(\gamma_P)\right)(1+2\alpha_R) > 0 \\ \Leftrightarrow\ & \underline{s}(\gamma_R) > s^f(\gamma_P). \end{aligned}$$

■

With a self-serving bias the number of cases, where the MAO exceeds the fair share $s^f(\gamma_P)$, increases compared to the case where both agents share the same reference allocation. Suppose both agents held the same reference allocation. Now, if the proposer is becoming biased, the fair share she attributes to the responder decreases in comparison, this is the very definition of a biased agent. Therefore, for the responder, the utility of the fair share $u_R\left(s_P^f, 1-s_P^f\right)$ is, in more cases, smaller than the utility of the outside option to him $u_R(\omega_R, \omega_P)$.

In case the MAO is larger than the fair share, we have to ensure that the proposer wants to offer more than her fair share to the responder. The efficiency gain from a bargain has to be sufficiently large as to compensate the proposer for the loss resulting from the disadvantageous deviation from her reference allocation. Lemma 2 establishes that the proposer is better off if she offers the MAO to the responder than if she is left with her outside option. In case the MAO exceeds the fair share, the proposer therefore prefers to offer the MAO, than to be left with her outside option.

**Lemma 2** *The MAO* $\underline{s}(\gamma_R)$ *of the responder is smaller than the MTO* $\overline{s}(\gamma_P)$ *of the proposer.*

**Proof.** Suppose agents are biased such that $\gamma_i \geq 1$, $\gamma_j \leq 1$ for $\omega_i > \omega_j$. This includes the case where each agent considers it fair that she or he gets the entire surplus. The MTO and the MAO can then be calculated as

$$\begin{aligned} u_P\left(1-\overline{s},\overline{s}\right) &= 1-\overline{s}-\alpha_P\left(2\overline{s}-1-\gamma_P\left(\omega_R-\omega_P\right)\right) \\ &= \omega_P-\alpha_P\left(1-\gamma_P\right)\left(\omega_R-\omega_P\right)=u_P\left(\omega_P,\omega_R\right) \end{aligned}$$

$$\overline{s}=\frac{\alpha_P+1-\omega_P-\alpha_P\left(\omega_P-\omega_R\right)}{1+2\alpha_P}, \tag{1.2}$$

and

$$\begin{aligned} u_R\left(\underline{s},1-\underline{s}\right) &= \underline{s}-\alpha_R\left(1-2\underline{s}-\gamma_R\left(\omega_P-\omega_R\right)\right) \\ &= \omega_R-\alpha_R\left(1-\gamma_R\right)\left(\omega_P-\omega_R\right)=u_R\left(\omega_R,\omega_P\right) \end{aligned}$$

$$\underline{s}=\frac{\alpha_R+\omega_R+\alpha_R\left(\omega_R-\omega_P\right)}{1+2\alpha_R}. \tag{1.3}$$

Algebraic transformations show that the MAO is smaller than the MTO if the sum of the outside options is smaller than the entire pie:

$$\underline{s}\leq\overline{s}\leftrightarrow\omega_R+\omega_P\leq 1,$$

see Appendix 1.A for further detail.

If agents become less partial as either $\gamma_i$ decreases or $\gamma_j$ increases (with $\gamma_i \geq \gamma_j$), the MTO weakly increases or the MAO weakly decreases, see Appendix 1.A for further detail. ■

To get some intuition, consider the following example. Let the outside option of the proposer be half the pie, $\omega_P = \frac{1}{2}$, while the responder has no postive outside option, $\omega_R = 0$. Further assume that the proposer's reference allocation is such that she gets the entire pie and the reference allocation of the responder is the *Equal Split*. Now, if the proposer suffers a lot from disadvantageous inequity, i.e. $\alpha_P$ is very large, one might think that she is not willing to deviate much from her reference allocation and is willing to give only a very small amount to the responder, $\varepsilon$. The responder with a high $\alpha_R$ might prefer to stay with the outside option constellation $\left(\frac{1}{2},0\right)$ rather than accept the devision $(1-\varepsilon,\varepsilon)$ as he suffers less from inequity aversion under the outside options. Why is this reasoning not correct? The proposer does not only suffer from inequity aversion when the bargain takes place and she gets less than the entire pie, but also when both agents get their outside options. In both situations, proposers with a very high $\alpha_P$ suffer a lot. Hence to avoid the suffering in the outside option constellation, she is willing to propose an offer that is substantially smaller than she thinks to be fair.

The result is robust to the following modification of the model. Suppose a participation decision precedes the game. Participation implying that agents forego the possibility to earn their outside option. In this version of the game, agents receive nothing in case they do not agree on a division of the surplus, just as in the standard case. However, the decision to pass on the outside option might still influence their perception of the fair allocation. As long as it influences the reference allocation whenever the bargain takes place as well as when it breaks down, the above result stays valid. For example, suppose that the outside option of the responder is bigger than the outside option of the proposer, $\omega_R \geq \omega_P$ and that both agents are biased such that $\gamma_P \leq 0$ and $\gamma_R \geq 0$. Then the utility for the responder in case the bargain breaks down and the agents do not get anything is given by $u_R(0) = -\alpha_R\gamma_R(\omega_R - \omega_P)$. His MAO is therefore $\underline{s} = \frac{\alpha_R}{1+2\alpha_R} < \frac{1}{2}$, just as in the case with no positive outside options. Analogously, the utility for the proposer is $u_P(0) = -\alpha_P\gamma_P(\omega_P - \omega_R)$ and her MTO is thus $\overline{s} = \frac{1+\alpha_P}{1+2\alpha_P} > \frac{1}{2}$, again as in the case of no outside options. The bargain takes place precisely because both players suffer from inequity aversion also in the case when the bargain breaks down. Thus, all the following results also hold for this slightly modified version.

The MAO could only exceed the MTO if the fair allocation depends on the difference in outside options in case the bargain takes place, but not when it breaks down. Hence, only when agents have different reference allocations in these two cases, the proposer might not be willing to offer the MAO.

The following proposition characterises the equilibrium of the ultimatum bargaining with sophisticated proposers, that is proposers who understand that they are biased.

**Proposition 1** *In perception perfect equilibrium, if* $u_R(\omega_R, \omega_P) \leq u_R\left(s^f(\gamma_P), 1 - s^f(\gamma_P)\right)$, *a sophisticated proposer offers a share*

$$s^* \begin{cases} = s^f(\gamma_P) & \beta_P > \frac{1}{2} \\ \in \left[\underline{s}(\gamma_R), s^f(\gamma_P)\right] & \beta_P = \frac{1}{2} \\ = \underline{s}(\gamma_R) & \beta_P < \frac{1}{2} \end{cases}.$$

*Otherwise, she proposes* $s^* = \underline{s}(\gamma_R)$. *The responder accepts the offer.*

**Proof.** If $u_R(\omega_R, \omega_P) \leq u_R\left(s^f(\gamma_P), 1 - s^f(\gamma_P)\right)$, the MAO $\underline{s}(\gamma_R)$ is smaller than the fair share $s^f(\gamma_P)$. The rest of the proof is analogous to the proof of proposition 1 in Fehr and Schmidt (1999).

If instead $u_R(\omega_R, \omega_P) > u_R\left(s^f(\gamma_P), 1 - s^f(\gamma_P)\right)$, the MAO $\underline{s}(\gamma_R)$ exceeds the fair share $s^f(\gamma_P)$. The proposer's utility of an offer above the fair share $s \geq s^f(\gamma_P)$ is given by $u_P(s) = 1 - s - \alpha_P(2s - 1 - \gamma_P(\omega_R - \omega_P))$ which is strictly decreasing in $s$. The proposer therefore never offers a share bigger than the MAO. By definition, the responder only accepts offers above the MAO. Lemma 2 shows that the proposer always prefers to offer the MAO than to get her outside option. Therefore, in equilibrium the proposer offers exactly the MAO. ■

Proposition 1 implies that a self-serving bias as such does not generate a bargaining breakdown. The proposer is always willing to render the responder at least indifferent between his outside option and the proposed share.[11] Note that the beliefs of sophisticated agents are correct and the perception perfect equilibrium coincides with the subgame perfect equilibrium.

To what extent do the results change if the proposer is biased and naive? Naive agents believe that other agents share the reference allocation with them. In the example, we have already seen that naiveté about the self-serving bias can lead to an offer that is not acceptable for the responder. The naive proposer underestimates the MAO. If she comes to propose the underestimated MAO in perception perfect equilibrium, the responder rejects the offer and the bargain breaks down.

Lemma 3 states the conditions under which naive and biased proposers predict the MAO to be strictly smaller than the actual MAO. Whether the naive proposer accurately predicts the MAO depends crucially on whether her fairness parameter is bigger or smaller than one. Remember that a fairness parameter of one implies the reference allocation of *Split the Difference*. With *Split the Difference*, the agent does not suffer from inequity in the outside option constellation. For illustrational purposes assume that the outside option of the responder is larger than of the proposer, just as in the example presented in section 1.3.1. As soon as the fairness parameter of the responder exceeds one, the responder suffers from disadvantageous inequity in the outside option constellation, even though he has got the larger outside option. The responder thus suffers in the same way from disadvantageous inequity, both, in the outside option constellation and when he gets his MAO. Therefore his MAO is independent of the fairness parameter as can be seen in (1.3). Contrary, if the fairness parameter of the responder is below the threshold of one, the responder suffers from advantageous inequity in the outside option constellation and from disadvantageous inequity when he gets his MAO. Therefore the MAO depends on the fairness parameter.

Now, a naive proposer thinks that her fairness perception is impartial and is thus shared by the responder. The belief about the MAO is based upon the fairness parameter of the proposer. Consider again the case where the outside option of the responder is larger than that of the proposer. The naive proposer predicts that the MAO is independent of the fairness parameter if her fairness parameter exceeds one, $\gamma_P \geq 1$. Otherwise the prediction depends upon the particular fairness parameter of the proposer. A wrong prediction can only occur when the proposer predicts that the MAO depends on the fairness parameter. In case she

[11] Babcock and Loewenstein (1997) propose the self-serving bias as source of bargaining impasse. They hypothese that a self-serving bias might eliminate the contract zone, that is the set of agreements that both sides prefer to their reservation value. The above argument shows that, within the framework of extended inequity aversion, a self-serving bias does not eliminate the contract zone in an ultimatum game.

predicts the MAO to be independent of the parameter, we know her fairness parameter is above one. The bias of the agents implies that the parameter of the responder is even bigger and therefore also bigger than one. Hence, the MAO is correctly predicted. However, if the proposer predicts the MAO to be dependent on the fairness parameter, the bias implies that she underestimates the actual MAO. The following lemma generalises this argument.

**Lemma 3** *A naive and biased proposer believes the MAO to be smaller than the actual MAO,* $\underline{s}(\gamma_P) < \underline{s}(\gamma_R)$ *if and only if*
*1)* $\omega_P < \omega_R$ *and* $\gamma_P < 1$*, or*
*2)* $\omega_P > \omega_R$ *and* $\gamma_P > 1$.

The proof of Lemma 3 is relegated to Appendix 1.A. Given the conditions of Lemma 3, a naive and biased proposer underestimates the MAO, i.e. $\underline{s}(\gamma_P) < \underline{s}(\gamma_R)$. Therefore, if she offers the predicted MAO in perception perfect equilibrium, her offer is too low and is rejected by the responder. The following proposition summarises the conditions for bargaining breakdown.

**Proposition 2** *Under the conditions of Lemma 3, a naive and biased proposer causes a breakdown (with positive probability) if*
*1)* $u_R(\omega_R, \omega_P) > u_R\left(s^f(\gamma_P), 1 - s^f(\gamma_P)\right)$ *or*
*2)* $u_R(\omega_R, \omega_P) \leq u_R\left(s^f(\gamma_P), 1 - s^f(\gamma_P)\right)$ *and* $\beta_P < \frac{1}{2}$ *(*$\beta_P = \frac{1}{2}$*).*

**Proof.** In equilibrium, the respondent accepts any offer above the true MAO $\underline{s}(\gamma_R)$. Under the conditions of Lemma 3, a naive and biased proposer predicts the MAO to be too small, that is $\underline{s}(\gamma_P) < \underline{s}(\gamma_R)$. If the utility of the outside option of the responder is larger than the utility of the fair share, $u_R(\omega_R, \omega_P) > u_R\left(s^f(\gamma_P), 1 - s^f(\gamma_P)\right)$, the MAO $\underline{s}(\gamma_R)$ is larger than the fair share of the proposer $s^f(\gamma_P)$. In perception perfect equilibrium, the maximally offered share is given by $\max\left\{s^f(\gamma_P), \underline{s}(\gamma_P)\right\}$, see Proposition 1. This is smaller than the actual MAO $\underline{s}(\gamma_R)$ and the bargain breaks down.

Otherwise, if the utility of the outside option of the responder is smaller than the utility of the fair share, $u_R(\omega_R, \omega_P) < u_R\left(s^f(\gamma_P), 1 - s^f(\gamma_P)\right)$, the MAO is smaller than the fair share the proposer attributes to the responder, $\underline{s}(\gamma_R) \leq s^f(\gamma_P)$. The proposer offers a share $s^* \begin{cases} = s^f(\gamma_P) & if\ \beta_P > \frac{1}{2} \\ \in \left[\underline{s}(\gamma_P), s^f(\gamma_P)\right] & if\ \beta_P = \frac{1}{2} \\ = \underline{s}(\gamma_P) & if\ \beta_P < \frac{1}{2} \end{cases}$ in perception perfect equilibrium. Therefore, if the parameter of advantageous inequity is smaller than $\frac{1}{2}$, the equilibrium share is smaller than the minimal share and the bargain breaks down. With a parameter $\beta_P = \frac{1}{2}$, the bargain breaks down with positive probability. ■

Proposition 2 characterises the circumstances under which there is bargaining breakdown with complete information concerning the parameters of the responder's utility function $\alpha_R$

and $\beta_R$. The analysis stresses that both characteristics of a self-serving bias are crucial for breakdown, namely, the bias as well as the ignorance of it.

On the one hand, the introduction of asymmetric outside options can increase the MAO such that it exceeds the fair share of the pie, on the other, the asymmetry might lead to a self-serving bias. There is no built-in mechanism that makes a self-serving bias more likely if the difference in outside options becomes more pronounced. The conditions of Lemma 3 do not get more or less restrictive if the difference in outside options increases. We therefore do not expect more bargaining breakdown because of self-serving biases when the difference in outside options increases. However, the increase in the difference of outside options might increase the likelihood of the case where the MAO exceeds the fair share and thus the likelihood of a bargaining breakdown. Remember, in case the MAO exceeds the fair share and the proposer underestimates the MAO, the bargain fails irrespective of $\beta_P$. However, the influence of an increase in the difference of outside options is indeterminate and depends on the parameters of the utility function. Nevertheless, even if the increase in asymmetry leads to more cases where the MAO exceeds the fair share, as agents do not underestimate more often the MAO, I expect the likelihood of bargaining breakdown to increase only by little.

So far, I analysed the perception perfect equilibrium given that the proposer knows the willingness of the responder to deviate from his reference allocation. Now, suppose the proposer does not know the parameters of the responder's utility, but believes that the parameter of disadvantageous $\alpha_R$ and advantageous $\beta_R$ inequity are distributed according to the joint cumulative distribution functions $F_{\alpha,\beta}(\alpha_R, \beta_R)$ on the support $[\underline{\alpha}, \overline{\alpha}] \times [\underline{\beta}, \overline{\beta}]$.[12] Denote $\underline{s}(\widehat{\gamma}_{PR})^{\max} = \max_{\alpha_R,\beta_R} \underline{s}(\alpha_R, \beta_R | \widehat{\gamma}_{PR})$ and $\underline{s}(\widehat{\gamma}_{PR})^{\min} = \min_{\alpha_R,\beta_R} \underline{s}(\alpha_R, \beta_R | \widehat{\gamma}_{PR})$.

**Proposition 3** *With $(\alpha_R, \beta_R) \sim F_{\alpha,\beta}[\underline{\alpha}, \overline{\alpha}] \times [\underline{\beta}, \overline{\beta}]$, the proposer offers*

$$s^*(\beta_P) \in \begin{cases} \left[s^f(\gamma_P), \max\left\{s^f(\gamma_P), \underline{s}(\widehat{\gamma}_{PR})^{\max}\right\}\right] & \text{if } \beta_P > \frac{1}{2} \\ \left[\min\left\{\underline{s}(\widehat{\gamma}_{PR})^{\max}, s^f(\gamma_P)\right\}, \max\left\{\underline{s}(\widehat{\gamma}_{PR})^{\max}, s^f(\gamma_P)\right\}\right] & \text{if } \beta_P = \frac{1}{2} \\ \left[\underline{s}(\widehat{\gamma}_{PR})^{\min}, \underline{s}(\widehat{\gamma}_{PR})^{\max}\right] & \text{if } \beta_P < \frac{1}{2} \end{cases}$$

*in the perception perfect equilibrium.*

**Proof.** This follows from Propositions 1, 2 and the proof of Proposition 1 in Fehr and Schmidt (1999). ■

The perception perfect equilibrium differs for sophisticated and naive proposers in essentially two features. First, the offered shares and second, the resulting propensity of bargaining breakdown. The share sophisticated proposers offer is weakly bigger than the share offered by a naive agent. Proposers face a trade-off between costs and the probability of acceptance.

[12] Note that the lower limits on the supports are bigger or equal to zero, $\underline{\alpha}, \underline{\beta} \geq 0$, and that the upper limit on the support of advantageous inequity is smaller or equal to one, i.e. $\overline{\beta} \leq 1$.

With increasing shares, the probability of acceptance increases as well as the associated costs to the proposer. Naive proposers assess the reference allocation of the responder wrongly. They believe the responder shares the reference allocation with themselves. We have seen that, under the conditions of Lemma 3, this leads to a wrong prediction of the MAO in the complete information case. For a given parameter pair $(\alpha_R, \beta_R)$, the prediction of the MAO is smaller than the true MAO. This implies that the assessment of the probability of acceptance of a share $s$ is bigger than the actual probability. Thus, naive proposers offer less than sophisticated proposers in perception perfect equilibrium. Given that the share a sophisticated proposer offers exceeds the share of a naive proposer, the probability of bargaining breakdown increases for a naive proposer. The following proposition summarises these two characteristics of the perception perfect equilibrium with incomplete information.

**Proposition 4** *With incomplete information, a naive proposer offers (weakly) less and the probability of bargaining breakdown is (weakly) higher than with a sophisticated proposer.*

**Proof.** The maximisation problem of the proposer is characterised by $\arg\max_s \left(u_P\left(1-s,s\right)-u_P\left(\omega_P,\omega_R\right)\right) prob\left(s \geq \underline{s}\left(\widehat{\gamma}_{PR}\right)\right)+u_P\left(\omega_P,\omega_R\right)$. Note that the probability is the estimated probability of acceptance of the share $s$. Lemma 2 tells us that the proposer is always better off proposing the MAO than with her outside option. The difference between the utility of the bargain with share $s$ and the outside option is thus always positive, $u_P\left(1-s,s\right)-u_P\left(\omega_P,\omega_R\right) \geq 0$ and weakly decreasing in $s$ on the interval of the equilibrium share $s^*$.

The maximisation problem is characterised by the trade-off between a higher probability of acceptance and the associated costs. If the conditions of Lemma 3 are met, the naive proposer underestimates the MAO. Thus she believes the probability of acceptance of share $s$ to be too high. The maximisation calculus thus results in a lower share for these proposers.

As shown above the share of a sophisticated proposer is weakly bigger than the share of a naive, $s_s \geq s_n$, where the subscripts $s, n$ denote sophisticated and naive. The probability of bargaining breakdown equals the probability of acceptance of a share. Thus the probability of breakdown is smaller with a sophisticated proposer, $prob\left(s \geq s_n\right) \geq prob\left(s \geq s_s\right)$. ■

The probability of a bargaining breakdown is higher if the proposer is naive than if she is sophisticated. The intuition for this result is straightforward. Naive and sophisticated proposers face uncertainty concerning the parameters that determine the loss resulting from a deviation from the responder's reference allocation. The decision how much of the pie to offer to the responder is thus based on expectations. In some cases, the proposed share is going to be too low for the responder to accept it. This is one source of bargaining breakdown which is identical for a naive and a sophisticated proposer. If the naive proposers share the belief about the responder's reference allocation with the sophisticated, they face the same propensity of bargaining breakdown out of uncertainty. However, generally the naive proposers do not share beliefs with sophisticated. Their belief about the responder's reference allocation is based on their own assessment of fairness. We have seen that this can lead to an offer that is below the actual MAO in the complete information case and a generally smaller offer than the offer of a sophisticated agent in the incomplete information case. This is an additional source of bargaining breakdown. Consequently, the probability of acceptance and therefore the probability of bargaining breakdown is larger with naive than with sophisticated proposers.

The model predicts that asymmetry compared to symmetry in outside options increases the probability of rejection. This contrasts with the predictions of the theory of inequity aversion by Fehr and Schmidt (1999), where agents, by assumption, share the same reference allocation of *Equal Split* and thus cannot fall prey to a self-serving bias. Fehr and Schmidt predict no difference in rejection rates across ultimatum games with symmetric and asymmetric outside options. This difference in predictions provides a test that discriminates between the theory of Fehr and Schmidt that does not allow for a self-serving bias and the enriched version presented in this model.

## 1.4 Investment game

The standard investment, or trust game as in Kreps (1990) or Berg, Dickhaut, and McCabe (1995) is a two stage game. In the first stage, the recipient ($R$) decides which fraction of her initial endowment ($\omega_R$) she sends to the second player. In the second stage, this amount gets multiplied by a positive number, typically strictly bigger than 1. Player 2, the dictator ($D$), then determines how to split this (enlarged) amount between himself and the recipient. Both players receive payoffs according to the division of the dictator and the recipient gets additionally the residual of her initial endowment. I simplify this set-up slightly by forcing the recipient to either invest her entire initial endowment or nothing.[13] The investment of the recipient could then be read as a simple participation decision, the initial endowment as an outside option. If the recipient participates in the dictator game, then both players enter the second stage in which the dictator decides on the distribution of a fixed surplus of one. Payoffs are determined according to the division proposed by the dictator. Otherwise, if the recipient does not participate, both agents receive their outside option, $\omega_i \geq 0$ for $i = D, R$ which is known to both agents.[14]

Under the assumption of pure self-interest, the dictator allocates the entire surplus to

[13] This is equivalent to the representation of the game by Van Huyck, Battalio, and Walters (1995). In their model, the agent either invests her entire endowment or nothing due to the linearity of the agent's utility function.

[14] This version of the game is equivalent to a game where the participation decision of the recipient as well as of the dictator is modelled. The participation decision of the dictator is trivial. As he decides at the second stage on the distribution of payoffs, he is able to make himself at least as well off as in the outside option constellation. He therefore always participates.

himself. Thus, the recipient in the first stage (weakly) prefers to refrain from participation and to receive her outside option $\omega_R \geq 0$. Analysing the game under the assumption of inequity averse agents as in Fehr and Schmidt (1999), with *Equal Split* as the reference allocation of both agents, generates the following prediction: The dictator either gives half of the surplus to the recipient, if he is sufficiently inequity averse (i.e. $\beta_D > \frac{1}{2}$), or otherwise keeps the entire surplus to himself. A recipient paired with a dictator who does not suffer much from inequity, $\beta_D < \frac{1}{2}$, refrains from participation. As the dictator is going to keep the entire surplus, the inequality increases in the dictator game. Moreover, the monetary payoff weakly decreases. The recipient is thus strictly worse off participating. However, a recipient encountering a dictator who suffers sufficiently from inequity aversion, $\beta_D > \frac{1}{2}$, receives half of the surplus which is the fair outcome for both agents. In this case the recipient prefers to participate, given that her outside option is not excessively large. As long as her outside option is smaller than half the surplus, she is evidently better off participating. However, if the outside option is bigger than half the surplus, there are some cases where the recipient prefers the outside option constellation.[15] In comparison to the case with solely self-interested agents, inequity aversion increases participation.

We have seen that standard inequity averse dictators, who base their decision on the *Equal Split* norm, are not influenced in their decision by the constellation of outside options. How much the recipient in relation to the dictator invested beforehand is irrelevant for their decision of how to split the surplus. Yet, intuitively, the size of relative investment might influence the dictator's decision. Allowing the reference allocation to be different from the *Equal Split* implies that the fair share depends on the difference in outside options. Analogously to the case with an *Equal Split* norm, the dictator gives the share she perceives as fair to the recipient if she is sufficiently inequity averse. With the fair share depending on the outside options, the distribution of surplus in the dictator game stage depends on the constellation of outside options. Let $s$ denote the share allocated by the dictator to the recipient. Lemma 4 characterises the behaviour of the dictator.

**Lemma 4** *The optimal strategy of the dictator is*

$$s\begin{cases} = \frac{1-\gamma_D(\omega_D-\omega_R)}{2} & if\ \beta_D > \frac{1}{2} \\ \in \left[0, \frac{1-\gamma_D(\omega_D-\omega_R)}{2}\right] & if\ \beta_D = \frac{1}{2} \\ 0 & if\ \beta_D < \frac{1}{2} \end{cases}.$$

**Proof.** It is obvious that it is never optimal for the dictator to offer more than she considers the fair share for the responder. For $s \leq s^f(\gamma_D)$ the utility function of the dictator is given by $u_D(s) = 1 - s - \beta_D \max\{1 - 2s - \gamma_D(\omega_D - \omega_R), 0\}$. If $\beta_D < \frac{1}{2}$, the utility is decreasing

[15] Whenever the outside option of the recipient is larger than half the surplus, $\omega_R > \frac{1}{2}$ the recipient is relatively better off in monetary terms in the outside option constellation, $\omega_R > \omega_D$. She thus prefers to participate as long as $\frac{1}{2} > \omega_R - \beta_R(\omega_R - \omega_D) \leftrightarrow \omega_R < \frac{1-2\beta_R\omega_D}{2(1-\beta_R)}$. See Appendix 1.B for further detail.

in the share $s$. It is hence maximal for $s = 0$. If instead $\beta_D > \frac{1}{2}$, the utility is increasing in the share $s$. It is optimal to allocate the fair share $s^f(\gamma_D) = \frac{1-\gamma_D(\omega_D-\omega_R)}{2}$ to the recipient. For $\beta_D = \frac{1}{2}$, the utility of the dictator is constant over the interval between 0 and $s^f(\gamma_D)$, the dictator is thus indifferent between any of these offers. ■

The fair share depends on the fairness parameter of the dictator $\gamma_D$. As soon as the reference allocation deviates from the *Equal Split* the outside option constellation matters.

Turning to the participation decision of the recipient, we see that the behaviour crucially depends on which kind of dictator the recipients encounter. Before we analyse the behaviour of the recipients in general, let us consider an example. Suppose the dictator has no positive outside option and that the responder's outside option is positive. Furthermore, assume that the reference allocation of the recipient is *Split the Difference* $\gamma_R = 1$ and of the dictator *Equal Split* $\gamma_D = 0$. Encountering a dictator whose parameter of advantageous inequity is small, $\beta_D < \frac{1}{2}$, the recipient is not going to participate. The utility of her outside option is strictly positive, while the utility of participating and earning nothing is negative. A recipient who is paired with a sufficiently inequity averse dictator, might however participate in the second stage of the game. A sophisticated recipient who understands that the reference allocation of the dictator is *Equal Split* rather than *Split the Difference* participates as long as $\omega_R \leq \frac{1}{2(1+\alpha_R)} = \widetilde{\omega}_R(\alpha_R)$. Note that for recipients who suffer heavily from disadvantageous inequity the threshold level $\widetilde{\omega}_R$ is small. Furthermore, participation of recipients with large outside options is less likely than participation of recipients with smaller outside options.

Contrary, a naive recipient believes the dictator shares the same reference allocation with her. She thus expects to receive more than half the pie from a $\left(\beta_D > \frac{1}{2}\right)$-dictator, namely $s = \frac{1+\omega_R}{2}$. She thus participates regardless of her outside option as the utility of the outside option $u(\omega_R, 0) = \omega_R$ is strictly smaller than the utility of the fair share $u\left(\frac{1+\omega_R}{2}, \frac{1-\omega_R}{2}\right) = \frac{1+\omega_R}{2}$. In the investment game, naiveté therefore enhances participation in comparison to sophistication.

More generally, recipients participate if their utility from the outside option constellation is smaller than the utility of the share they receive from the dictator, that is if $u_R(\omega) \leq u_R(s, 1-s)$. Whenever the recipient is paired with a dictator who does not suffer much from inequity, that is $\beta_D < \frac{1}{2}$, the share the dictator sends back is zero. Participation is thus very unlikely to occur. There are some extreme reference allocations that render the recipient willing to participate even in this case. This occurs if the agent believes that the dictator deserves most of the pie, see Appendix 1.B for further detail. The decision of the recipient exclusively depends on her own reference allocation, as the dictator sends back nothing irrespective of his own reference allocation. Differing reference allocations thus do not generate any effect. Whether the recipient is naive or sophisticated is irrelevant to her behaviour.

Conversely, if the recipient is paired with a $\left(\beta_D > \frac{1}{2}\right)$-dictator, the share that is sent back equals the share the dictator believes to be fair. Participation occurs if

$$u_R(\omega) \leq u_R\left(s^f(\gamma_D), 1 - s^f(\gamma_D)\right).^{16} \tag{1.4}$$

The share the dictator allocates to the recipient depends on his own reference allocation. As agents might be biased, the fair share for the recipient from the point of view of the dictator is (weakly) smaller than from the point of view of the recipient. Thus, the recipient suffers from (weakly) disadvantageous inequity whenever the surplus is divided according to the dictator, $u_R\left(s^f(\gamma_D), 1 - s^f(\gamma_D)\right) = s^f(\gamma_D) - \alpha_R(\gamma_D - \gamma_R)(\omega_D - \omega_R)$. This representation of the left hand side of inequality (1.4) illustrates that the larger the self-serving bias, the less likely it becomes that the recipient participates.[17] Suppose the outside option of the dictator is larger than of the recipient. Biasedness implies that the fairness parameter of the dictator is (weakly) larger than the fairness parameter of the recipient, $\gamma_D \geq \gamma_R$. For a given fairness parameter of the recipient, if biasedness augments by increasing $\gamma_D$, the utility of the dictator game outcome decreases for the recipient. Thus inequality (1.4) is less likely to be satisfied and the sophisticated recipient is less likely to participate.

In this situation, the belief about the reference allocation of the dictator influences the behaviour of the recipient. There are cases where the sophisticated recipient refrains from participation because she correctly believes that the fairness parameter of the dictator strongly favours the dictator. The fair share he sends back is thus small and the recipient prefers her outside option. A naive recipient however believes that the dictator shares the reference allocation with her. She thus believes that the fair share he sends back is going to be larger rendering her participation more likely.

In the investment game naiveté plays a somehow converse role to the ultimatum game. In the latter game, naiveté induced bargaining breakdown. That is, one agent preferred his outside option to the outcome of the bargaining game. Conversely, in the investment game naiveté induces larger participation of the recipients, meaning that one agent turns her outside option down.

The impact of self-serving biases on the behaviour of recipients is however unclear. On the one hand, the self-serving bias reduces participation, on the other hand naiveté increases it. Therefore, it seems hard to test the theory in terms of the behaviour of recipients. Furthermore, allowing for heterogeneity in reference allocations implies that the share sent back by the dictator can depend on the outside options of both agents. The particular way it depends on the outside options, however, is determined by the specific fairness parameter $\gamma_D$. Thus, there is unfortunately no unambiguous prediction that could be tested.

[16] For a more rigorous analysis of recipient behaviour, see Appendix 1.B.

[17] Appendix 1.B shows that there are parameter configurations where the recipient refrains from participation.

## 1.5 Evidence

This section investigates whether we can find evidence in the experimental economics literature that can be explained by the presented theory of self-serving biases in asymmetric bargaining environments. There are some experiments on ultimatum bargaining that introduce asymmetry in outside options. In an experiment by Knez and Camerer (1995), proposer and responder have positive and asymmetric outside option. The proposer's outside option amounts to 30% of a $10-pie, while the respondents are divided into two groups. The first half of the responders (R1) gets a smaller outside option than the proposers, namely 20% of the pie, and the second half of the responders (R2) gets a higher option of 40%. Offers to the responder with the small outside option are significantly lower than to the responder with the high outside option. Moreover, MAO of the R1 responder are significantly lower than of the R2 responder. This impact of the outside options on offers and MAO can be explained with inequity averse agents. Furthermore, Knez and Camerer (1995) find that rejection rates are around 45%-48%. This is much higher than the rejection rates found for two player ultimatum games with no outside options which are around 20%, see tables on pages 53-55 in Camerer (2003). A likely cause for the increase in the rejection rate is the introduction of asymmetric outside options. The remaining experimental set-up is identical to other ultimatum bargaining experiments in western countries. If agents are inequity averse with symmetric reference allocations as postulated in Fehr and Schmidt (1999), then rejection rates should not be influenced by the introduction of asymmetric outside options. However, the existence of self-servingly biased agents can account for part of the additional inefficiencies. As we have seen, naive proposers underestimate the MAO and are thus likely to propose a share that is not acceptable for the responder. Hence, the bargain breaks down more frequently than in the case where agents are sophisticated about their bias or where they are not biased at all.

Unfortunately, the experimental set-up does not allow to distinguish whether the mere introduction of outside options has caused rejection rates to increase or whether the attached asymmetry of outside options is the driving force. Economic theory predicts that the introduction of symmetric outside options does not cause rejection rates to increase. A game with positive symmetric outside options is equivalent to a game where the pie is reduced by the sum of the outside options. Nevertheless, there might be some cognitive processes that render the game with positive symmetric outside options different to a game with no outside options and therefore breakdown might occur more frequently than with no outside options.

In the ultimatum experiments by Buchan, Croson, and Johnson (2004) and Schmitt (2004), solely one of the two players is endowed with a positive outside option. Both studies find that offers and MAO decrease with a higher outside option of the proposer. Moreover, both studies find high rejection rates. Schmitt finds that rejection rates are 50% in the treatments where proposers have the positive outside option and around 30-40% in the treatments with positive outside options for the responders. The experiment by Buchan, Croson, and Johnson is run in the US and Japan. For the US, the rejection rate in the condition with a positive outside option for the proposer is significantly larger than in the condition with no positive outside option for either player, whereas there is no significant difference in Japan.[18]

Falk, Fehr, and Fischbacher (2001) present an experiment on a reduced ultimatum game with a positive outside option for the respondent. There, the proposer can choose between a split which gives herself 8 and the responder 12 and a split where she gets 5 and the proposer gets 15. Whenever the responder rejects the offer, the proposer goes home with nothing and the responder gets his outside option of 10. They argue that: "Since both offers give the responder a higher payoff than the proposer they cannot be viewed as unfair from the responder respectively. Thus resistance to unfairness cannot explain rejections in this game." They observe that 24% of the responders reject the 8/12 offer, while only 4% reject the 5/15 offer with the difference being significant at the 1%-level. They take this result as a case for the presence of spitefulness which they define as the willingness to sanction in order to increase the payoff difference between two agents. As the 8/12 offer decreases the payoff difference in comparison to the 0/10 outcome, spiteful subjects reject the 8/12 offer. In contrast the 5/15 offer does not change the payoff difference and therefore spitefulness cannot be a reason for rejection.

The evidence from this experiment can also be explained by self-serving biases in the perception of the fair allocation. If the proposer thinks that both subjects unanimously believe that the 8/12 split is the closest to a fair outcome, she proposes this split. But she could be coupled with a responder that is convinced that splitting the difference between the pie and his outside option is fair and is therefore going to reject the inequitable share of 8/12. This provides another explanation to why the rejection of the 8/12 offer is significantly higher than the 5/15 offer. Which of these explanations suits the case better is yet to be determined.

Moreover Falk, Fehr, and Fischbacher (2001) present the results of a baseline reduced ultimatum game where the responder does not have any outside option. The possible offers are analogously 8/2 and 5/5. If the responder rejects, then none of the subjects gets any monetary payoff. They report a rejection rate of the 8/2 split of 56.3%. This is much above the rejection rate of the 8/12 offer in the game with outside options. This finding is also consistent with the above theory. Suppose an individual rejects the 8/2 split in the game without an outside option for the responder. Confronted with the 8/12 split in the game with the outside option, the same individual might be willing to accept this split. The reason is that the minimal offer the individual is willing to accept in the setting with the outside option is bigger than the minimal offer in the setting with no outside option. However, it is not increased by as much as the entire outside option. Hence, potentially more subjects

[18] Buchan, Croson, and Johnson do not report rejection rates. I thus compute these for each of their treatments using their original data set.

reject the 8/2 offer than the 8/12 offer.

In the case of the investment game, the predictions of the model presented in this paper are not clear-cut and therefore difficult to verify.[19] Apart from that, the evidence found in the literature also partly contradicts each other. To my knowledge, there are two experimental studies on investment games that vary the initial endowment of the recipient or investor systematically. Given that the first agent invests, Van Huyck, Battalio, and Walters (1995) find that most of the dictators either keep the entire surplus or send back a fraction of the surplus that exceeds the endowment invested by the first agent. This implies in particular that the amount send back by the dictator varies with the endowment of the investor. Contrary, Dufwenberg and Gneezy (2000) do not find any correlation between the fraction of the surplus send back and the size of the outside option. There are two main differences between the two studies. First, in the study by Van Huyck, Battalio, and Walters investors are allowed to invest any fraction of their endowment, while in Dufwenberg and Gneezy they can only invest everything or nothing. Second, Dufwenberg and Gneezy use the strategy method to obtain more information on dictator behaviour, while the other study does not. In both studies investment decreases with the size of the outside option. This is compatible with the predictions of the behaviour of an inequity averse investor, with and without self-serving bias.

## 1.6 Conclusion

There is strong empirical evidence that in bargaining situations with asymmetric outside options people exhibit self-serving biases concerning their fairness judgements and that these self-serving biases are a driving force of bargaining impasse. This paper provides a theoretical framework for analysing the behaviour of self-servingly biased agents in simple bargaining situations. I build on the notion of inequity aversion and extend it to incorporate self-serving biases due to asymmetric outside options. I distinguish between sophisticated and naive agents, that is, those agents who understand their bias and those who do not. I then apply the framework to analyse the behaviour of naive and sophisticated biased agents in ultimatum and investment games. For ultimatum bargaining with complete information, I find that bargaining can only break down, if biased proposers are not aware of their self-serving bias. In the incomplete information case, the propensity of bargaining breakdown is

[19] Keep in mind that the investment game is composed of two stages with the second stage being strategically identical to the dictator game. In dictator games, one agent decides on the division of the surplus. Empirical evidence suggests that the behaviour of agents in dictator type games is not entirely captured by inequity aversion. The prediction with inequity averse agents is that the dictator either keeps the entire pie to himself or sends the fair share to the recipient; the fair share supposedly being half of the pie. For example, in the dictator experiment by Forsythe, Horowitz, Savin, and Sefton (1994) around 20% of the dictators offered half the pie and another 20% kept all the surplus to themselves, the rest of the offers being distributed in between these offers with a mean offer of 25% of the surplus. Similar results are reported by Frohlich and Oppenheimer (2001).

higher with naive than with sophisticated agents. In the investment game, a self-serving bias leads to reduced participation, but naiveté about the bias increases participation.

So far, the framework only incorporates one prominent form of asymmetry due to outside options. One path of further research could be to think of incorporating other forms of asymmetries in bargaining games that might bias the perception of fairness, such as asymmetric payoff possibilities. Kagel, Kim, and Moser (1996) have run ultimatum experiments with asymmetric payoff possibilities. Players bargain over the distribution of chips with different exchange rates and different information concerning these rates. If both players are fully informed and proposers have higher exchange rates, conflicting fairness norms seem to develop. This is reflected in unusually high rejection rates.

# Appendix

## 1.A Proofs

**Details to the proof of Lemma 2** Lemma 2 states that the amount the proposer is maximally willing to give (MTO) exceeds the acceptance threshold of the responder (MAO). To prove this, we have to show that even in case of the most extreme biasedness, i.e. $\gamma_i \geq 1$, $\gamma_j \leq 1$ for $\omega_i > \omega_j$, the MAO is smaller than the MTO. Suppose $\gamma_i \geq 1$, $\gamma_j \leq 1$ for $\omega_i > \omega_j$. Then the MTO and MAO are given by (1.2) and (1.3). The following calculations show that the MAO is smaller than the MTO:

$$\begin{aligned}
\underline{s} &= \frac{\omega_R + \alpha_R - \alpha_R(\omega_P - \omega_R)}{(1+2\alpha_R)} \leq \frac{1 + \alpha_P - \omega_P + \alpha_P(\omega_R - \omega_P)}{(1+2\alpha_P)} = \overline{s} \\
&\leftrightarrow \omega_R(1+2\alpha_P) + \alpha_R(1+2\alpha_P) - \alpha_R(1+2\alpha_P)(\omega_P - \omega_R) \\
&\leq 1 + 2\alpha_R + \alpha_P(1+2\alpha_R) - \omega_P(1+2\alpha_R) + \alpha_P(1+2\alpha_R)(\omega_R - \omega_P) \\
&\leftrightarrow (\omega_R + \omega_P)(1 + \alpha_R + \alpha_P) \leq 1 + \alpha_R + \alpha_P.
\end{aligned}$$

Next, I show that the MTO given by (1.2) is the smallest MTO and that the MAO given by (1.3) is the largest MAO. Generally, the MTO is computed as $\overline{s} = \frac{\alpha_P + 1 - \omega_P}{1+2\alpha_P} + \frac{\alpha_P \gamma_P(\omega_R - \omega_P)}{1+2\alpha_P} + \frac{\alpha_P \max\{(1-\gamma_P)(\omega_R - \omega_P),0\}}{1+2\alpha_P} + \frac{\beta_P \max\{-(1-\gamma_P)(\omega_R - \omega_P),0\}}{(1+2\alpha_P)}$. The minimum of the MTO occurs at $\overline{s}^{\min} = \left\{ \begin{array}{ll} \overline{s}(\gamma_P \geq 1) & if\ \omega_P > \omega_R \\ \overline{s}(\gamma_P \leq 1) & else \end{array} \right\} = \frac{\alpha_P + 1 - \omega_P + \alpha_P(\omega_R - \omega_P)}{1+2\alpha_P}$. Similarly, the MAO can be expressed as $\underline{s} = \frac{\omega_R + \alpha_R - \alpha_R \gamma_R(\omega_P - \omega_R)}{1+2\alpha_R} - \frac{\alpha_R \max\{(1-\gamma_R)(\omega_P - \omega_R),0\}}{1+2\alpha_R} - \frac{\beta_R \max\{(1-\gamma_R)(\omega_R - \omega_P),0\}}{1+2\alpha_R}$. The maximum of the MAO occurs at $\underline{s}^{\max} = \left\{ \begin{array}{ll} \underline{s}(\gamma_R \leq 1) & if\ \omega_P > \omega_R \\ \underline{s}(\gamma_R \geq 1) & else \end{array} \right\} = \frac{\alpha_R + \omega_R + \alpha_R(\omega_R - \omega_P)}{1+2\alpha_R}$. Thus MTO exceeds the MAO. ■

**Proof of Lemma 3** Lemma 3 establishes the conditions under which a biased and naive proposer underestimates the MAO. The belief of the naive proposer concerning the MAO is given by

$$\underline{s}(\gamma_P) = \begin{cases} \frac{\alpha_R+\omega_R+\alpha_R(\omega_R-\omega_P)}{1+2\alpha_R} & \begin{array}{c}\text{if } \omega_R > \omega_P \text{ and } \gamma_P \geq 1 \\ \text{or } \omega_R < \omega_P \text{ and } \gamma_P \leq 1\end{array} \\ \frac{\alpha_R+\omega_R+((\alpha_R+\beta_R)\gamma_P-\beta_R)(\omega_R-\omega_P)}{1+2\alpha_R} & \text{else} \end{cases}.$$

1) Suppose $\underline{s}(\gamma_P) < \underline{s}(\gamma_R)$ and neither condition 1) nor 2) are satisfied. Then, if $\omega_R > \omega_P$ ($\omega_R < \omega_P$) the fairness parameter of the proposer is $\gamma_P \geq 1$ ($\gamma_P \leq 1$). As the proposer is biased, the true fairness parameter of the responder is larger (smaller) than the parameter of the proposer, $\gamma_R > \gamma_P \geq 1$ ($\gamma_R < \gamma_P \leq 1$). Thus both, the true MAO and the belief of the proposer about the MAO, are independent of the fairness parameter, $\underline{s}(\gamma_i) = \frac{\alpha_R+\omega_R+\alpha_R(\omega_R-\omega_P)}{1+2\alpha_R}$ for $i = P, R$. Hence the assumption of $\underline{s}(\gamma_P) < \underline{s}(\gamma_R)$ is violated.

2) Now, suppose condition 1) (or 2)) is satisfied, $\omega_R > \omega_P$ and $\gamma_P < 1$ (or $\omega_R < \omega_P$ and $\gamma_P < 1$). Then the belief of the proposer is $\underline{s}(\gamma_P) = \frac{\alpha_R+\omega_R+((\alpha_R+\beta_R)\gamma_P-\beta_R)(\omega_R-\omega_P)}{1+2\alpha_R}$ and we have to show that this is smaller than the true MAO, $\underline{s}(\gamma_R)$. As the proposer is biased, it holds that $\gamma_R > \gamma_P$ (resp. $\gamma_R < \gamma_P$). If the true MAO is $\underline{s}(\gamma_R) = \frac{\alpha_R+\omega_R+((\alpha_R+\beta_R)\gamma_R-\beta_R)(\omega_R-\omega_P)}{1+2\alpha_R}$, then $\underline{s}(\gamma_R) > \underline{s}(\gamma_P)$ as the MAO is increasing (decreasing) in the fairness parameter. If the true MAO is $\underline{s}(\gamma_R) = \frac{\alpha_R+\omega_R+\alpha_R(\omega_R-\omega_P)}{1+2\alpha_R}$, then $\underline{s}(\gamma_R) > \underline{s}(\gamma_P)$ as $\gamma_P > 1$ ($\gamma_P < 1$). ■

## 1.B Recipient Behaviour in Investment Games

Recipients participate as long as the value of the outside option is less than the value of the division of the surplus. If the dictator keeps the entire surplus to herself, the difference in payoffs can only increase when participating. However, there might be some recipients that believe that an increase in the difference in payoffs is fair. These recipients might be willing to sacrifice their outside option to get closer to their fair allocation. Yet, most of the recipients refrain from participation.

A $\left(\beta_D < \frac{1}{2}\right)$-dictator keeps the entire surplus to herself. The utility of participation to the recipient is thus $u_R(0,1) = -\alpha_R(1-\gamma_R(\omega_D-\omega_R))$. Furthermore, she values her outside option at $u_R(\omega_D,\omega_R) = \omega_R - \alpha_R \max\{(\omega_D-\omega_R)(1-\gamma_R),0\} - \beta_R \max\{(\omega_R-\omega_D)(1-\gamma_R),0\}$. Now, if the recipient suffers from disadvantageous inequity in the outside option constellation, that is, if $(\omega_D-\omega_R)(1-\gamma_R) \geq 0$, she always refuses to participate as

$$\begin{aligned} \omega_R - \alpha_R(\omega_D-\omega_R)(1-\gamma_R) &> -\alpha_R(1-\gamma_R(\omega_D-\omega_R)) \\ \leftrightarrow \omega_R &> -\alpha_R(1-(\omega_D-\omega_R)). \end{aligned}$$

Otherwise the recipient refuses to participates if

$$\begin{aligned} \omega_R - \beta_R (\omega_R - \omega_D)(1-\gamma_R) &> -\alpha_R (1 - \gamma_R (\omega_D - \omega_R)) \\ \gamma_R (\omega_D - \omega_R) &< \frac{\omega_R + \alpha_R + \beta_R (\omega_D - \omega_R)}{(\alpha_R + \beta_R)}. \end{aligned} \tag{1.5}$$

The above result implies that only recipients that have extreme attitudes towards fair allocations participate in a game that leaves them with no share of the surplus. Suppose the outside option of the recipient is larger than of the dictator. Then the reference allocation is the more favourable to the recipient, the larger is $\gamma_R$. Condition (1.5) implies that the recipient participates if her fairness parameter is sufficiently small. That is, if $\gamma_R \leq \frac{\omega_R+\alpha_R+\beta_R(\omega_D-\omega_R)}{(\alpha_R+\beta_R)(\omega_D-\omega_R)}$ $(< 0)$. This translates to a reference allocation attributing most of the surplus to the other party, although the recipient herself has the larger outside option. Hence, we expect most recipients to refuse participation.

Contrary, recipients who are paired with a dictator that sends the share he considers as fair, participate more frequently. In this case, participation is only prevented if the dictator's reference allocation favours himself very much. To see this, consider that a recipient paired with a $\left(\beta_D > \frac{1}{2}\right)$-dictator receives the share $s^f(\gamma_D)$. Her utility is thus $u_R\left(s^f(\gamma_D), 1 - s^f(\gamma_D)\right) = s^f(\gamma_D) - \alpha_R(\gamma_D - \gamma_R)(\omega_D - \omega_R)$. If the recipient suffers from disadvantageous inequity in the outside option constellation, that is, if $(\omega_D - \omega_R)(1-\gamma_R) \geq 0$, she participates as long as

$$\gamma_D (\omega_D - \omega_R) \leq \frac{1 - 2\omega_R + 2\alpha_R (\omega_D - \omega_R)}{(1+\alpha_R)}. \tag{1.6}$$

This condition reads: If the reference allocation of the dictator is sufficiently favourable to himself, the recipient is not going to participate. Note that there are some values of $\gamma_D$ for which condition (1.6) is not satisfied and the recipient does not participate. Suppose $\omega_D > \omega_R$, then the parameter range on $\gamma_D$ is $\left[-\frac{1}{\omega_D-\omega_R}, \frac{1}{\omega_D-\omega_R}\right]$. To see that there are some parameter constellations that do not satisfy the condition we have to establish that

$$\begin{aligned} \gamma_D \leq \frac{1 - 2\omega_R + 2\alpha_R(\omega_D - \omega_R)}{(1+\alpha_R)(\omega_D - \omega_R)} &< \frac{1}{\omega_D - \omega_R} \\ \leftrightarrow \alpha_R (2(\omega_D - \omega_R) - 1) &< 2\omega_R. \end{aligned}$$

Whenever the difference in outside options is smaller than $\frac{1}{2}$ the left hand side becomes negative, while the right hand side is positive. Hence, there are some values of $\gamma_D$ where the recipient does not participate. Note that condition (1.6) hinges solely on the fairness parameter of the dictator as the recipient suffers from disadvantageous inequity in the same way in both situations, the dictator outcome and the outside option.

If the recipient suffers from advantageous inequity in the outside option constellation, that is, if $(\omega_D - \omega_R)(1 - \gamma_R) \leq 0$, she participates as long as

$$\omega_R - \beta_R(\omega_R - \omega_D)(1 - \gamma_R) \leq \frac{1 - \gamma_D(\omega_D - \omega_R)}{2} - \alpha_R(\gamma_D - \gamma_R)(\omega_D - \omega_R).$$

This implies that the more biased agents are, the less likely participation of the recipient becomes.

# Chapter 2

# Asymmetric Outside Options in Ultimatum Bargaining: An Experiment*

## 2.1 Introduction

Individual decision making is at the core of economic theory. Under specific behavioural assumptions, economic models predict the outcomes of decision situations and derive policy implications based on these predictions. Hence, it is of great importance to verify our underlying assumptions. This paper is concerned with individual behaviour in the strategic environment of bilateral bargaining with asymmetric outside options. We try to understand how offers and demands evolve with increasing asymmetry in an ultimatum game experiment. In particular, we are interested in the effect of asymmetric outside options on bargaining breakdown.

There is experimental evidence that agents fall prey to a self-serving bias in fairness perception in the presence of asymmetry. Furthermore, this bias might be a major source of bargaining breakdown. Typically, the experiments concerned with self-serving biases and bargaining breakdown are conducted under a rich experimental framework. For example, Messick and Sentis (1979) place their experiment in a work place environment where subjects have to state the fair wage of different groups of workers. In another experiment by Babcock, Loewenstein, Issacharoff, and Camerer (1995), subjects find themselves in the roles of prosecutor and defendant in a juridical case. The question there concerns the fair settlement. In contrast, the present paper deals with the question whether asymmetry changes behaviour and leads to increasing bargaining breakdown in one of the simplest bargaining environments, the ultimatum bargaining game. In the ultimatum bargaining game, two sub-

*This chapter is based on joint work with Heike Hennig-Schmidt from the University of Bonn.

jects have to agree on how to divide a fixed pie. The first subject, the proposer, can suggest a division of the pie whereas the second subject, the responder, can either accept or reject the division. This structure is the backbone of most bargaining behaviour and one of the simplest frameworks in which we can analyse the role of asymmetry. We introduce asymmetry through outside options into the bargaining. Outside options are an important source of asymmetry. In reality, asymmetric outside options are widespread. They are present when children bargain with their parents over a certain privilege, when an employer and a job seeker discuss wages, or when countries negotiate their contribution to some global public good. Furthermore, outside options can be measured quantitatively and thus can be easily varied in an experiment.

The main focus of this paper is to analyse the impact of varying outside options on the behaviour of subjects, and in particular the impact on the breakdown of the bargaining. There are a couple of ultimatum game experiments that introduce asymmetry through outside options or different payoff possibilities (see the next section for an overview of the literature). However, none of these vary the asymmetry systematically. For the experiments that introduce outside options, it is typical to study the difference between two particular outside option constellations only. This paper contributes to the literature in that it systematically increases the outside option of one of the players. We can therefore investigate the evolution of behaviour with increasing outside options. In particular, we can also study whether a small amount of asymmetry is already sufficient to induce a shift in behaviour or whether significant changes are triggered by large asymmetries only. However, this study does not claim to investigate in detail the underlying causes for the changes in behaviour. In particular, it cannot single out whether subjects are self-servingly biased in their fairness perceptions due to different outside options.

We conduct an ultimatum game experiment where subjects are repeatedly confronted with decision situations that differ systematically with respect to the outside options of both players involved. We use the strategy method in each situation to extract the minimum acceptable offer (MAO) for the responder. Between the decision situations, subjects do not get any feedback on any of the previous situations. We run two treatments. In the first treatment, the outside option of the responder is nearly always larger than the outside option of the proposer. In the second treatment, the reverse holds.

We find that offers and MAO increase with increasing outside options of the responder and decrease with increasing outside options of the proposer. Furthermore, the probability of bargaining breakdown increases with increasing asymmetry, irrespective as to whether the asymmetry is to the advantage of the proposer or the responder. This finding suggests that, even in this relatively simple setting, there might exist a self-serving bias in fairness perception. Additionally to offers, MAO and rejections, we analyse the propensity to offer or demand an equal split. We find that the introduction of a small outside option is sufficient to decrease the propensity of equal splits. This holds in particular, when the outside option is on the side of the proposer implying that the power to propose is bundled with the advantage in terms of outside options. Moreover, we find evidence that subjects might get biased by the nature of the previous decision situations. It seems that subjects change their behaviour when they realise that repeatedly the outside option is on the side of the *responder*. In this case, the proposer offers significantly less in a situation where the asymmetry in outside options slips away whereas the responder asks significantly more. We do not find such an effect when the *proposer* has repeatedly the larger outside option. This is an interesting effect as it suggests that subjects do not regard each situation independently, but that in particular situations they are influenced by previous decision circumstances.

Next, we give an overview of the related experimental literature. Section 2.3 explains the design of the experiment in detail. We then turn to the data analysis. We start in section 2.4.1 with an across treatment analysis which is followed in section 2.4.2 by a within treatment analysis. Section 2.5 first compares our results to the results of a similar study. It then discusses some of the design features of the experiment. In the last section, we conclude.

## 2.2 Literature

The standard ultimatum game certainly is the game most intensely studied in economic experiments, besides possibly the prisoner's dilemma game. Results are regular and robust. Average offers are 30-40% of the pie, modal and median offers are 40-50%, more than half of the offers below 20% are rejected, offers below 10% and offers above 50% are rare, and the overall rejection rate is around 20%, see pages 50-55 in Camerer (2003).

Asymmetry, however, has rarely been investigated in the ultimatum game paradigm. An early study is the paper by Knez and Camerer (1995). They introduce positive and asymmetric outside options for both players. In two different treatments, responders have outside options of 20% and 40% of a $10 pie, whereas proposers always have an outside option of 30%. Knez and Camerer use the strategy method where responders state their minimal acceptable offer (MAO). On average, proposers send 38% (42.5%) of the pie in the low (high) outside option condition whereas responders demand 42.7% (49.6%) on average. The paper reports surprisingly high rejection rates of around 45-48% compared to a much lower percentage in other studies. Knez and Camerer attribute this finding to the ambiguity in fairness perceptions. Players might "... self-servingly disagree about what constitutes a fair offer" (p. 66).

Buchan, Croson, and Johnson (2004) study an asymmetric ultimatum game experiment run in the USA and Japan. The outside option of the proposer is either 20% of a 1000 token pie (worth $10 in the US and ¥2000 in Japan) in the high power condition or zero in the low power condition. The responder's outside option is always zero. After 10 rounds of ultimatum bargaining with different anonymous partners, subjects completed a post-experimental questionnaire asking for their belief on the fair offer/demand. Proposers in both cultures offer significantly less in the high power condition. In the US, offers, averaged over the 10 rounds, decrease from around 45% to 40% in the high power condition. In Japan, the offers decrease from 51% to 46%. American responders demand less in the high power condition (average MAO of 33%) than in the control condition (average MAO of 38%). Contrary, Japanese responders demand significantly more in the high power condition (average MAO of 44%) than in the control condition (average MAO of 39%). Except for Japanese proposers, actual offers/demands are significantly correlated with subjects' beliefs on "fair" offers/demands. Buchan, Croson, and Johnson do not report rejection rates, we thus compute these for each of their treatments.[20] For behaviour in the first round, we find that the rejection rates are relatively high in both conditions, around 40% in the no power condition and 40-55% in the high power condition. For the US, the rejection rate in the high power condition is significantly larger than in the no power condition, whereas there is no significant difference in Japan.[21]

The study by Schmitt (2004) investigates an ultimatum game with varying asymmetric chip valuations, asymmetric outside options and informational settings. Agents bargain over a 100 token pie where the conversion rate is either \$0.10 or \$0.30 per token, the outside option is either \$2 or nothing and agents are either entirely informed about the conversion rate and outside option of their opponent or they only know that her/his conversion rate and outside option differ from their own. Schmitt finds that offers are higher when responders rather than proposers are endowed with the positive outside option. In the first round, the average offer is 53 (26.4) chips in the treatment with the high (low) conversion rate for the proposer and the high outside option for the responder, whereas the average offer is 26.4 (17.4) chips in the treatment with high (low) conversion rate for the proposer and positive outside option for the proposer. Similarly to Knez and Camerer (1995), rejection rates are very high. In the first round, they are 50% in the treatments where proposer have the positive outside option and around 30-40% in the treatments with positive outside options for the responders.[22]

Falk, Fehr, and Fischbacher (2001) present an experiment on a reduced ultimatum game with an outside option for the respondent. There, the proposer can choose between a split which gives herself 8 and the responder 12 and a split where she gets 5 and the responder gets 15. Whenever the responder rejects the offer, the proposer goes home with nothing and the responder gets an outside option of 10. The authors observe that 24% of the responders

[20] We use the original data set, sent to us by Buchan, Croson, and Johnson. To calculate the rejection rates, we match each proposer offer with each responder demand in each round. We find that rejection rates decrease over the 10 rounds.

[21] Note that Buchan, Croson, and Johnson frame the ultimatum bargaining in a buyer-seller context giving richer context to the game than most other studies summarised here.

[22] Unlike the other studies cited, Schmitt does not use the strategy method. In her experiment, responders have to cicle whether they accept a particular offer made to them by their matched proposer.

reject the 8/12 offer while only 4% reject the 5/15 offer with the difference being significant at the 1%-level. They take this result as a case for the presence of spitefulness which they define as the willingness to sanction in order to increase the payoff difference between two agents.

All ultimatum game experiments, cited above, introduce outside options. However, each of them studies the behaviour in two particular outside option settings only. Moreover, the characteristics of these settings are similar. The (positive) outside options of both players are relatively small, ranging from approximately 7% to 40% of the pie, only in the study by Falk, Fehr, and Fischbacher (2001) does the responder have an outside option of 50% of the pie. Accordingly, the difference in outside options is small, spreading from 7% to 20% of the pie, with the exception of 50% in the Falk, Fehr, and Fischbacher paper. In contrast, the present study varies outside options systematically from no outside option to an outside option of 72% of the pie. The minimum difference in outside options is no difference at all and the maximum is 72% of the pie.

Finally, we report on a study by Güth, Huck, and Müller (2001). They conduct three reduced ultimatum games where the proposers chose between a "fair" and an "unfair" offer. The unfair offer attributes 85% of the pie to the proposer. The three games vary with respect to the fair allocation. In the first version, the alternative to the unfair offer is to split the pie exactly equally. In the second version, the alternative is slightly tipped to the favour of the responder, he gets 55%, whereas in the third version the proposer gets the more favourable share of 55%. Güth, Huck, and Müller find that proposers are less likely to make a fair offer if the equal split is replaced by a nearly equal split. Responders reject unfair offers less frequently if the nearly equal split is tipped to the side of the proposer.

## 2.3 Experimental design

In the ultimatum bargaining game, a proposer and a responder bargain over the division of a fixed pie. The proposer ($P$) announces a division of the pie. The responder ($R$) in turn accepts or rejects the proposal.[23] If he accepts, the pie is divided according to the proposed allocation. If he rejects, each player gets her or his outside option denoted by $\omega_i \geq 0$ for $i = P, R$. Both agents know the outside options of either player.

We conducted an ultimatum bargaining game with symmetric and asymmetric outside options. Subjects had to divide a pie of 22 tokens, the experimental currency. We chose a pie size of 22 tokens to enable the choice of different fairness norms and to separate between the equal split and the prominent choice of 10 tokens. Participants were confronted with 11 constellations, each differing in players' outside options. The roles of the players remained the same throughout the experiment and each subject played with the same partner in all

[23] In what follows, we denote the proposer as female and the responder as male.

constellations. Players did *not* get any feedback in between the constellations. As the order of play might influence the behaviour of each individual, we chose to randomise the order in which subjects faced the different constellations to mitigate order effects at the group level.

In each constellation, the proposer had to chose one of the 23 possible (integer) partitions of the pie by clicking on a particular partition. Simultaneously, the responder selected the partition that provided him with his minimal acceptable offer (MAO). Before making their decisions, players were informed about each other's outside option on a separate screen. Appendix 2.C shows an example of the screens for the responder.[24] Players did *not* learn about their partner's choice in any situation. After all 11 choices, subjects were informed about the finally chosen constellation which was C5 in both treatments. They were reminded of the outside options and their own decision in C5. They were told the choice of their counterpart and their earnings.

We run two treatments. In treatment 1 (T1), the responder almost always got a higher outside option, whereas in treatment T2 (T2), the situation was reversed. Table 2.1 gives an account of all 11 constellations. The constellations of outside options were chosen to study the influence of increasing asymmetry on agents' behaviour. Furthermore, C9 and C11 are chosen to enable a comparison with the study by Knez and Camerer (1995).

Table 2.1: Description of treatments

| | treatment 1 (T1) | | treatment 2 (T2) | |
|---|---|---|---|---|
| constellation | proposer outside option | responder outside option | proposer outside option | responder outside option |
| C1 | 0 | 0 | 0 | 0 |
| C2 | 0 | 2 | 2 | 0 |
| C3 | 0 | 4 | 4 | 0 |
| C4 | 0 | 6 | 6 | 0 |
| **C5** | **0** | **8** | **8** | **0** |
| C6 | 0 | 10 | 10 | 0 |
| C7 | 0 | 12 | 12 | 0 |
| C8 | 0 | 16 | 16 | 0 |
| C9 | 6 | 4 | 4 | 6 |
| C10 | 6 | 6 | 6 | 6 |
| C11 | 6 | 8 | 8 | 6 |

Remark: Outside options denoted in tokens

Instructions regarding the experimental protocol were provided at the outset of each session. The translation can be found in Appendix 2.C. We took great care to ensure that

[24] We do not request our subjects to decide for all constellations on one screen, but rather present them each constellation on a different screen. We chose this presentation for two reasons: Firstly, we wanted to present all possible divisions of the pie on the screen in a clearly arranged format. We thus provided a table with rows for each division of the pie. Subjects could then click the division they wanted to propose or accept. Secondly, we intended to stress the differences between the different constellations. Therefore we chose to present a screen in between each round describing the outside option situation in the following round.

participants understood the game and the incentives, since we wanted to measure subjects' preferences as accurately as possible. Therefore, after subjects had read the instructions, they had to answer several computerised control questions that tested their understanding of the decision situation. We did not proceed until all subjects had answered all questions correctly. Thus, we can safely assume that people understood the game and the incentives. For the control questions, we refer to Appendix 2.C.

The experiment was run in *B*onn*E*con*L*ab, the Laboratory of Experimental Economics at Bonn University, in May and June 2004. We conducted 4 sessions with 20 subjects each, two sessions for each treatment. We thus base our analysis on 20 independent observations on proposers and responders in each treatment. Subjects were recruited by the online recruiting system ORSEE (Greiner, 2004) promising a monetary reward for participation in a decision-making task. In each session, subjects were randomly allocated to the cubicles, where they took their decisions in complete anonymity from the other participants. The random allocation to cubicles also determined subjects' roles as proposers or responders. All participants were fully informed on all features of the experimental design and the procedures. The experiment was programmed using the software Z-Tree (Fischbacher, 1999).

Sessions lasted for about one hour. The exchange rate of the experimental currency was €0.6 per token. On average subjects earned €8.80 including a show-up fee of €4. The lowest payoff was €4, the highest €18. The main characteristics of the composition of the 80 participants are displayed in Table 2.2. There is no major difference in terms of gender, age and nationality across treatments. Note that only a small fraction of about 20% of the participants in each treatment are economics students.

Table 2.2: Composition of treatments

| | T1 | T2 |
|---|---|---|
| sex | 18 (F)/22(M) | 20(F)/20(M) |
| mean age | 24.85 | 23.50 |
| economics major | 22.50% | 20.00% |
| knowledge in microeconomics* | 32.50 % | 35.00% |
| German nationality | 92.50% | 100.00% |

*: includes students taking economics as a minor

## 2.4 Results

We first report major descriptive statistics of experimental behaviour, see Table 2.3 and Table 2.4. Columns 2 and 3 show the outside option of the proposer and the responder. Column 4 displays proposers' mean offer for each outside option constellation in both treatments. Column 5 shows responders' mean MAO. Standard deviations are given in brackets. Column

6 displays the rejection rate measured by matching each offer with each MAO while column 7 reports the "empirical" rejection rate which results from the pair actually matched in the experiment.[25] A remarkable first finding is that mean offers are generally lower than mean MAO. Considering each constellation separately, mean MAO exceed mean offers in 55% of all constellations in T1 and in 91% of all constellations in T2. Furthermore, we find rejection rates ranging from 25% to 46% in T1, and from 23% to 68% in T2. Average rejection rates are 39% in T1 and 52% in T2.

The remainder of this section is organised as follows. First, we investigate whether the two treatments are readily comparable. To this aim, we compare the symmetric constellations C1 and C10 across treatments. Second, we turn to the main focus of the experiment, the evolution of offers and MAO with asymmetric outside options.

### 2.4.1 Across treatment analysis

We start our analysis by comparing outcomes across the two treatments. In particular, we ask the question whether there is a treatment effect, i.e. whether the different experimental environments have an effect on behaviour. We explore this issue by looking at constellations that are identical in both treatments, i.e. the symmetric constellations C1 and C10. This crosscheck of the experimental design is necessary to analyse whether the randomisation of the order of constellations eliminates significant order effects or whether there is some path dependence in the outcomes.

In this part of the analysis, we do not assume that subjects have particular preferences. We rather argue that most standard preference representations in economics do not consider that the history of play might matter for behaviour. In each constellation, subjects are expected to decide independently of previous constellations. This leads to the prediction that no differences in behaviour occur in situations that are identical across the two treatments. Moreover, as we did not give any feedback on the outcomes of previous decisions, we can rule out a classical learning effect. Additionally, we ensured that subjects are familiar with the nature of the decision situations before we started the experiment. Thus, there should be no effect of stronger familiarity with the decision task.

**Hypothesis 0**: *In symmetric constellations, there exists no difference in behaviour across treatments.*

Constellations C1 and C10 respectively are identical decision problems for both players in both treatments. The only differences stem from the history of play preceding the constellation. The history differs in the number of previous decision problems the player had to

[25] These two rates differ slightly from each other, as shown in tables 2.3 and 2.4. However, subjects did not get any feedback during the 11 constellations and, thus, did not have any chance of learning. Therefore, the difference is not substantial. Unless stated otherwise, we use the rejection rate resulting from matching each offer with each MAO.

Table 2.3: **T1** - Average offers and MAO's by constellation

| constellation | prop | resp | mean offer | mean MAO | rejection | emp. rejection |
|---|---|---|---|---|---|---|
| C1 | 0 | 0 | 8.10 (4.12) | 9.85 (1.60) | 37.00% | 35.00% |
| C2 | 0 | 2 | 10.35 (4.90) | 9.35 (4.13) | 35.75% | 40.00% |
| C3 | 0 | 4 | 9.70 (2.94) | 10.95 (3.15) | 42.50% | 45.00% |
| C4 | 0 | 6 | 10.40 (2.44) | 10.60 (3.28) | 40.75% | 45.00% |
| C5 | 0 | 8 | 11.45 (2.91) | 11.60 (2.62) | 45.50% | 60.00% |
| C6 | 0 | 10 | 12.10 (1.89) | 11.75 (2.67) | 25.00% | 35.00% |
| C7 | 0 | 12 | 13.10 (3.54) | 12.85 (4.09) | 40.50% | 40.00% |
| C8 | 0 | 16 | 15.80 (3.38) | 15.95 (3.38) | 43.50% | 50.00% |
| C9 | 6 | 4 | 9.35 (2.99) | 9.15 (2.89) | 37.25% | 45.00% |
| C10 | 6 | 6 | 9.45 (1.93) | 10.35 (2.01) | 46.25% | 55.00% |
| C11 | 6 | 8 | 10.40 (1.43) | 10.25 (1.45) | 35.25% | 40.00% |
| all | - | - | 10.93 (3.66) | 11.15 (3.44) | 39.02% | 45.00% |

Note: Standard deviations in brackets

Table 2.4: **T2** - Average offers and MAO's by constellation

| constellation | prop | resp | mean offer | mean MAO | rejection | emp. rejection |
|---|---|---|---|---|---|---|
| C1 | 0 | 0 | 10.10 (2.85) | 10.55 (3.44) | 23.25% | 25.00% |
| C2 | 2 | 0 | 8.15 (4.16) | 10.45 (4.42) | 50.50% | 50.00% |
| C3 | 4 | 0 | 6.80 (4.16) | 8.35 (3.94) | 54.25% | 55.00% |
| C4 | 6 | 0 | 7.35 (4.77) | 9.45 (3.36) | 64.25% | 65.00% |
| C5 | 8 | 0 | 6.55 (4.72) | 8.50 (3.00) | 62.25% | 55.00% |
| C6 | 10 | 0 | 6.30 (4.80) | 7.75 (2.99) | 60.50% | 60.00% |
| C7 | 12 | 0 | 4.20 (3.32) | 6.45 (2.67) | 67.75% | 60.00% |
| C8 | 16 | 0 | 3.60 (2.84) | 5.15 (4.90) | 56.75% | 60.00% |
| C9 | 4 | 6 | 11.05 (2.04) | 10.75 (2.78) | 47.00% | 60.00% |
| C10 | 6 | 6 | 9.75 (1.99) | 10.00 (2.58) | 37.00% | 30.00% |
| C11 | 8 | 6 | 9.20 (1.85) | 9.95 (1.85) | 52.00% | 50.00% |
| all | - | - | 7.55 (4.18) | 8.85 (3.68) | 52.32% | 51.82% |

Note: Standard deviations in brackets

face until constellation C1 or C10 turned up and from differences in the outside options of previous decision situations. We can thus test the above hypothesis by comparing the distribution of offers and MAO in constellations C1 and C10 across treatments. The upper half of Table 2.5 compares offers and MAO in identical constellations over treatments. Columns 2 and 3 give the mean offers, columns 4 and 5 the mean MAO in C1 and C10 respectively. Row 3 displays significance levels for comparisons across treatments using the two-sample Wilcoxon-Mann-Whitney test. The lower half of Table 2.5 compares the propensity to offer/ demand an equal split in the two symmetric constellations C1 and C10 over treatments. The equal split represents a subclass of offers and MAO, which is frequently chosen. We thus decided to analyse equal splits separately. Columns 2 and 3 give the percentages of equal splits in offers and columns 4 and 5 the percentages of equal splits in MAO. Row 7 displays significance levels for comparisons across treatments using the $\chi^2$-test.

Table 2.5: Test results in symmetric constellations

| | Mean Offer | | Mean MAO | |
|---|---|---|---|---|
| | C1 | C10 | C1 | C10 |
| T1 | 8.10 | 9.45 | 9.85 | 10.35 |
| sign.[a] | ****(.0426)** | n.s. | n.s. | n.s. |
| T2 | 10.10 | 9.75 | 10.55 | 10.00 |
| | Equal Splits in offers | | Equal Splits in MAO | |
| | C1 | C10 | C1 | C10 |
| T1 | 60% | 45% | 50% | 40% |
| sign.[b] | ****(.028)** | n.s. | n.s. | n.s. |
| T2 | 90% | 60% | 55% | 50% |

***, **, * significant at the 1-, 5-, 10-% level, n.s.: not significant
[a]: Wilcoxon-Mann-Whitney test (two-tailed), [b]: $\chi^2$-test
p-values given in brackets

We do not find a significant difference in responder behaviour across treatments. In C1 as well as in C10, MAO do not vary significantly between T1 and T2. The percentage of equal splits in MAO does not vary significantly either. Proposer behaviour is ambivalent though. In C10, we find no significant difference in mean offers or the propensity of equal splits in offers. Contrary in C1, mean offers are significantly higher in T2 as compared to T1. Further, we find that in C1 proposers offer the equal split significantly more often in T2 than in T1. The results on proposer behaviour contradict the hypothesis of no difference in identical situations.

Further analysis suggests that the effect in proposer behaviour is due to the order in which subjects are confronted with constellations. We sequentially excluded observations where C1 came up during the last round, the second last round, the third last round and so on. It turns out that when we drop those observations where C1 occurred in the last and second last round in either treatment, the difference in offers is no longer significant. This procedure involves dropping 9 out of 40 observations. This suggests that the order of play does matter in the symmetric constellation C1. To explore the differences in treatments further, we examine how the order of play influences behaviour *within* each treatment. To this aim, we classify C1-offers depending on whether C1 turned up early or not. We construct a dummy variable indicating whether C1 came up in the first 6 rounds or afterwards. By means of a Wilcoxon-Mann-Whitney test, we then compare *early* to *late* C1-offers. In treatment T1, we find that offers are significantly higher ($p = 6.66\%$) when C1 shows up early than when it shows up late. In contrast, in treatment T2 we find no such effect.

Even though we do not find any significant difference in responder behaviour *across* the two treatments, we check whether we can find a similar effect of the order of play *within* treatments. To this aim, we classify C1-MAO depending on whether C1 turns up early or not. A Wilcoxon-Mann-Whitney test finds significantly smaller MAO ($p = 1.7\%$) when C1 turns up early in treatment T1. In treatment T2, however, there is no order effect on MAO. As already reported, we do not find any significant difference in responder behaviour in C1 across treatments. This suggests that the impact of the order of play on responder behaviour exists, but is less strong than the impact on proposer behaviour.

The effect in T1 is robust to a variation of the definition of *early* and *late*. If we define early as during the first 5 rounds, then offers are significantly higher with $p = 4.87\%$ and MAO are significantly smaller with $p = 1.7\%$. Furthermore, we do not find any effect in T2. Apart from constellation C1 in T1 , there is no robust early-late effect in any other constellation on neither offers nor MAO. This suggests that the C1 constellation is special in treatment T1 and that in all other constellations, there is no strong effect of the order of play. In the following, we refer to the effect caused by the fact that C1 occurs early or late in the order of play as the *order effect*. Due to the order effect, we have to be careful in the within treatment analysis when comparing outcomes of the C1 constellation with other constellations in T1.

**Summary 1** *In the symmetric constellation C1, proposers offer significantly less in T1 than in T2, moreover the propensity to offer an equal split is significantly smaller in T1 than in T2. There is evidence that this difference is induced by the order of play. It seems that subjects do not consider each constellation independently and that history matters, even though participants do not learn about any of their partners' decisions.*

According to the above analysis, behaviour changes depending on whether C1 turns up early or late. For the remainder of the paper, we assume that this is the case. We thus assume that in T1, subjects actually respond to an early C1 situation differently than to a late C1 situation. This seems the simplest behavioural model that is consistent with the findings. Furthermore, the following simple story fits the behavioural model. In the course of the experiment, *proposers* in treatment T1 realise that they almost always have the lower outside option. This might lead them to act aggressively when the symmetric C1 constellation turns up late. They might want to take advantage of the fact that the imbalance slips away in the C1 constellation with no outside options. Moreover, in T1, *responders* are almost always endowed with the higher outside option. Early on in the experiment, responders might not have realised that they are often in a powerful position. As they do not feel that strong yet, they might ask less when C1 turns up early than when it turns up late. Contrary, in treatment T2, the order of play does not seem to have a strong effect on the behaviour in the symmetric situation C1. This might be due to the following difference in the two treatments. In the ultimatum game, the right to propose a division of the pie induces an asymmetry in bargaining power between the two players. In treatment T1, this bargaining power is counteracted by the high outside options of the responder, whereas in T2, it is reinforced by the high outside option of the proposer herself. This difference in treatments might induce the difference in behaviour discussed above.

However, with our experimental data, we cannot detect what exactly causes the difference in proposer and responder behaviour in C1 in treatment T1. For example, instead of C1 occurring early or late being the cause for the change in behaviour, it could be that only if C8 has been one of the previous decision situations, did proposers change their behaviour in C1. We cannot discover a pattern like that. However, if this is the behavioural pattern, then classifying subjects into groups that got the C1 early or late seems a good approximation; the later C1 turns up in the order of play, the larger is the probability that C8 has turned up as one of the previous screens.

Next, we turn to analyse the impact of outside options on offers, MAO, rejection rates and equal splits.

### 2.4.2 Within treatment analysis

Up to now, we have compared behaviour across treatments. We now turn to the main focus of the study, the analysis of the variation in behaviour due to variation in the outside options. This requires comparing outcomes within treatments.

In general, different behavioural assumptions generate different model predictions. We analyse the strategic interaction with two opposing behavioural assumptions in mind. Self-interested agents who care for their own material well-being only and inequity averse agents who dislike outcomes they perceive as unequal or unfair. Experimental evidence suggests that a large fraction of agents reveals other-regarding behaviour. We chose to contrast the pure self-interest model with the Fehr and Schmidt (1999) model of inequity aversion, a particular model of other-regarding preferences, for two reasons: Firstly, inequity aversion has proven to be able to predict subjects' behaviour well in a variety of different experiments. Secondly, the Fehr and Schmidt model generates specific predictions as to the change in behaviour with varying outside options.

In the subgame-perfect equilibrium under the assumption of purely self-interested agents, the proposer offers the responder his outside option $\omega_R$ (or the outside option plus one unit of the experimental currency). This offer is accepted by the responder. In treatment T1 with increasing outside options of the responder, the assumption of self-interest thus implies that we should observe a one by one increase in the proposer's offer as well as in the responder's demand in constellations C1 to C8 and C9 to C11. In constellations C1 to C8 in treatment T2, however, the outside option of the responder remains zero. We should thus observe no change in behaviour of either player across the constellations.

We now turn to the predictions of the Fehr and Schmidt model. Inequity aversion is represented in the linear utility function $u_i(x) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}$, where $x_{i,j}$ represents monetary payoffs for agents $i$ and $j$ respectively. The utility parameters $\alpha_i$ resp. $\beta_i$ measure agent $i$'s utility loss caused by a deviation from the equal split to her disadvantage resp. advantage, with $\alpha_i \geq \beta_i \geq 0$ and $\beta_i < 1$.[26] The inequity averse responder is willing to accept any offer above the threshold $\underline{s} = \frac{22\alpha_R + \omega_R - \alpha_R \max\{\omega_P - \omega_R, 0\} - \beta_R \max\{\omega_R - \omega_P, 0\}}{1+2\alpha_R}$, the MAO. With complete information on the utility parameters $\alpha_R$, $\beta_R$, the equilibrium offer of an inequity averse proposer is given by

$$s \begin{cases} = \max\{\underline{s}, 11\} & if\ \beta_P > \frac{1}{2} \\ \in [\underline{s}, \max\{\underline{s}, 11\}] & if\ \beta_P = \frac{1}{2} \\ = \underline{s} & if\ \beta_P < \frac{1}{2} \end{cases} .^{27}$$

With incomplete information concerning the utility parameters $\alpha_R$, $\beta_R$, the equilibrium offer depends on the joint distribution of the parameters. In comparison to the pure self-interest model, the model with inequity aversion predicts offers to be positive irrespective of agents' outside options. Furthermore, a positive outside option for the responder increases the MAO. A positive outside option for the proposer reduces or increases the MAO depending on the difference in outside options. Note that in treatment T1, $\omega_R$ is (nearly) always positive, whereas in treatment T2, the reverse holds. Thus, the MAO of an inequity averse responder should increase with his own outside option $\omega_R$ by $\frac{1-\beta_R}{1+2\alpha_R}$ $(\leq 1)$ in treatment T1 and should decrease with the outside option of the proposer $\omega_P$ by $\frac{\alpha_R}{1+2\alpha_R}$ $\left(\leq \frac{1}{2}\right)$ in treatment T2.

Concerning *responder* behaviour, we can reject the assumption of self-interested agents if we observe responders (i) demanding more than their outside option, (ii) increasing their demand by less than the increase in their outside option and (iii) decreasing their demand

[26] One might argue that the outside option of an agent represents a kind of safe lower bound of income to the agent and that with increasing fixed income from the experiment, the attitude towards risk of the agent might change. This could influence the behaviour during the experiment. The linearity of the utility function, implying risk neutrality, excludes behavioural effects of this sort. We henceforth abstract from risk effects. We analyse behaviour with respect to this lower bound in Appendix 2.B. We find that some agents do not seem to regard the outside options as bounds.

[27] See proposition 1 in Kohnz (2004). The equilibrium offer looks slightly more complicated than in Fehr and Schmidt (1999). Introducing a positive outside option for the responder affects his MAO. With the outside option of the responder high enough, the MAO might be larger than half the pie. This matters in constellations C7 and C8. The equilibrium abstracts from the issue of a smallest currency unit in the experimental setting.

if the outside option of the proposer increases. With regard to *proposer* behaviour, it is more difficult to assert the adequate behavioural assumption. First of all, in an experiment, the proposer does not know the utility parameters $\alpha_R$, $\beta_R$ of the responder. Furthermore, the proposer's behaviour might be motivated by concerns for the other player's payoff *or* he might act strategically when offering more than the responder's outside option, understanding that the responder dislikes inequity. Even when the MAO exceeds the equal split, a purely self-interested proposer prefers to offer the MAO over a bargaining breakdown, see Kohnz (2004). It is hard to disentangle the strategic motive from the inequity concerns in an ultimatum game. However, in the presence of self-interested proposers who merely anticipate the inequity aversion of the responder, we expect the impact of positive outside options on offers and MAO to be identical. Contrary, with genuinely inequity averse proposers, we expect the impact of outside options on offers to be less strong than on MAO. Outside options do not change the behaviour of proposers who offer the fair allocation of half the pie. However, they do influence those proposers who are not very inequity averse, $\beta_P < \frac{1}{2}$, and offer the MAO. Thus, overall, the influence should be less pronounced for proposers.

To investigate the impact of outside options on offers and MAO, we run an OLS regression with clustered error terms to account for the dependence of observations resulting from the repeated decisions of each individual.[28] The results are reported in Table 2.6. Columns 2/ 4 and 3/ 5 report the results of the estimation for offers and MAO respectively. Note, that the regression includes constellations C1 to C8 only, because in these constellations we have a systematic increase in outside options of responders/ proposers, respectively. We report on similar results of an analysis around the symmetric constellation C10 with positive outside options for both players in Appendix 2.A.

As proposer behaviour in the symmetric constellation C1 is significantly different, we chose to run separate estimations for each treatment. We include dummy variables for each constellation capturing the impact of outside options on the behaviour of subjects. We chose a dummy variable representation over a linear or quadratic form, because it does not restrict the impact of outside options on offers and MAO to a particular functional form. Within that specification, we test the linear hypothesis that all dummies lie on a straight line. For all four regressions presented in Table 2.6, we find that the linear hypothesis can be rejected on a 1% significance level using the Wald test.[29] Additionally, the regression contains a dummy variable "C1 late" that indicates whether constellation C1 turns up early or late in the order of play. The dummy is 1 for C1 observations that turn up late in the order of play, where we define *late,* analogously to section 2.4.1, as during the last 5 rounds.[30] In treatment T1, we

[28] For information on heteroskedasticity, see chapter 12 in Greene (2000).

[29] For example in treatment T2, the null hypothesis is that the coefficient of the "T2 Outside option P2" dummy equals twice the coefficient of the "T2 Outside option P4" dummy and trice the coefficient of the "T2 Outside option P6" dummy and so on.

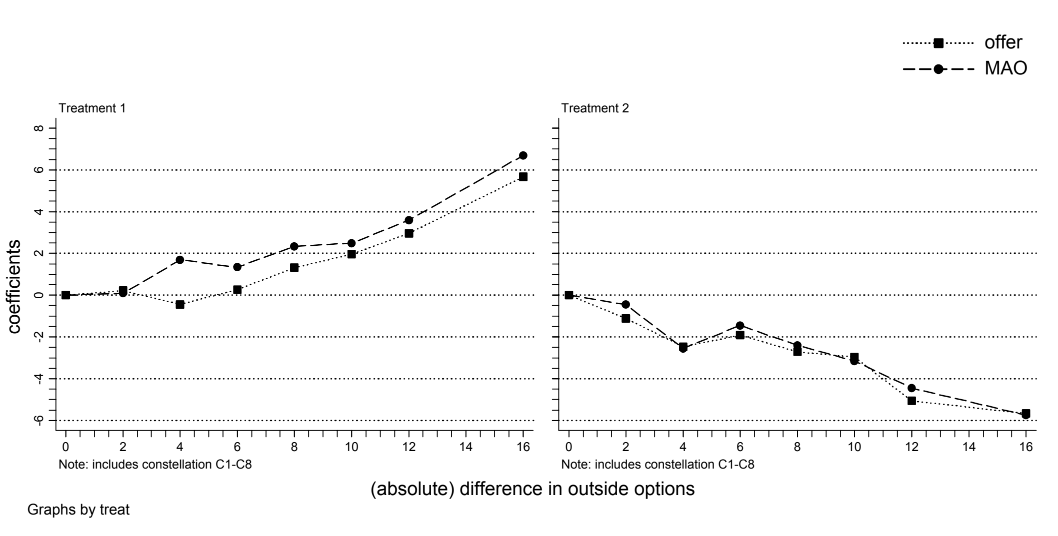
[30] The results are robust to a variation of the definition of *late* as during the last 6 rounds. See Appendix 2.A for further detail on the robustness results.

Table 2.6: OLS regression with clustered errors on offers and MAO

| | T1 | | T2 | |
|---|---|---|---|---|
| | offer | MAO | offer | MAO |
| late | **-3.134*** | **1.689***** | 1.857 | -.699 |
| | (1.556) | (.535) | (1.364) | (1.551) |
| T2 Outside option P2 | | | -1.114 | -.450 |
| | | | (1.024) | (1.124) |
| T2 Outside option P4 | | | **-2.464**** | -2.550 |
| | | | (1.043) | (1.491) |
| T2 Outside option P6 | | | -1.914 | -1.450 |
| | | | (1.413) | (1.194) |
| T2 Outside option P8 | | | **-2.714**** | **-2.400**** |
| | | | (1.090) | (1.123) |
| T2 Outside option P10 | | | **-2.964***** | **-3.150**** |
| | | | (1.016) | (1.207) |
| T2 Outside option P12 | | | **-5.064***** | **-4.450***** |
| | | | (1.252) | (1.178) |
| T2 Outside option P16 | | | **-5.664***** | **-5.750***** |
| | | | (1.513) | (1.729) |
| T1 Outside option R2 | .213 | .091 | | |
| | (1.164) | (.994) | | |
| T1 Outside option R4 | -.437 | **1.691**** | | |
| | (1.075) | (.770) | | |
| T1 Outside option R6 | .263 | 1.341 | | |
| | (.742) | (.879) | | |
| T1 Outside option R8 | 1.313 | **2.341***** | | |
| | (.945) | (.647) | | |
| T1 Outside option R10 | **1.963**** | **2.491***** | | |
| | (.824) | (.685) | | |
| T1 Outside option R12 | **2.963***** | **3.591***** | | |
| | (.961) | (1.058) | | |
| T1 Outside option R16 | **5.663***** | **6.691***** | | |
| | (1.027) | (.866) | | |
| age | -.0145 | .049 | **-.088*** | -.198 |
| | (.037) | (.209) | (.049) | (.221) |
| sex | -.062 | -.139 | .787 | .201 |
| | (.813) | (.926) | (1.626) | (.890) |
| micro | .124 | -.988 | -.466 | -.553 |
| | (.915) | (.870) | (1.390) | (.848) |
| game | -.369 | .033 | .194 | .356 |
| | (.774) | (.567) | (1.393) | (1.108) |
| constant | **6.128**** | **8.843**** | **14.639**** | **13.621**** |
| no. of observations | 160 | 160 | 160 | 160 |
| $R^2$ | 0.4458 | 0.3045 | 0.2588 | 0.2230 |

***,**, * : significantly different from zero at the 1, 5, 10% level resp.
standard deviation given in brackets, estimation includes constellations C1-C8

Figure 2.1: Plotted dummy coefficients




find that the coefficient of "C1 late" is significantly different from zero for the offer- as well as the MAO-regression. Offers in C1 decrease by approximately 3.1 tokens when they turn up late in comparison to early, whereas MAO in C1 increase by approximately 1.7 tokens. This confirms the earlier result of an *order effect*, namely that the order of play influences behaviour of both players in C1. Conversely, in treatment T2, the coefficients in both regressions are not significantly different from zero which is also consistent with the previous finding. As further explanatory variables we include age, gender, knowledge in microeconomics and game theory. The gender dummy "sex" is 1 if the subject is male. None of these variables is significant in any estimation, apart from age in the offer estimation of treatment T2.

The coefficients on the outside option dummies indicate how offers or MAO in a particular constellation change in comparison to offers/ MAO in the *early* symmetric constellation C1 with no outside option for either player. For example, the third row of Table 2.6 indicates that, in T2, offers decrease by approximately 1.1 tokens when the outside option of the proposer is 2 tokens in comparison to the constellation where both players have no outside option and this constellation turns up early in the order of play. Figure 2.1 illustrates the estimated coefficients of the dummy variables. The x-axis depicts the difference in outside options and the y-axis presents the coefficients. It visualises that from an outside option of 4 tokens onwards an increase in the outside option in T1 (T2) increases (decreases) offers and MAO approximately linearly.

**Hypothesis 1***: The outside option of the* ***proposer*** *does not influence behaviour of neither proposer nor responder.*

As Table 2.6 shows, offers as well as MAO decrease with an increase of the outside option of the proposer. All coefficients of the dummy variables in T2 are negative and the absolute values of the coefficients increase with increasing outside options of the proposer. However, not all coefficients are significant. Smaller outside options have no statistically significant impact, while larger outside options have a strong impact on both offers and MAO. The hypothesis of no influence can thus be rejected.

**Hypothesis 2**: *The outside option of the* ***responder*** *increases the offer as well as the MAO one by one.*

In T1, offers as well as MAO increase when the outside option of the responder increases. Apart from the impact of an outside option of 4 tokens on offers, all dummy variables have a positive sign, as expected and coefficients increase with increasing outside options of the responder. Coefficients on larger outside options are significant whereas coefficients on most smaller outside options are not. Thus, again we do not find an impact for smaller outside options while we do for larger outside options. An increase on a one by one basis would imply that coefficients are close to the level of the positive outside option, meaning that, for example, the coefficient for the dummy "Outside option R8" should be close to 8. As all the coefficients are relatively small, the hypothesis of a one by one increase can be rejected.

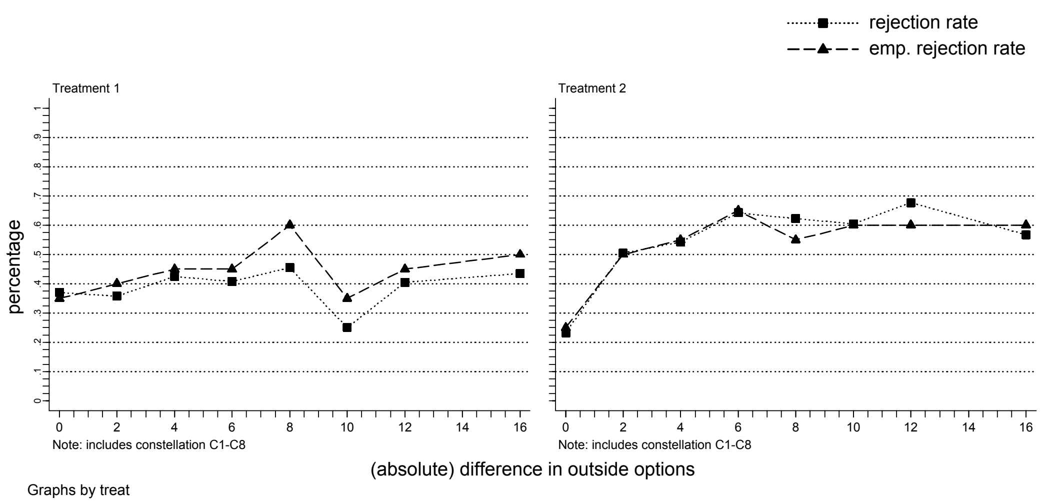
We expect that proposers, who act strategically rather than out of the concern for others, change their behaviour according to the reaction of responders to outside options. Conversely, genuinely inequity averse proposers respond less than self-interested proposers as some of them offer the fair share regardless of the outside options. We thus expect an overall effect of outside options that is less strong for proposers than for responders.

**Hypothesis 3**: *The impact of outside options is identical on MAO and offers.*

In treatment T1, the coefficients of the outside option dummies are generally larger for the MAO regression than the offer regression. This indicates a less pronounced impact on offers than on MAO when the outside option of the responder increases. In T2, there is no such clear trend. Some of the coefficients in the MAO regression are smaller than in the offer regression, some are larger. However, to test the hypothesis rigorously, we have to compare confidence intervals for the coefficients. A confidence interval specifies the bounds within which the true coefficient lies with a certain probability. If the confidence intervals given a certain probability level do not overlap, than we can reject the hypothesis that the impact of outside options is identical on offers and MAO with that probability. Considering the 95% confidence intervals, we find that these overlap considerably. We can thus not reject the hypothesis of identical impacts.

Next, we turn to the impact of outside options on rejection rates. Under the assumption of self-interest, the proposer knows the MAO of the responder, being the responder's outside option, and there is no bargaining breakdown at all. Tables 2.3 and 2.4 illustrate, however, that a considerable fraction of bargains do fail. Inequity aversion under incomplete infor-

Figure 2.2: Rejection rates




mation predicts some bargaining failure in equilibrium. However, inefficiencies should not increase with an increasing difference in outside options. Remember that the experimental subjects knew the outside option of either party before deciding in a particular constellation.

**Hypothesis 4**: *Outside options do not influence rejection rates.*

Figure 2.2 shows the development of rejection rates with outside options. The x-axis depicts the difference in outside options within constellations. The y-axis shows the percentage of rejections. The results of treatment T1 (T2) can be found in the left (right) panel. In T2, the introduction of outside options clearly increases rejection rates, both the rejection rate that measures behaviour by matching each offer with each MAO as well as the "empirical" rejection rate computed on the basis of the actual match of proposer and responder in the experiment. For T1, results are not that clear-cut.

To further examine the impact of outside options on the propensity of bargaining breakdown, we run a logit estimation, whose results are reported in Table 2.7.[31] The binary dependent variable is one if bargaining breaks down in a particular match of proposer and responder. For the logit regression, we match each proposer with each responder. Observations are not independent because each player decides several times and each player is matched with all 20 possible partners. To account for the problem of dependent observations, we include a set of dummy variables, one for each player. Again, we run two separate

[31] In a logit estimation, the coefficients can only be interpretated in terms of the sign, not in terms of the magnitude. The marginal effect of regressor $x_i$ is computed as $f(X'\beta)\beta_i$ where $\beta_i$ represents the regression coefficient of regressor $x_i$ and $f(\cdot)$ is the density function of the logistic distribution. In Table 2.7, we report the marginal effect evaluated at the mean of each regressor $x_i$. For further information, see chapter 19 in Greene (2000).

regressions, one for each treatment.[32] To tackle the *order effect*, we include dummies that indicate whether C1 turned up early or late. As both players influence the rejection in a particular match of subjects, we construct three dummies to account for the four possible combinations: Constellation C1 can turn up early for both players, it can turn up early for the proposer and late for the responder or vice versa, and last, it can turn up late for both players. We include a dummy for each of the last three combinations and leave out the first combination where both players encounter C1 early. The dummy "C1: early P late R" ("C1: late P early R") is 1 when the constellation C1 turns up early (late) for the proposer and late (early) for the responder, whereas "C1: late P late R" is 1 when C1 turns up late for both players. *Late* is again defined as during the last 5 rounds.[33] In the regression for treatment T1, the marginal effects of all three order dummies are positive, and only the first marginal effect is not significant. This implies that whenever the C1 constellation turns up late for the proposer the probability of a rejection is increased. This is consistent with the fact that offers decrease when C1 turns up late in comparison to early, as a decreased offer implies a higher probability of bargaining breakdown given the same distribution of MAO. Furthermore, the marginal effect of the second dummy where constellation C1 turns up early for the responder is smaller than the marginal effect of the third dummy where C1 turns up late for the responder. This is also consistent as MAO increase significantly when C1 turns up late instead of early. However, this effect does not seem strong enough on its own. We therefore do not observe a significant increase in the rejection probability when the proposer encounters C1 early and the responder late as compared to C1 turning up early for both players. In T2, the marginal effects of all three order dummies are significantly negative implying that the probability of a rejection decreases when the C1 constellation turns up late instead of early for either player. Even though the change in offers and MAO are not significant, Table 2.6 shows the coefficients on "C1 late" have the correct sign implying that the offer increases and the MAO decreases when C1 turns up late. By matching each proposer with each responder in the rejection regression, this effect is strengthened. This results in a significant impact of the order dummies. So the fact that in T2 the marginal effects of the order dummies are significantly negative is also consistent with previous findings.

The estimation results clearly reject the null hypothesis of no influence of the outside options on rejections. All dummy variables have positive marginal effects and are significant on the 1% or 5% level, apart from the introduction of a small outside option of 2 tokens in T2 which is significant at the 10% level only. This indicates that a positive outside option for the proposer or responder increases the probability of a rejection. This is in line with earlier experiments that find unusually high rejection rates in the presence of asymmetries, see for example Knez and Camerer (1995) and Kagel, Kim, and Moser (1996).

[32] The estimation includes constellation C1-C8. For a similar analysis around the constellation C10, see Appendix 2.A.

[33] For robustness results with respect to the definition of *late*, see Appendix 2.A.

Table 2.7: Logit estimation on rejections

| | T1 | | T2 | |
|---|---|---|---|---|
| C1: early P late R | .127 | (.140) | **-.199**** | (.090) |
| C1: late P early R | **.435***** | (.073) | **-.424***** | (.056) |
| C1: late P late R | **.523***** | (.057) | **-.544***** | (.026) |
| Outside option P2 | | | **.124*** | (.066) |
| Outside option P4 | | | **.177***** | (.061) |
| Outside option P6 | | | **.302***** | (.048) |
| Outside option P8 | | | **.279***** | (.051) |
| Outside option P10 | | | **.258***** | (.053) |
| Outside option P12 | | | **.340***** | (.043) |
| Outside option P16 | | | **.211***** | (.058) |
| Outside option R2 | **.352***** | (.082) | | |
| Outside option R4 | **.432***** | (.073) | | |
| Outside option R6 | **.413***** | (.076) | | |
| Outside option R8 | **.464***** | (.069) | | |
| Outside option R10 | **.193**** | (.091) | | |
| Outside option R12 | **.410***** | (.076) | | |
| Outside option R16 | **.443***** | (.072) | | |
| no. of observations | 3200 | | 3200 | |
| loglikelihood | -1661.930 | | -1401.655 | |
| Pseudo $R^2$ | 0.2224 | | 0.3636 | |

***,**, * : significantly different from zero at the 1, 5, 10% level resp.
reports marginal effects evaluated at the mean
standard deviation given in brackets, estimation includes constellations C1-C8

Psychological experiments suggest that with the introduction of asymmetry in the bargaining environment various fairness perceptions emerge, see Babcock and Loewenstein (1997). The same study (p. 111) states that "self-serving assessments of fairness are likely to occur in morally ambiguous settings in which there are competing "focal points" - that is, settlements that could plausibly be viewed as fair." Thus, the introduction of asymmetry in outside options might enable the emergence of various conflicting fairness perceptions. Subjects choose different reference points for measuring inequity, which might be different from the equal split. In an extended framework of inequity aversion that allows to measure inequity aversion according to different reference points, Kohnz (2004) shows that with self-servingly biased subjects, bargaining failure increases if we introduce asymmetry in outside options. The data from the present experiment is thus consistent with the theoretical predictions of inequity aversion allowing for a self-serving bias in fairness perceptions.

**Summary 2** *Inequity aversion explains the evolution of offers and MAO with outside options well. The analysis suggests that the probability of bargaining breakdown increases with increasing asymmetry in terms of outside options. This pattern cannot be explained by simple inequality aversion. However, it is consistent with the extended framework of inequity aversion that allows for a self-serving bias.*

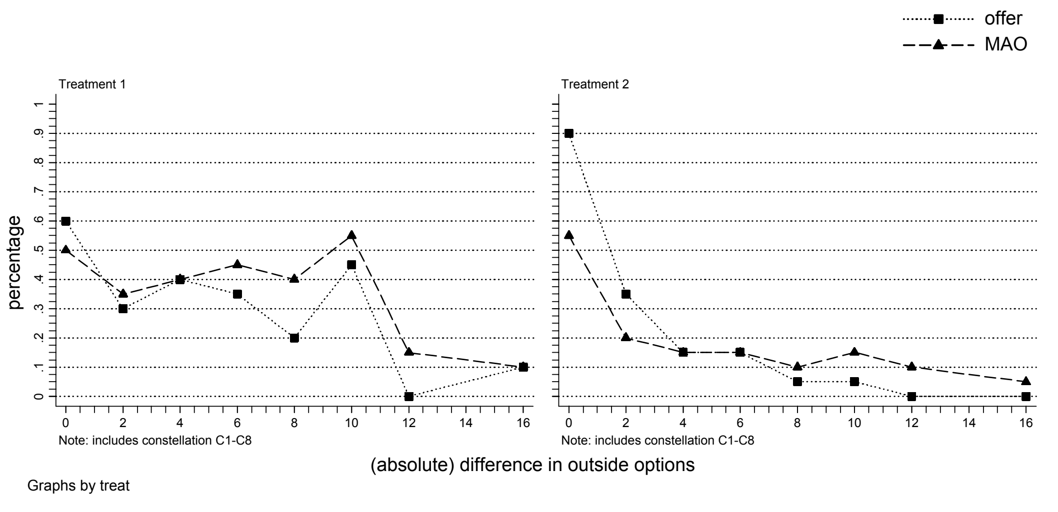
Additional insights can be gained by examining the propensities to offer or demand an equal split of the pie, i.e. 11 out of the 22 tokens. The equal split is a crucial variable in the Fehr and Schmidt (1999) theory of inequity aversion being the reference point for measuring inequity. In the standard ultimatum game, the equal split has been found to be the *modal offer* in most experiments, see Table 2.2 in Camerer (2003). We conjecture that, due to a self-serving bias, the equal split loses importance in asymmetric constellations.

**Hypothesis 5***: Asymmetric outside options do not affect the propensity to offer or demand equal splits.*

Figure 2.3 shows the evolution of equal splits in offers and MAO in both treatments. The x-axis depicts the difference in outside options in constellations, while the y-axis shows the percentage of equal splits. The results of treatment T1 can be found in the left panel. We see that the percentages of equal splits in offers differ between 60% in C1 and 0% in C7, whereas the percentages of equal splits in MAO differ between 55% in C6 and 10% in C8. The right panel of Figure 2.3 illustrates the results of treatment T2 where percentages in offers (MAO) vary between 90% (55%) in C1 and 0% (5%) in C8. In both treatments, the percentages of equal splits in offers and MAO move very much in line.

We find that in treatment T1, the equal split is the modal offer, whenever the responder's outside option is smaller than 11 tokens (C1-C6 and C9-C11). In T2, the equal split remains the mode in the two symmetric constellations C1 and C10. Yet, it ceases to be the modal offer in all asymmetric constellations, except for constellation C2 which is only weakly asymmetric.

Figure 2.3: Equal Splits




The propensity to offer or demand an equal split clearly decreases when the outside option of the proposer is increased. The right panel of Figure 2.3 visualises this. It shows in particular that the introduction of a small asymmetry induces a strong reaction in offers and MAO. In treatment T2, the decrease in equal split offers from 90% in C1 to 35% in C2 is very pronounced. According to the Binomial test, this change is highly significant, see Table 2.8.[34] Note that the difference in outside options increases by only 2 tokens, i.e. 9 percent of the pie. Responders behave similarly. They demand the equal split less often. The percentage decreases from 55% to 20% with the change being significant at the 1% level.

Table 2.8: Percentages of Equal Splits in offers and MAO

| proposers | C1 | sign. | C2 | C9 | sign. | C10 | sign. | C11 |
|---|---|---|---|---|---|---|---|---|
| T1 | 60% | ****(.035)** | 30% | 35% | n.s. | 45% | n.s. | 40% |
| T2 | 90% | *****(.001)** | 35% | 30% | ****(.044)** | 60% | ****(.011)** | 15% |
| responders | C1 | sign. | C2 | C9 | sign. | C10 | sign. | C11 |
| T1 | 50% | n.s. | 35% | 30% | n.s. | 40% | n.s. | 30% |
| T2 | 55% | *****(.008)** | 20% | 30% | n.s. | 50% | ****(.031)** | 25% |

Binomial Test: ***, **, * significant at the 1-, 5-, 10-% level, n.s.: not significant
p-values given in brackets

The reaction to an increase in the outside option of the responder is less clear-cut. Still, both proposer and responder offer/demand the equal split less frequently when little asymmetry is present. In T1, the propensity to offer (demand) an equal split falls from 60% (50%)

[34] The total number of changes is smaller than 10, we thus use the Binomial test. For more information, see the Appendix 2.D on statistical tests.

in C1 to 30% (35%) in C2. The drop in equal split offers is significant at the 5% level, see Table 2.8. However, Figure 2.3 shows that as the outside option of the responder increases further, the percentage of equal split offers and demands remains constant/ increases until the outside option of the responder is larger than the equal split. Thus the equal split does not seem to lose much importance in treatment T1. One explanation for this effect could be that in treatment T1 two forces of asymmetry work in opposing directions. On the one hand, the proposer has a higher bargaining power than the responder. On the other hand, the responder gains weight through higher outside options. These opposing forces might render the emergence of a reference allocation other than the equal split difficult as compared to treatment T2 which units the two forces on the side of the proposer.

C10 is another symmetric situation where both agents are endowed with an outside option of 6 tokens. C9 and C11 vary around this symmetric situation by giving a slightly higher outside option to the proposer or responder. Table 2.8 shows the percentages of equal splits in all of these constellations. The equal splits are most frequent in C10, as expected. Percentages decrease when asymmetry is introduced in C9 and C11. This finding supports the hypothesis that small asymmetries can decrease the propensity to propose or demand an equal split. However, the differences are significant solely in treatment T2. This supports again the conjecture that in treatment T2 the bundling of forces on the side of the proposer drives offers and MAO away from the equal split as compared to T1.

The results in equal splits can be seen in line with the results of Güth, Huck, and Müller (2001).[35] They find that proposers are less likely to make a fair offer if the equal split is replaced by a nearly equal split. Responders meanwhile reject unfair offers less frequently if the nearly equal split is tipped to the side of the proposer. In our study, in particular the latter finding can be related to the fact that the drop in equal splits is more pronounced in the second than in the first treatment. One possible explanation in both experiments is that responders resign to the fact that they are in the unfavourable position, not having much bargaining power and additionally being weak in terms of outside options.

**Summary 3** *Small asymmetries (significantly) reduce the propensity to offer an equal split. The reaction is more pronounced whenever asymmetry works in favour of the proposer than of the responder.*

## 2.5 Discussion

As pointed out, there are only a few ultimatum experiments that introduce outside options into the bargaining setting. Our results are fairly comparable to the results of Knez and Camerer (1995), who introduce positive outside options for both players. Proposers have an

[35] For a detailed description of their experimental design, see the literature overview in section 2.2.

outside option of 30% of a $10-pie. Responders are split into two groups. One group has an outside option of 20% and the other group of 40%. In our experiment, this is roughly comparable to constellations 9 and 11 of treatment T1, where the proposer has an outside option of 27% and the responder of either 18% or 36%. Simply comparing means of offers and MAO highlights that mostly behaviour seems similar. In the experiment by Knez and Camerer, responders demanded on average a share of 42.7% when they had an outside option of 20% whereas subjects in our experiment demanded 41.6%. So they demanded slightly less with a slightly smaller outside option. With an outside option of 40% in Knez and Camerer, responders asked on average 49.6%, where our subjects asked for 46.6%, again slightly less. Behaviour of proposers seem to differ a little more. Their offers average 38% (42.5%) in the experiment by Knez and Camerer (in our experiment) when paired with a responder that receives a smaller outside option. However, when paired with the responder of the high outside option type offers seem to be rather comparable, with a mean of 46.6% (47.3%) in Knez and Camerer (in our experiment). Rejection rates in both studies are high at 45% (37.25%) when the proposer has the higher outside option and 48% (35.25%) otherwise.[36]

Next, we discuss some features of our experimental design. First, we elaborate upon the fact that our subjects do not have a chance to learn and second, we discuss the influence of the strategy method upon the results.

### 2.5.1 No learning

We deliberately decided against using a repeated environment with feedback inbetween repetitions. To our opinion, one shot designs are best to understand the nature of preferences, particularly in such simple environments as ultimatum games. Analysing behaviour in a repeated framework is complicated by, above all, learning effects. However, we are not interested in analysing the particular learning pattern, but rather want to understand subjects' initial preferences. The focus of our study is to examine the preference of agents absent any strategic reasoning resulting from the repetition of events.

We took great care in making sure that our subjects did understand the implications of their decisions. After the instructions had been read out to them by one of the experimenters, they had plenty of time to re-read them again. Only after no one had any questions concerning the rules of the game did we start the computers. Subjects' first task was to answer a couple of computerised control questions ensuring that they did actually understand the nature of the task. A typical control question consisted of the description of the outcome of a particular constellation (which was different from all constellations in both treatments) and the request to fill in the payoff allocation of that outcome. This, in particular, ensured that subjects did understand the role of outside options. For the exact translation of the control questions, see Appendix 2.C. We proceeded with the decision part of the experiment only when every

[36] Knez and Camerer measure rejection rates also by matching each offer with each MAO.

subject had answered the control questions correctly. The introductory part took about 20 to 25 minutes.

### 2.5.2 Strategy method

In each constellation, we apply the strategy method introduced by Selten (1967)) to elicit information on responder behaviour in those information sets that are not reached in the experiment. Instead of reacting to a particular offer by the proposer, responders report the minimum offer they are still willing to accept, the MAO.[37] This does not only provide more information on responder behaviour, but also renders a design with no feedback for both players feasible.

Roth (1995) points out that the strategy method may suffer from several shortcomings that might induce participants to behave differently than under the non-strategy method. In particular, with the strategy method participants are required to make several decisions simultaneously. The representation of the game might appear to be more complex and incentives to think carefully about the decision at each contingency might get diluted. Moreover, events that do occur appear to be more real or vivid, emotions might therefore influence participants less when deciding under the strategy method. In case there exists an effect of the strategy method on the behaviour of responders, we expect this effect to increase the acceptance threshold of offers in the ultimatum game; responders under the strategy method do not feel the immediate loss when they decide which offers to reject whereas responders in the sequential mode do.

In the sequential mode, the equilibrium is (essentially) unique. Given pure self-interest, the responder accepts any offer larger than his outside option and the proposer offers exactly the outside option of the responder. The inequity averse responder accepts any offer above the mimimum acceptable offer. With complete information, the equilibrium offer of the inequity averse proposer is either the fair share or the minimum acceptable offer, depending on the degree of advantageous inequity aversion of the proposer. In contrast, the strategy method gives rise to a multiplicity of equilibria. With pure self-interest, any demand above the outside option of the responder can be sustained in equilibrium; and with inequity aversion, any share above the MAO and below the share that the proposer is maximally willing to send can be the equilibrium share. However, the discussed equilibrium remains special as it marks the equilibrium with the minimum payoff for the responder among all equilibrium payoffs.

Experimental results are inconclusive as to the behavioural effects of simultaneous versus sequential choices. Brandts and Charness (2000), Cason and Mui (1998) and Oxoby and McLeish (2004) do not find a significant difference in behaviour whereas Brosig, Weimann,

[37] If we take for granted that all offers higher than the MAO are accepted as well, subjects give a complete strategy vector for the game. For an overview on experimental studies reporting non-monotone strategies, see Hennig-Schmidt and Yang (2004).

and Yang (2003) and Güth, Huck, and Müller (2001) do.[38] Given our research questions, however, the strategy method is a necessary tool for our study. Whether it has an effect on the behaviour of responders or not, does not really affect the analysis of our data as long as the effect is similar across constellations within treatments and across both treatments. However, when comparing our results to other ultimatum experiments, that do not use the strategy method, one should keep in mind that our rejection rates might be slightly higher due to the strategy method.

## 2.6 Conclusion

In our ultimatum game experiment, we find that offers and MAO increase (decrease) with increasing outside options of the responder (proposer). Furthermore, with increasing asymmetry the probability that the bargaining breaks down increases. It does not matter whether it is the outside option of the proposer or the responder that has increased. This finding suggests that, even in this relatively simple setting, there might exist a self-serving bias in fairness perception that underlies the increase in bargaining breakdown. Analysing the propensity to offer or demand an equal split, we find that the introduction of a small outside option is sufficient to decrease the propensity of equal splits. This holds in particular when the outside option is on the side of the proposer implying that the power to propose is bundled with the advantage in terms of outside options.

Moreover in the symmetric situation with no outside options, we have seen that it makes a difference whether subjects belong to the treatment that attributes repeatedly the larger outside option to the proposer or to the responder. In the latter treatment, agents behave differently when the constellation turns up early than when it turns up late. Proposers offer significantly less when the symmetric constellations turns up late whereas responders ask for significantly more. We do not find such an effect in the treatment where the proposer has nearly always the larger outside option. This finding suggests that the nature of the previous decision situations influences behaviour. However, it only seems to do so in particular circumstances, namely when there are two opposing forces. These forces are in our experiment the power to propose the division of the pie and the advantage in outside options. However, we can only speculate about the underlying forces driving people to take the previous situations into account or not. It would be interesting to analyse whether our finding turns up in similar experiments and whether we can tackle the question what is driving this effect.

[38] Blount and Bazerman (1996) find a behavioural effect depending on how the strategy method is presented. They conduct an ultimatum experiment and elicit the behaviour of the responder in two ways. In the first condition, responders are asked to state their MAO. In the second condition, responders are asked to explicitly indicate for each possible offer whether they accept or reject. They found that participants are more willing to accept unequal payoffs when asked to separately reject or accept each possible offer.

# Appendix

## 2.A Additional regressions

Table 2.9 gives account of the estimation results when C10, the symmetric constellation with positive outside options for both players, is the point of reference. It includes constellation C9-C11. In the symmetric constellation C10 with positive outside options for both players, behaviour does not differ significantly between treatments. We thus run one regression rather than separate regressions for each treatment. This implies in the OLS regression that we restrict the influence of all other explanatory variables like age and gender to be identical in both treatments. However, this is not a strong restriction as the results for the OLS regression on constellations C1-C8 show no significant effect on any explanatory variable other than the outside options.

Table 2.9: Estimation results around C10

| | OLS regression with clustered errors | | | | logit estimation | |
|---|---|---|---|---|---|---|
| | offer | | MAO | | rejection [a)] | |
| T2 Outside Option P4 | **1.221**** | (.502) | .653 | (.526) | **.205***** | (.052) |
| T2 Outside Option P8 | -.628 | (.415) | -.147 | (.394) | **.300***** | (.049) |
| T1 Outside Option R4 | -.022 | (.604) | **-1.103*** | (.626) | **-.141***** | (.039) |
| T1 Outside Option R8 | **1.028**** | (.397) | -.003 | (.529) | **-.171***** | (.038) |
| age | -.046 | (.033) | .023 | (.032) | | |
| sex | .640 | (.547) | **1.325**** | (.578) | | |
| micro | .631 | (.561) | -.673 | (.617) | | |
| game | -.470 | (.537) | -.187 | (.542) | | |
| fairresp | -.121 | (.092) | .130 | (.106) | | |
| constant | **11.787***** | (1.240) | **7.932***** | (1.738) | | |
| no. of observations | 120 | | 120 | | 2400 | |
| loglikelihood | | | | | -959.656 | |
| $R^2$ | 0.1866 | | 0.1549 | | [b)]0.4135 | |

***,**, * : significantly different from zero at the 1, 5, 10% level
standard deviations given in brackets,[a)]: reports marginal effects evaluated at the mean
[b)]: reports Pseudo-$R^2$, estimation includes constellations C9-C11

The second and third column display the results of the OLS regression with clustered errors on offers and MAO whereas the last column reports the results of the logit estimation on rejections. Rows 2-5 show the impact of asymmetry on the offers and MAO with comparison to the symmetric constellation C10 where both players have an outside option of 6 tokens. In the OLS regressions, all dummy variables exhibit the expected sign. An increase in the outside option of the proposer decreases the offer as well as the MAO, whereas an increase in the outside option of the responder increases offers and MAO. However, most of the coefficients are not significant.

For the logit estimation on rejections, we match again each proposer with each responder and take account of the dependence of observations through dummy variables for each individual. Rows 2-5 show the impact of asymmetry on the probability of breakdown. In T2, as expected, the introduction of asymmetry significantly increases the probability of a rejection. However, in T1, the reverse holds. Asymmetry reduces the probability of a rejection significantly.

The following two tables show the robustness results with respect to the definition of the "C1 late" variable. Throughout the paper, we present the results derived by defining the "C1 late" dummy as 1 whenever the C1 constellation turns up in the last **5** rounds. Changing the definition such that the dummy is 1 whenever the C1 constellation turns up in the last **6** rounds does not impact the results much. In the OLS regression with clustered errors, the different definition changes the signs of two coefficients. The impact of small outside options of the responder on offers in treatment T1 becomes negative. However, as before the coefficients are not significantly different from zero. There exists no other considerable impact of the exact definition on the OLS results.

In the logit estimation on rejections, the sign of none of the marginal effects changes in both treatments implying that the introduction of asymmetry increases breakdown. The overall explanatory power of the regression stays the same in both treatments. However, in treatment T1, we cannot compute the marginal effect of the order dummy "C1: early P late R". This is due to the fact that a) there are only very few observations for which the dummy is 1 and b) in these few instances the dummy predicts rejection perfectly.

Table 2.10: OLS regression with clustered errors - Robustness results

| | T1 | | T2 | |
|---|---|---|---|---|
| | offer | MAO | offer | MAO |
| C1 late | **-3.939*** ** | **1.689*** ** | 1.857 | -.551 |
| | (1.238) | (.535) | (1.364) | (1.642) |
| T2 Outside option P2 | | | -1.114 | -.403 |
| | | | (1.024) | (1.236) |
| T2 Outside option P4 | | | **-2.464** ** | -2.503 |
| | | | (1.043) | (1.621) |
| T2 Outside option P6 | | | -1.914 | -1.403 |
| | | | (1.413) | (1.325) |
| T2 Outside option P8 | | | **-2.714** ** | **-2.353* ** |
| | | | (1.090) | (1.257) |
| T2 Outside option P10 | | | **-2.964*** ** | **-3.103** ** |
| | | | (1.016) | (1.336) |
| T2 Outside option P12 | | | **-5.064*** ** | **-4.403*** ** |
| | | | (1.252) | (1.327) |
| T2 Outside option P16 | | | **-5.664*** ** | **-5.703*** ** |
| | | | (1.513) | (1.866) |
| T1 Outside option R2 | -.704 | .091 | | |
| | (1.159) | (.994) | | |
| T1 Outside option R4 | -1.354 | **1.691** ** | | |
| | (.816) | (.770) | | |
| T1 Outside option R6 | -.654 | 1.341 | | |
| | (.573) | (.879) | | |
| T1 Outside option R8 | .396 | **2.341*** ** | | |
| | (.623) | (.647) | | |
| T1 Outside option R10 | **1.046* ** | **2.491*** ** | | |
| | (.505) | (.685) | | |
| T1 Outside option R12 | **2.046*** ** | **3.591*** ** | | |
| | (.698) | (1.058) | | |
| T1 Outside option R16 | **4.746*** ** | **6.691*** ** | | |
| | (.650) | (.866) | | |
| age | -.014 | .049 | **-.088* ** | -.199 |
| | (.037) | (.209) | (.049) | (.221) |
| sex | -.088 | -.139 | .787 | .207 |
| | (.811) | (.926) | (1.626) | (.892) |
| micro | .227 | -.988 | -.466 | -.555 |
| | (.910) | (.870) | (1.390) | (.855) |
| game | -.271 | .0327 | .194 | .380 |
| | (.767) | (.567) | (1.393) | (1.110) |
| constant | **6.815** ** | **8.843** ** | **14.639** ** | **13.614** ** |
| | (2.425) | (3.974) | (5.399) | (5.245) |
| no. of observations | 160 | 160 | 160 | 160 |
| $R^2$ | 0.4513 | 0.3045 | 0.2588 | 0.2227 |

***,**, * : significantly different from zero at the 1, 5, 10% level resp.
standard deviation given in brackets, estimation includes constellations C1-C8

Table 2.11: Logit estimation - Robustness results

| | T1 | | T2 | |
|---|---|---|---|---|
| C1: early P late R | | | **-.281*****| (.081) |
| C1: late P early R | **.902***** | (.007) | **-.417***** | (.059) |
| C1: late P late R | **.840***** | (.008) | **-.563***** | (.021) |
| Outside option P2 | | | .069 | (.072) |
| Outside option P4 | | | **.126*** | (.068) |
| Outside option P6 | | | **.262***** | (.055) |
| Outside option P8 | | | **.237***** | (.058) |
| Outside option P10 | | | **.214***** | (.060) |
| Outside option P12 | | | **.305***** | (.049) |
| Outside option P16 | | | **.162**** | (.065) |
| Outside option R2 | **.970***** | (.004) | | |
| Outside option R4 | **.971***** | (.005) | | |
| Outside option R6 | **.971***** | (.005) | | |
| Outside option R8 | **.972***** | (.005) | | |
| Outside option R10 | **.967***** | (.004) | | |
| Outside option R12 | **.971***** | (.005) | | |
| Outside option R16 | **.972***** | (.005) | | |
| no. of observations | 3200 | | 3200 | |
| loglikelihood | -1633.165 | | -1396.308 | |
| Pseudo $R^2$ | 0.2296 | | 0.3660 | |

***,**, * : significantly different from zero at the 1, 5, 10% level resp.
reports marginal effects evaluated at the mean
standard deviation given in brackets, estimation includes constellations C1-C8

## 2.B Seemingly irrational behaviour

Another interesting feature of our data is that some subjects do not seem to see the outside options as lower or upper bound for demands and offers. Some responders ask for less than their own outside option while others ask for more than the pie minus the outside option of the proposer. Equivalently, some proposers offer less than their counterpart's outside option and some offer more than the pie minus their own outside option. Naturally, such "irrational" choices are not the rule, but rather the exception. However, they account for 4.25% of overall behaviour. That is, we find 34 instances of such behaviour out of 800 observations, excluding the behaviour in constellation C1 where both outside options are zero.

Table 2.12 gives a detailed account of extreme behaviour of subjects. It indicates the percentage (over all constellations, i.e. 200 observations per treatment and role) of offers or MAO that either fall short of the players own outside option or the other players outside option. The second column gives the percentage of offers/ MAO that are smaller than the outside option of the responder in T1. The third column gives the percentage of offers/ MAO that are bigger than the pie minus the outside option of the proposer in T2. In T1, where the outside option is favourable to the responder, proposers never offer too much. However, they do offer less than the responders' outside option in 4% of the cases. In T2, where the outside option is to their own favour, proposers keep less for themselves than is guaranteed by their outside option in 2.5% of the cases. None of these offers involve the equal split. In total, there are 8 distinct proposers behaving seemingly irrational.

Table 2.12: Extreme offers and MAO

| | T1: too small | T2: too big |
|---|---|---|
| offer | 4,0% | 2,5% |
| MAO | 6,0% | 3,5% |

The picture looks a little bit different for responders. Responders ask for less than their outside option in 6% of the cases in T1. Responders demand five times an equal split (3 in C7, 2 in C8); in seven cases responders claim other shares (1 in C4 and C6, 3 in C7, 2 in C8). Responders granting less to the proposers than their outside option occurs in 3.5% of the cases in T2. Three equal-split MAO are involved (2 in C7, 1 in C8) and four MAO of other shares (1 in C4 and C6, 2 in C8). Mostly, these extreme offers were made in extreme situations such as C7 and C8. Furthermore, for the responders it is 11 different individuals behaving in the seemingly irrational way.

It is a puzzling phenomenon that subjects should behave in the above manner. One is tempted to reason that these subjects did not understand the game from the outset. However, the ultimatum game is a relatively simple experimental set-up, plus everyone had to answer the control questions correctly before we proceeded with the decision part. Thus, it is relatively unlikely that all these 19 subjects should have misunderstood the game. The finding suggests that experimentalists should be careful in limiting subjects' choice sets as it might restrict behaviour. In the study by Knez and Camerer (1995), experimental subjects are limited in their choice of offers and MAO. Responders circle a MAO from a range of numbers between their own outside option and the pie minus the outside option of the proposers. The proposers played in two conditions. In the first condition, proposers simply announce an offer to the responder. In the second condition, they circle a number, just as the responder, from a range of numbers between the outside option of the responder and the pie minus their own outside option. They find that offers are significantly higher in the second condition. They trace this effect back to the possibility that the offer range can influence how players define the surplus and therefore players' notion of the fair allocation. Our finding can add to the explanation of the effect by pointing out that some choices were no longer available to the proposer in the second condition and that this might possibly constitute a binding restriction to proposer behaviour.

## 2.C Instructions, control questions and example screens

**Instructions (Original in German)**

**Welcome to the Experiment!**

General explanations concerning the experiment

You are participating in an economic experiment funded by the Deutsche Forschungsgemeinschaft. According to your performance you can earn money. Please read the following instructions carefully.

Independently of the outcome of the experiment each participant will receive an amount of **€4**.

**During the experiment, it is not allowed to communicate with any of the other participants in any way.** If you have any questions, please **contact the experimenter**.

During the experiment, we do not talk of Euro, but of „Taler“. Your payoff will first be calculated in Taler. In the end, the total payoff in Taler, that you earned during the experiment, is going to be transferred into Euro at an exchange rate of **1 Taler = €0.6**.

At the end of the experiment we will pay you the amount of Taler, you earned during the experiment transferred into Euro, plus the participation fee of €4 **in cash**.

At the beginning of the experiment, all participants are randomly divided into dyads interacting with each other during the whole experiment. Yet, **you will not be informed who**

**the other person in you dyad is**, neither during nor after the experiment.

The experiment consists of two parts. First, in each dyad a decision has to be taken **how to divide a fixed amount of Taler between a Sender and Receiver**. This decision has to be taken **in different situations**. Then, you are asked to answer some questions.

In the next paragraphs, we will explain the experimental procedure. Following this introduction, we ask you to answer some computerized control questions which will help you to get acquainted with the decision situation.

## The decision in the experiment

In this experiment, **Senders** and **Receivers** are participating. Whether you are a Sender or a Receiver is determined by your drawing a terminal number and is revealed to you at the beginning of the experiment.

You will have to **divide an amount of 22 Taler between the Sender and the Receiver**. The decisions explained in the following will be made **simultaneously**.

The **Sender** decides on **how to divide the 22 Taler** between herself and the Receiver. If you are a Sender you will see a computer screen displaying all possible partitions of the 22 Taler. **By clicking on the corresponding button**, the Sender decides on a **particular partition**. Each Sender can click on one button only.

The **Receiver** decides on **acceptance or rejection** of the Sender's proposed partition. He therefore has to state the minimal amount of Taler that he has to receive in order to accept the partition. If you are a Receiver you will see a computer screen displaying all possible partitions of the 22 Taler. **By clicking on the corresponding button** the Receiver decides on the **minimal amount of Taler** he needs to receive in order to **accept the partition**. Each Receiver can click on one button only. The Receiver decides before knowing the Sender's actual offer. If the actual offer of the Sender is higher or equal to the minimal amount the Receiver is willing to accept, the partition is accepted. Conversely, the partition is rejected if the Sender's actual offer is smaller than the minimal amount the Receiver is willing to accept.

After both the Sender and the Receiver have made their decisions, choices are compared. If the Receiver is willing to **accept** the partition chosen by the Sender

- the amount is divided according to the Sender's choice.

If the Receiver is **not** willing to accept the partition

- the Sender and the Receiver are paid a guaranteed amount called **conflict payment**. The conflict payment can be identical or distinct for the Sender and the Receiver. The conflict payment is known to both the Sender and the Receiver before each of them is making his/her decision.

After having made your decision, you will be confronted with other situations characterized by **varying** conflict payments for Senders and Receivers. In these situations as well, the **Sender** decides **on the partition of the 22 Taler** within the dyad and the **Receiver** decides on the **minimal partition he is willing to accept**. You will not be informed about your co-player's choice in the previous situation.

Only **one** of the decision situations will be **relevant for final payment**. The decision you made in this situation is decisive for the payment from the experiment. After having made their decisions in **all** situations, both Senders and Receivers will be informed on the situation that was **chosen** and **thus is relevant for your payment**.

- If the Sender's partition was **accepted** the amount will be allocated according to the Sender's choice.
- If the Sender's partition was **rejected** each participant receives his/her conflict payment.

Your payment will correspond to your decision in the **chosen** situation. Therefore, please reflect your decision **carefully** in **each** situation.

When taking your decisions you will **not** be informed about the choices of the other person in your dyad. **At the end of the first part** of the experiment, however, we will communicate the decision the other person in your dyad has taken in the **chosen** situation**.**

After the first part of the experiment we ask you to answer some questions.

Please read the introduction carefully again to make sure that you understood everything. If there are any **questions** left, please **raise your hand**. We will come to your cubicle and answer your questions.

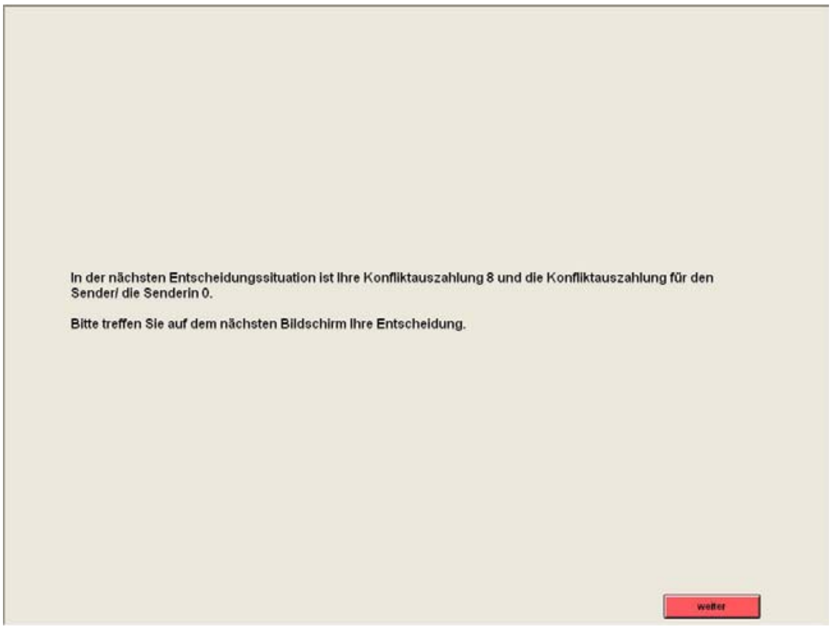
**Control questions (Original in German)**

Question 1 (2): The amount to be divided is 22 Taler. The conflict payment for both participants is 0 (3) Taler. Please think of the following result of the decision situation: the Sender offers 1 (21) Taler, the Receiver accepts this allocation. How many Taler do the Sender and the Receiver get?

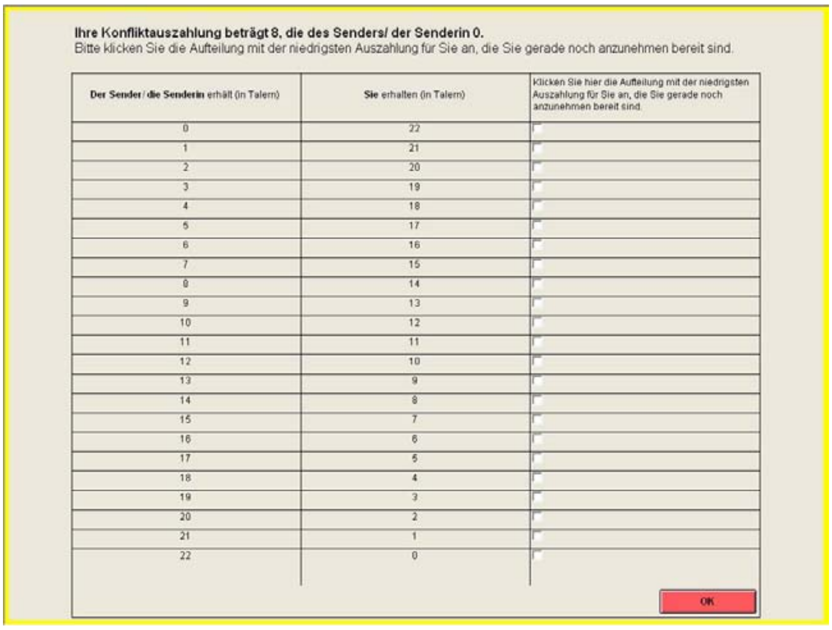
Question 3: The amount to be divided is 22 Taler. The conflict payment for the Sender is 5 Taler and for the Receiver 4 Taler. Please think of the following result of the decision situation: the Sender offers 0 Taler, the Receiver rejects this allocation. How many Taler do the Sender and the Receiver get?

**Experimental computer screens**

The following two pictures show the computer screens presented to the subjects. The first screen announces the specific outside options constellation that is relevant for the next decision situation. It is followed by the second screen picture in which subjects can click their desired devision of the pie.




Screen 1 for responder in C5/ T1



**Ihre Konfliktauszahlung beträgt 8, die des Senders/ der Senderin 0.**
Bitte klicken Sie die Aufteilung mit der niedrigsten Auszahlung für Sie an, die Sie gerade noch anzunehmen bereit sind.

| **Der Sender / die Senderin** erhält (in Talern) | **Sie** erhalten (in Talern) | Klicken Sie hier die Aufteilung mit der niedrigsten Auszahlung für Sie an, die Sie gerade noch anzunehmen bereit sind. |
|---|---|---|
| 0 | 22 | |
| 1 | 21 | |
| 2 | 20 | |
| 3 | 19 | |
| 4 | 18 | |
| 5 | 17 | |
| 6 | 16 | |
| 7 | 15 | |
| 8 | 14 | |
| 9 | 13 | |
| 10 | 12 | |
| 11 | 11 | |
| 12 | 10 | |
| 13 | 9 | |
| 14 | 8 | |
| 15 | 7 | |
| 16 | 6 | |
| 17 | 5 | |
| 18 | 4 | |
| 19 | 3 | |
| 20 | 2 | |
| 21 | 1 | |
| 22 | 0 | |

OK

Screen 2 for responder in C5/ T1

## 2.D Statistical tests

The choice of the appropriate test statistic is a very important issue for the evaluation of experimental data. In general, there are two types of tests available, parametric and non-parametric tests. Parametric tests assume a particular distribution function (usually the normal distribution) of the data and aim at rejecting hypotheses about specific parameters, such as mean or median. Although it has been shown that small deviations from the assumptions do not change results dramatically, it is not clear what constitutes a small deviation. Therefore in drawing inference from parametric testing, one has to be careful about possible mistakes. In contrast, nonparametric tests are based on a statistical model that specifies only very general conditions. In particular, nonparametric testing does not rely on a specific assumption concerning the distribution underlying the sampling. This is primarily important as sample sizes in experiments are usually limited and the central limit theorem is not readily applicable. Furthermore, non-parametric tests can be easily used with ordinal data. These are the main reasons for experimental economists to use mainly non-parametric tests. For further discussion on advantages and disadvantages of parametric and non-parametric tests, see Siegel and N. John Castellan (1988). In the following, I briefly explain the test statistics used in this study.

The **$\chi^2$-test** for two independent samples examines the significance of differences between two independent groups. The null hypothesis is that the two groups do not differ in a particular categorical characteristic. Under the null hypothesis, we thus expect the frequencies over the categories to be equal for both groups. The test is based on whether the deviations from the expected frequencies we observe in the data are large enough to exclude that they are the result of chance. The Fisher exact test draws on the same logic. However, the $\chi^2$-test is applicable in larger samples (where the number of observations exceeds 20), while the Fisher exact test performs better in smaller samples, for the statistical properties of the tests see Siegel and N. John Castellan (1988).

The McNemar change test investigates changes in a binary variable when two samples are related. The null hypothesis is that changes in either direction, that is from zero to one or from one to zero, are equally likely. Therefore the expected frequency for a change in either direction is simply the sum of the total changes divided by two. The test statistic gives us a measure of how likely the deviation from the expected change is. This test is only applicable when the total number of changes exceeds 10. If this is not the case, the approximation of the test statistic to be $\chi^2$ distributed is poor and it is better to use a version of the **Binomial test**. This test is very similar to the McNemar change test, only that it does not rely on the approximation. The Binomial test treats the changes as a binomial variable with parameter $\frac{1}{2}$. The statistic therefore builds on how likely it is that we observe particular frequencies given the distribution parameter of $\frac{1}{2}$. As most of the changes occurring in our data set are small in number, we use the Binomial test.

The **Wilcoxon-Mann-Whitney** test studies whether two independent groups are sampled from the same population where the characteristic in question is at least of ordinal type. The null hypothesis is that both groups are drawn from the same distribution. Under the null hypothesis, we expect that the probability that one observation is strictly larger than another equals $\frac{1}{2}$. The test relies on a comparison of the ranks assigned to each observation of the two groups. If high ranks are mainly assigned to one particular group while low ranks remain with the other group, this indicates that the better part of the observation of the first group is larger than the second group. The Wilcoxon-Mann-Whitney test calculates the probability that this particular ranking occurs under the null hypothesis.

# Chapter 3

# Ratification Quotas in International Agreements: An Example of Emission Reduction

## 3.1 Introduction

For the international provision of global public goods, country representatives gather and bargain over individual contribution levels. The resulting agreement not only specifies each country's contribution. Typically, it also contains a ratification quota, the minimum number of countries that have to ratify the agreement to render it legally binding for the ratifying countries. This paper explores the effect of a ratification quota on the provision of a trans-boundary public good. It focuses on two channels via which the quota might impact the provision of the public good. On the one hand, the higher the quota, the higher is the level of the public good provision whenever the agreement comes into effect. On the other hand, a higher quota may also increase the chance of a contractual breakdown, as an insufficient number of countries ratify the agreement. The present paper captures this trade-off in a three stage model where countries first determine the ratification quota, they then decide whether to ratify and at the last stage provide the public good. The three stage nature is inspired by reality where the national representatives first negotiate the agreement, before typically the legislative body of the country decides on ratification and therewith on the provision of the public good. I examine the differences between the situations where each country decides to ratify simultaneously or where it does so sequentially and identify circumstances where it is optimal to have a high or a low quota.[38]

I consider the special case of emission of a global pollutant, like carbondioxides fostering global warming. I restrict the setting to a binary type space; countries have either high or low benefits from local emission. When deciding on the ratification, countries know their own benefit parameter, but not the parameter of any other country. Their expectations of how many other countries ratify are thus crucial in the analysis. In their ratification decision, countries trade-off the expected gain with the expected costs of ratification. I distinguish between the case where the benefit of the high type is smaller or where it is larger than the social costs. With *simultaneous* ratification, I find that the optimal quota is 100% if individual benefits of the high type do not exceed the social costs of emission. Otherwise, if the high type's benefits are larger than the social costs, the optimal quota is as small as possible such that it still induces the low type to ratify. In this latter case, the optimal quota increases with the benefit parameter as well as with the probability of the low type. Furthermore, the optimal quota increases with the number of countries and the optimal quota relative to the number of countries decreases with the number of countries. With *sequential* ratification, the optimal quota is again 100% if the individual benefits of the high types are smaller than the social costs. However, when the individual benefit of the high type is larger than the social costs, the aggregate expected surplus decreases with respect to the simultaneous case. The sequential structure potentially discloses information inducing some low types to refrain from ratification, whereas in the simultaneous structure all low types ratify.

This paper is closely related to Black, Levi, and de Meza (1993). They simulate the effect of a minimum ratification quota on the provision of a global public good. In a simultaneous ratification setting where the types of countries are continuous, Black, Levi, and de Meza find that the optimal ratification quota relative to the number of countries is relatively robust to variations in the number of countries and in the cost parameter, but that it is not robust with respect to the distribution of benefits. In contrast, for the binary type setting I find that the optimal quota decreases with the number of countries, increases with the cost parameter and also increases with the probability of a low type. These differences stem from the different modelling of types. In a setting with binary types, all low types ratify; whereas with continuous types, there exists a critical value such that only those types exceeding this value ratify. This critical value is influenced by all model parameters.

A strand of coalition theory examines the provision of public goods and in particular

[38] In reality, ratification quotas differ. The Kyoto Protocol specifies that 55 countries (out of 166 signatories) which have to cover at least 55% of total emissions in 1990 have to ratify. The Convention for the Regulation of Whaling indicates 6 countries (out of 14 signatories) including the Netherlands, Norway, the Union of Soviet Socialist Republics, the UK, and the USA. Finally, for the International Criminal Court to come into force, 60 countries (out of 139 signatories) had to ratify. Mostly, ratification quotas are substantially different from 100% as well as from effectively no quota.

emission reduction using a cooperative game theory approach.[39] Part of this literature focuses on the effect of ratification quotas in international environmental treaties, called minimum participation levels in that literature. The paper by Rutz (2001) introduces a minimum participation rule in the context of a two-stage coalition model. It shows that a participation rule can potentially overcome the free-rider problem of transboundary pollution. Carraro, Marchiori, and Oreffice (2003) endogenise the participation rule by extending the model via a preceding stage where countries determine the minimum participation level. They establish that the grand coalition is stable with a 100% quota. Furthermore, they determine conditions under which all players agree to a 100% quota.

The literature on step-level public goods centers on the participation issue in the provision of discrete public goods. Palfrey and Rosenthal (1984) analyse participation to the provision of a binary public good where a fixed number of contributors is needed to provide the public good. In contrast to my paper, they assume complete information of the symmetric players and find that the efficient number of players contribute in equilibrium. They do not consider the role of the fixed number of contributors.

Next, I outline the model and its basic assumptions. For simultaneous ratification, section 3.3 characterises the optimal quota as well as some comparative statics results. Section 3.4 presents basic results for the case of sequential ratification. The last section summarises the results and indicates future paths of research.

## 3.2 Model

$N$ countries participate in an international bargaining process over the emission of a global pollutant. At the first stage, countries bargain over a ratification quota. A ratification quota $Q \in \{1, 2, ..., N\}$ is defined as the absolute number of countries that have to ratify the agreement to render it legally binding. Remember, only those countries that ratify the agreement will enforce the provisions of the agreement. At the second stage, each country decides whether to ratify. At the last stage, if a country ratified, it fulfills its obligations of the agreement, or otherwise emits freely.

Countries benefit from their own, local emission through production and consumption activities. But they suffer from the sum of all emissions, called global emission as these reduce environmental quality globally. The relation between the benefits of local emission and the harm of global emission is expressed by the parameter $\theta$. Countries' utility functions

[39] For example, Carraro and Siniscalco (1993), D'Aspremont, Jacquemin, Gabszewicz, and Weymark (1983), Finus and Rundshagen (2001), Barrett (1994) and Diamantoudi and Sartzetakis (2002) look at the provision of a public good within a two-stage coalition model. On the first stage, a single coalition is formed by simultaneous decisions of all countries. On the second stage, countries contribute in a static or dynamic game. In these models, the stable coalitions are generally small regardless of the number of participating countries. The grand coalition is always efficient, though not stable. For an introduction to the literature on international environmental agreements and coalition theory, see Barrett (2003), in particular chapter 7.

are captured by

$$u_i\left(e_i, \sum_{j=1}^{N} e_j\right) = \theta_i e_i - \sum_{j=1}^{N} e_j$$

where $e_i \in \{0,1\}$ indicates the emission level of country $i$. I treat the emission decision as binary.[40] Whenever a country is not member to the agreement, it chooses an emission level independently of all other countries' emissions. If the benefits of emission are small in relation to the costs, i.e. $\theta_i \leq 1$, country $i$ does not emit, while in the reverse case, whenever $\theta_i > 1$, it does emit.

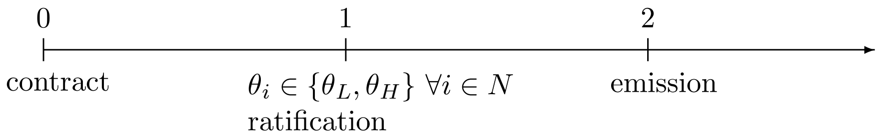
All countries that ratified the agreement satisfy their obligations in case the agreement gets legally binding. I neglect the problem of compliance. With this assumption and the specified utility function, the emission choice on the third stage of the game is uniquely determined.

At the time of deciding on a quota, countries have no information on any of the benefit-damage parameters $\theta_i$. However, they know the distribution function $F(\theta_i)$ which is assumed to be identical and independent for each country $i$. During the time which elapses between the bargaining over the agreement and its ratification, new information outcrops. Each country $i$ learns its own value of $\theta_i$, but not the realisations of the other countries' parameters. This formulation can be seen as a benchmark analysis to the case where countries do not learn the exact value of their benefit parameter, but get a more precise signal on it. I assume that the parameter can amount two distinct values, a low value $\theta_L$ and a high value $\theta_H$, with $\theta_L < \theta_H$. The probability of a low value is denoted by $p = prob(\theta = \theta_L)$. After the revelation of information, countries decide whether to ratify the agreement. I look at two distinct scenarios, the first where countries ratify simultaneously and the second where they do so sequentially, one after the other. In the latter version, when deciding on a quota, countries do not know the order of play. However, they know that it is equally likely to be in any position. Each possible order has the same probability of being drawn after the countries have agreed on the ratification quota. Figure 3.1 illustrates the time structure of the game.

In a *first best* world countries internalise the negative externality of their local emission on the rest of the world. As long as the benefit $\theta_i$ of local emissions to country $i$ is smaller than the social cost of local emissions $N$, country $i$ does not emit in a first best world. However, whenever countries are free to emit, they only take into account the damage of their local emission caused on themselves. They neglect the effect of their emission on all other countries. An agreement with a minimum participation level can help to ameliorate this classical free-rider problem. With a minimum participation rule, countries take into

[40] This assumption is particularly helpful as it reduces the dimension of the social maximisation problem. When it comes to the determination of the optimal emission level specified in the agreement, the interesting case is one in which no emission for the ratifying countries is desirable. The agreement therefore essentially determines a level of the ratification quota.

Figure 3.1: Timing of the game




account that they may eventually cause a breakdown of the agreement with their decision. An optimal quota is defined to be the number of necessary ratifications that maximises the aggregate expected surplus. A quota that achieves the first best outcome is clearly optimal. Whenever the first best level of pollution is achieved in the Nash equilibrium of the emission game, the design and ratification process of an agreement is of no interest. In these cases, an agreement cannot improve the allocation. The focus of this paper is thus on situations where the Nash equilibrium of the emission game does not achieve the first best outcome. As all players are symmetric at the time of deciding on the quota, the voting process is not essential. All players solve the same maximisation problem and thus decide unanimously on a quota. Therefore, I neglect the design of the voting mechanism.

In the ratification game, players have identical binary action sets $A_i = \{0, 1\}$ with actions labeled {do not ratify, ratify}. A strategy profile $s_i(\theta_i|Q)$ of country $i$ assigns an action $a_i \in A_i$ to every type $\theta_i \in \{\theta_L, \theta_H\}$ given a quota $Q \in \{1, ..., N\}$. There might be multiple equilibria in the ratification game. The aggregate expected surplus therefore depends not only on the ratification quota but also on the specific equilibrium.[41]

**Definition 1** *The optimal quota maximises the aggregated expected utility, i.e.* $Q^* = \arg\max_{Q\in\{1,...,N\}} \left\{\arg\max_{s^*(\theta|Q)} N * Eu\left(\left(s^*\left(\theta|Q\right)\right)\right)\right\}$ *where* $s^*\left(\theta|Q\right) = \left(s_1^*\left(\theta_1|Q\right), ..., s_N^*\left(\theta_N|Q\right)\right)$ *constitutes a symmetric Bayesian Nash Equilibrium (BNE) in pure strategies of the simultaneous ratification game or a symmetric Perfect Bayesian Equilibrium (PBE) in pure strategies of the sequential ratification game.*

The definition of the optimal quota relates to pure strategies. The focus of this paper is to ask what is implementable in pure strategies. The interpretation of mixed strategies is debated among game theorists, see chapter 3 in Osborne and Rubinstein (1994). Considering exclusively pure strategies, avoids the interpretation of mixed strategies in the ratification setting and finds the minimum expected utility that can be implemented in the stronger concept of pure strategies. However, in what follows, I also discuss how the results might change when we allow for mixed strategies.

[41] For the definition of Bayesian Nash Equilibrium and Perfect Bayesian Equilibrium, see chapter 2.6 and 12.3 in Osborne and Rubinstein (1994).

Next, I solve for the optimal quota in the simultaneous case and establish when it achieves the first best allocation. Furthermore, I conduct some comparative statics on the optimal quota.

## 3.3 Simultaneous ratification

In this section, I analyse the case where all countries ratify simultaneously. With simultaneous ratification, countries do not know how many other countries are ratifying. They can merely infer the probability that there is a sufficient number of ratifications such that the agreement gets legally binding. Suppose a country is pivotal, i.e. without this country's ratification the agreement fails. Then, ratification reduces global emissions by the sum of the reduced local emissions of all the ratifying countries. The expected gain of ratification equals the sum of reductions minus the benefit of local emission times the probability of being pivotal. Now, suppose a country is not pivotal. Ratification then implies that this country incurs costs amounting to the foregone benefits minus the damage of local emissions in case the agreement becomes binding. The expected loss of ratification of the country equals these costs times the probability that the quota is satisfied without this country's ratification. A country ratifies if the sum of these net expected gains from ratification exceed naught.

I distinguish between a situation where the agreement aims at inducing cooperation among all types and a situation where it aims at low types only. These two situations differ substantially. In the first case, the optimal quota is as large as possible, while in the second, the reverse is true. The following sections show the reasoning and intuition behind these results.

**Case 1: Participation of High Types** In this section, I analyse the cases where both types of countries prefer no emissions by any country to maximum emissions by every country. This occurs if the benefits of both types are smaller than the social costs of emission, $\theta_L, \theta_H \leq N$. In the first best allocation, no country emits pollutants and the sum of utilities is zero. However, whenever countries are free to decide on emission, they emit as long as their individual benefits are large enough. That is as long as $\theta > 1$.

An agreement, specifying that a certain amount of countries have to ratify before the agreement comes into force, can ameliorate the free-rider problem inherent in that situation. Suppose the ratification quota is 100%. Then, each country is pivotal for the emergence of the agreement. The potential loss in case of contractual breakdown is large, as emissions might rise from none at all to the maximal level of emissions $N$. This is an extreme scenario as the agreement induces all countries to take their decision on the background of comparing a situation with no emission to a situation with full emission by all countries.[42] Ratification of all countries and therefore the first best allocation can be achieved.

The above argument relies on the assumption that all countries, regardless of their type, emit if the agreement fails. However, benefits of the low type countries might be smaller than the individual costs, i.e. $\theta_L \leq 1$. Low types might thus abstain from emission regardless of the agreement. In this case, the agreement tries to establish cooperation foremost among the high types. Suppose again that the ratification quota is 100% and every country is pivotal. Whenever the agreement fails, the loss for the high type countries is smaller than in the above scenario as low type countries do not emit for sure. For the agreement to successfully lure high types to participation, the expected gain from ratification must exceed the loss $\theta_H$. Countries that ratify forego the damage of their own emission plus the expected damage of emissions by the remaining $N-1$ countries. Therefore, the expected gain of ratification amounts to $1+(N-1)(1-p)$. Given that gains exceed losses, ratification of all countries can be achieved by a 100% quota. Proposition 1 summarises this by characterising the conditions under which the symmetric pure strategy to ratify regardless of the benefit type is a BNE and stating the optimal quota.

**Proposition 1** *Given $\theta_H \leq N$, a quota $Q = N$ is a necessary condition for the symmetric pure strategy $s(\theta_j|Q) = 1$ for $\theta_j \in \{\theta_L, \theta_H\}$ to be a BNE. If $\theta_L > 1$, this is sufficient. Otherwise, additionally $\theta_H \leq 1 + (N-1)(1-p)$ has to be satisfied. The optimal quota is $Q^* = N$.*

**Proof.** If $Q < N$, there exists an incentive to deviate from the proposed equilibrium strategy. Given all other countries follow the proposed strategy to ratify regardless of their type, country $i$'s incentive to deviate is

$$\begin{aligned} D(\theta_j, Q < N) &= E(u(0|\theta_j, Q < N)) - E(u(1|\theta_j, Q < N)) \\ &= \theta_j - 1 \end{aligned} \tag{3.1}$$

for $j \in \{L, H\}$. At least for the high type $\theta_H$, equation (3.1) is positive and thus the country has an incentive to deviate from the proposed strategy.

If $Q = N$ and $\theta_L > 1$, there exists no incentive to deviate as each country is pivotal and $D(\theta_j, N) = \theta_j - N \leq 0$ for $j \in \{L, H\}$. If $Q = N$, but $\theta_L \leq 1$ holds, then there is no incentive

[42] Suppose one country abstains from ratification. One might think that it is not credible that there will be no cooperation among the ratifying countries. The literature on coalition formation shows, however, that the maximum number of countries forming a stable coalitions in public good environments is very small, see Carraro and Siniscalco (1993) and D'Aspremont, Jacquemin, Gabszewicz, and Weymark (1983). Therefore the situation is close to the one modelled.

to deviate for the high types if

$$\begin{aligned} D(\theta_H, N) &= \theta_H - 1 - \sum_{P=0}^{N-1} b(P|N-1,p)(N-1-P) \\ &= \theta_H - N + (N-1)p \le 0 \end{aligned}$$

where $b(P|N-1,p)$ denotes the binomial distribution
$b(P|N-1,p) = \begin{pmatrix} N-1 \\ P \end{pmatrix} p^P (1-p)^{N-1-P}$. This implies the stated condition.

For $\theta_L > 1$ or $\theta_L \le 1$ and $\theta_H \le 1 + (N-1)(1-p)$, a 100% quota implies that there is no global emission. Hence $Q = N$ achieves the first best outcome and is thus optimal. For $\theta_L \le 1$ and $\theta_H > 1 + (N-1)(1-p)$, all low types do not emit while all high types emit when there is no agreement. No agreement irrespective of the quota can improve on this result. By definition, any quota is therefore optimal. ■

The intuition is straightforward. If each country has a relatively small benefit from its own emission, a 100% quota forces participation of all countries and induces the first best allocation with no emission.[43] The restriction to pure strategies is innocuous in this case. Whenever the pure strategy to ratify regardless of the type can be implemented with a 100% quota, this quota achieves the first best and no other equilibrium can improve on that. Furthermore, if the low type's benefit $\theta_L$ is smaller than 1 and the condition that $\theta_H \le 1 + (N-1)(1-p)$ is not met, then mixed strategies cannot improve the outcome either. Keep in mind that in this situation we cannot induce high types to ratify even if all other countries do ratify. The expected gain from a working agreement is simply not large enough. If we allow high types to mix, then the expected utility of ratification decreases further. Appendix 3.A.1 illustrates that even though the expected utility of no ratification also decreases, it always exceeds the expected utility of ratification. There is thus no mixed strategy equilibrium.

We now turn to the situation where the benefit of the high types exceeds the social costs and high types are thus never going to participate.

**Case 2: Participation of Low Types** The situation is different when we look at cases where the high types' benefits are above the social costs, $\theta_H \ge N$. It is not possible to induce these countries to ratify an agreement which obliges them to abstain from emission.[44] Hence, an agreement can only build up cooperation among the low benefit types. The candidate symmetric pure strategy is to ratify if the country is of low type and to abstain otherwise,

[43] In the paper by Black, Levi, and de Meza (1993) this case is not considered. In their simulations, they assume that the benefit parameter is fixed $\theta > 1$ and the costs are drawn from the unit interval $c \in (0,1)$. Thus, types with low costs prefer to emit than to ratify.

[44] We abstract from the possibility of side payments, see the conclusion for further discussion.

$$s(\theta_i|Q) = \begin{cases} 1 & \text{if } \theta_i = \theta_L \\ 0 & \text{else} \end{cases}.$$

Given a ratification quota $Q$, there are $N - Q + 1$ possible states at the last stage of the game. There is either no binding agreement or an agreement that is binding for $P \in \{Q, ..., N\}$ countries. The incentive to deviate from the proposed strategy given all other $N-1$ countries follow the strategy is given by the incentive expression

$$\begin{aligned} D(\theta_L, Q) &= E(u(0|\theta_L, Q)) - E(u(1|\theta_L, Q)) \\ &= (\theta_L - 1)\sum_{P=Q}^{N-1} b(P|N-1, p) - (Q - \theta_L)\, b(Q-1|N-1, p), \end{aligned} \tag{3.2}$$

where $(P|N-1, p)$ again denotes the binomial distribution with parameters $N-1$ and $p$.[45] The proposed strategy to ratify if of low type and not otherwise is only sustainable in equilibrium if the incentive expression (3.2) is negative.

The first term captures the expected costs of ratification whenever the agreement gets legally binding and the country is not pivotal. It is the probability that more than $Q$ countries ratify times the benefit of local emission $\theta_L$ minus the damage of local emission of 1. Whereas the second term signifies the gain of ratification if the agreement gets binding and the country is pivotal, i.e. exactly $Q-1$ other countries ratified. A straightforward insight is that whenever there is no gain from ratification, that is, whenever the quota is smaller than the benefits, the proposed strategy cannot be sustained in equilibrium. This implies a lower bound on the optimal quota $Q > \theta_L$. It is not possible to solve analytically for the quota at which equation (3.2) equals zero (resp. is just negative). However, Proposition 2 shows that this minimum quota is optimal.

**Proposition 2** *With $1 < \theta_L \leq N < \theta_H$ and $p \in (0,1)$, there exists a unique optimal quota $Q^*$ that is the smallest integer for which $D(\theta_L, Q^*) \leq 0$.*

**Proof.** The aggregate expected surplus is maximised if the individual expected utility is maximised, $pE(u(s(\theta_L|Q))) + (1-p)E(u(s(\theta_H|Q))) \to \max_Q$. Given each country follows the strategy $s(\theta_i|Q) = \begin{cases} 1 & \text{if } \theta_i = \theta_L \\ 0 & \text{else} \end{cases}$, taking first differences of the individual expected utility yields

$$b(Q-1|N-1, p)\, p\, (\theta_L - N) < 0. \tag{3.3}$$

The derivation of (3.3) is found in Appendix 3.A.3. Thus, the quota that maximises individual and therewith aggregate surplus is as small as possible, still satisfying equation (3.2).

Next, I turn to characterise the incentive expression (3.2). For all $Q \in [1, \theta_L]$, the incentive expression is positive as there is no gain from being pivotal, i.e. $D(\theta_L, Q) \geq$

[45] A detailed derivation of the incentive expression can be found in Appendix 3.A.2.

$(\theta_L - 1) \sum_{i=\lceil\theta_L\rceil}^{N-1} b(i|N-1,p) > 0$. Moreover, whenever the quota is 100%, the incentive expression becomes strictly negative,
$D(\theta_L, N) = b(N-1|N-1,p)(\theta_L - N) < 0$. Taking first differences of the incentive expression using the fact that

$$b(Q|N-1,p) = b(Q-1|N-1,p)\frac{p}{1-p}\frac{N-Q}{Q} \tag{3.4}$$

yields

$$D(\theta_L, Q+1) - D(\theta_L, Q) = -b(Q-1|N-1,p)\left(\frac{pN-Q}{1-p} + \theta_L\right)$$

The incentive expression is therefore decreasing in $Q$ if $Q \leq Q' = p(N-\theta_L) + \theta_L$ and is increasing if $Q \geq Q' = p(N-\theta_L) + \theta_L$. This implies that $D(\theta_L, \cdot)$ jumps the x-axis exactly once. The minimum quota that still satisfies the incentive expression (3.2) is located at the jump (or just after). ■

Proposition 2 states that it is optimal to set the ratification quota as low as possible to the level where the low type country is just willing to ratify. Intuitively, this is appealing. If the quota is lower than this level, all the low type countries switch to a strategy of no ratification. The agreement does not get legally binding and every country emits. This cannot be optimal. Though, if the quota is higher, then all low type countries ratify. The higher quota implies that more countries have to be of low type. Therefore, the risk of breakdown of the agreement is increased as more low type countries are needed for ratification. This cannot be optimal either.

Next, I discuss some comparative statics results. The impact of an increase in the benefit parameter of the low type is straightforward. The incentive expression (3.2) increases with an increase in the low type parameter $\theta_L$, as the costs increase in case the country is not pivotal and at the same time the gains decrease in case the country is pivotal. The optimal quota must therefore be increased.

**Proposition 3** *With $1 < \theta_L \leq N < \theta_H$ and $p \in (0,1)$, the optimal quota $Q^*(\theta_L, p, N)$ is increasing in $\theta_L$.*

**Proof.** The incentive expression is increasing with $\theta_L$ as

$$\frac{\partial D(\theta_L, p, N)}{\partial \theta_L} = \sum_{i=Q-1}^{N-1} b(i|N-1) > 0$$

Furthermore, we know that the incentive expression decreases with $Q$ for all $Q \leq Q' = p(N-\theta_L) + \theta_L$ and increases for all $Q > Q' = p(N-\theta_L) + \theta_L$. As the incentive expression starts from a positive and ends with a negative value, the cutoff level $Q'$ has to be larger than the optimal quota, $Q^* \leq Q'$. The incentive expression is decreasing around the optimal quota

$Q^*$. This implies that the optimal quota $Q^*(\theta_L, p, N)$ increases with the benefit parameter $\theta_L$. ■

The impact of an increase in the probability of the low type is less straightforward. On the one hand, as the probability of the low type increases, the probability that there are sufficient ratifications from low type countries increases. On the other hand, the effect on the probability of being pivotal is undetermined; the probability could be increasing or decreasing. The following proposition shows that even if the probability of being pivotal increases with an increase of $p$, the gain associated with this is offset by the costs of the increased probability of sufficient ratifications. The incentive expression increases and the optimal quota increases with the probability of a low type.

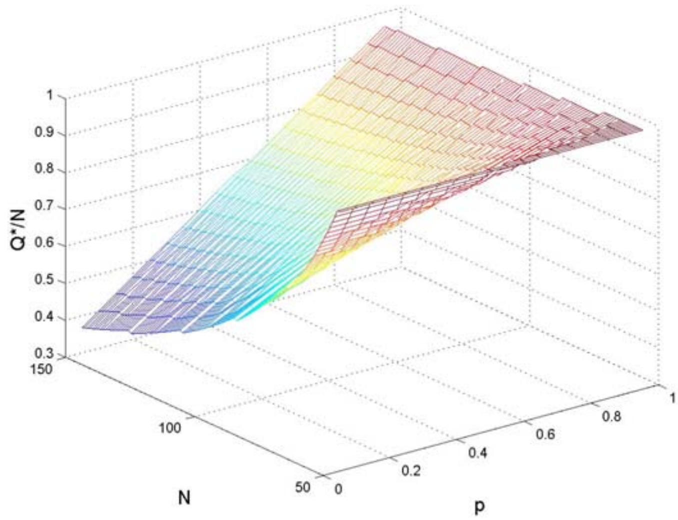
**Proposition 4** *With* $1 < \theta_L \leq N < \theta_H$ *and* $p \in (0,1)$*, the optimal quota* $Q^*(\theta_L, p, N)$ *is increasing in* $p \in (0,1)$*.*

The proof is delegated to Appendix 3.A.3. At first, the impact of a variation of the number of countries $N$ seems ambiguous. On the one hand, as the number of countries increases, the probability that there are sufficient ratifications from low type countries increases. This is due to the fact that the cumulative distribution function of the binomial distribution with parameter $N$ first order stochastically dominates the binomial distribution with parameter $N-1$. On the other hand, the effect on the probability of being pivotal is undetermined; the probability could be increasing or decreasing with an increase in $N$. Intuitively, the optimal quota should be increasing in the number of countries $N$. A constant or even decreasing quota with increasing $N$ does not seem plausible as the probability that sufficient countries out of the $N-1$ other countries are of low type $\sum_{P=Q}^{N-1} b(P|N-1, p)$ is converging to one when $N$ becomes large while the probability of being pivotal $b(Q-1|N-1, p)$ goes to zero. Hence, low type countries have an incentive to abstain from ratification. Therefore, we should observe an increase in the quota with $N$. Proposition 5 confirms this intuition.

**Proposition 5** *With* $1 < \theta_L \leq N < \theta_H$ *and* $p \in (0,1)$*, the optimal quota* $Q^*(\theta_L, p, N)$ *is increasing in* $N$*.*

The proof is delegated to Appendix 3.A.3. One might also be interested in how the relative optimal quota $q^*(\theta_L, p, N) = \frac{Q^*(\theta_L, p, N)}{N}$ evolves with an increasing number of countries $N$. To see this, I run simulations of the relative optimal quota over a parameter range of $N \in \{3, 4, ..., 150\}$, $p \in \{0.05, 0.1, ..., 0.95\}$ and $\theta_L \in \{2, 4, ..., N\}$. Figure 3.2 shows the simulated evolution of the relative optimal quota $q^*$ with the number of countries $N$ and the probability $p$ for a given $\theta_L = 50$. The value of $\theta_L$ is picked arbitrarily and is in no way particular. The x-axis depicts the number of countries $N$, while the y-axis shows the values of the probability of a low type $p$. The graph shows that the relative optimal quota starts with a value close

Figure 3.2: Relative Optimal Quota $Q^*/N$ for $\theta_L = 50$




to or equal to one at $N = \theta_L$, regardless of the probability. With an increasing number of countries $N$ the relative optimal quota decreases. This implies that the (absolute) optimal quota $Q^*$ increases underproportionately to the increase in $N$. Furthermore, the smaller the probability of the low type, the steeper is the decrease in the relative optimal quota. The graph also illustrates the result that the optimal quota is increasing in the probability of the low type $p$.

So far, we answered the question what can be optimally achieved when countries either ratify or not. Unlike case 1, the restriction to pure strategies might be binding. If low type countries chose to ratify with probability $q < 1$, it might be possible to increase aggregate expected utility by lowering the quota. Suppose all low types follow the symmetric mixed strategy to ratify with probability $q$. Given all other countries employ the mixed strategy, country $i$ has no incentive to deviate if the expected utility from ratification equals the expected utility from no ratification. The incentive expression (3.2), where the probability $p$ is substituted by the probability $\widetilde{p} = pq$, has to hold with equality,

$$(\theta_L - 1) \sum_{P=Q}^{N-1} b\left(P|N-1, \widetilde{p}\right) - (Q - \theta_L)\, b\left(Q-1|N-1, \widetilde{p}\right) = 0. \tag{3.2'}$$

Whether there exists a $q \neq 0$ that solves this equation, depends on the parameter constellations.[46] In the following example, we find that the modified incentive expression has got an inner solution with $q \in (0,1)$ and that the aggregate expected utility increases in comparison to the pure strategy equilibrium. Suppose the number of countries is $N = 7$, the low benefit parameter is $\theta_L = 4$ and the probability of a low type is $p = \frac{1}{2}$. With pure strategies, the incentive expression (3.2) is positive for a quota of $Q = 5$ at $D(4,5) = \frac{6}{64}$ and becomes negative for $Q = 6$ with $D(4,6) = -\frac{9}{128}$. Thus, the optimal quota is $Q = 6$ resulting in an expected aggregate utility of $-2.8359 + \frac{1}{2}\theta_H$. Contrary, when we allow for mixed strategies, the incentive expression (3.2') equals zero at the mixing probability $q = \frac{8-\sqrt{14}}{5}$ for a quota of $Q = 5$. The attached aggregate expected utility amounts to $-2.7252 + \frac{1}{2}\theta_H$, which exceeds the expected utility with pure strategies.

Simulations can give an indication how results change when we allow for mixed strategies. I simulate the scenario with mixed strategies for a parameter range of $N \in \{3, 4, ..., 150\}$, $\theta_L \in \{2, 4, ..., N\}$ and $p \in \{0.05, 0.1, ..., 0.95\}$. First, I compute the probability of ratification $q(Q)$ given a quota $Q$. Next, I calculate the quota $Q^*$ that maximises the aggregate expected utility given the probability $q(Q^*)$. For a low type probability of $p = 0.5$, Figure 3.3 shows the simulated mixing probability $q(Q^*)$ given the optimal quota $Q^*$. The x-axis depicts the number of countries $N$, while the y-axis shows the benefit parameter of the low type $\theta_L$. I find that the corner solution, where the probability is $q(Q^*) = 1$, is the rule rather than the exception. There is only an inner solution when the benefit parameter $\theta_L$ is relatively small. Moreover, the smallest value for the mixing probability is $q(Q^*) = 0.83$. It occurs when the number of countries is $N = 7$, the benefit parameter equals $\theta_L = 2$ and the optimal quota is $Q^* = 3$. The simulation results for $p = 0.5$ are in no way different to those with other values for $p$. We include further simulation results in Appendix 3.A.4. The simulation results illustrate that the limitation to pure strategies is not very restrictive and that it loses importance with an increasing benefit parameter $\theta_L$.

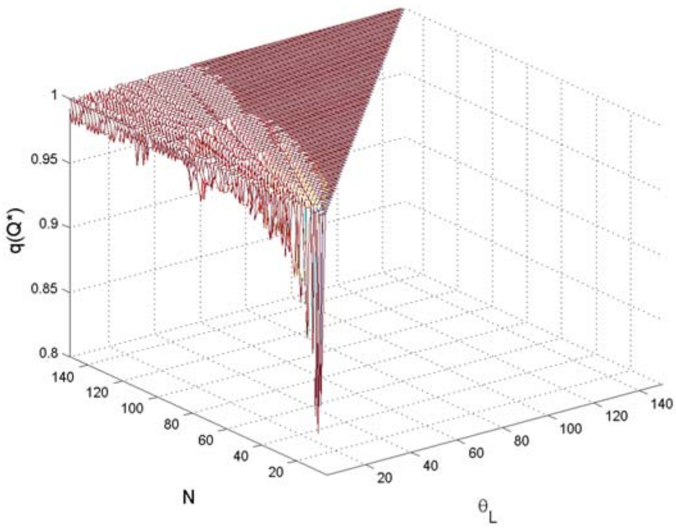
Summarising the case of simultaneous ratifications, we have seen that, if the benefit of the high type is smaller than the social costs, the optimal quota is 100% and the first best can be achieved. Otherwise, if the benefit of the high type exceeds the social costs, the optimal quota is as low as possible, taking into account that low type countries must still have an incentive to ratify. Thus, the results in theses two cases are diametrically opposed.

## 3.4 Sequential Ratification

In this section, I investigate the implications of sequential rather than simultaneous play. Instead of deciding all at once, countries ratify one after the other, observing the decisions

[46] As shown in Appendix 3.A.3, the incentive expression is increasing in $\widetilde{p}$ for all $\widetilde{p} \geq \frac{Q-\theta_L}{N-\theta_L}$. Thus, it starts in the origin, decreases with $q$ until $q = \frac{Q-\theta_L}{(N-\theta_L)p}$ and increases thereafter. If the reversal point $\frac{Q-\theta_L}{(N-\theta_L)p}$ is large, it is likely that there exists no solution $q \neq 0$. The expression is thus negative for all $q \neq 0$, and we are in a corner solution where each low type country choses to ratify with certainty, $q = 1$.

Figure 3.3: Mixing probability $q\left(Q^{*}\right)$ given $p = 0.5$




made by all previous countries. One crucial difference between these two scenarios is that with sequential ratification, the process of ratifying stops as soon as a sufficient number of ratifications occurred. No country has an incentive to keep on ratifying after the sufficient number of ratifications is reached as long as its benefit type exceeds one, i.e. $\theta_i > 1$. Therefore, with sequential ratification, the equilibrium outcome is either no binding agreement or an agreement with exactly $Q$ ratifying parties.

Sequential ratification is a game of incomplete information. The country only faces uncertainty regarding the type of the countries moving after it. The solution concept is Perfect Bayesian Equilibrium (PBE). In contrast to the case of simultaneous ratification, the PBE is essentially unique.[47] The history of the game can be summarised by the number of countries that have ratified so far. The restriction to pure strategies is innocuous when countries ratify sequentially. There exists no mixed strategy equilibrium. Consider the ratification decision of the last country when this country is pivotal. If the benefit parameter $\theta_i$ is smaller than $Q$, the last country's optimal strategy is to ratify with certainty. Further, if the second last country is pivotal, it is also going to ratify for sure. However, if it is not pivotal and the quota is not yet satisfied, then the country compares the expected utility from ratifying with the expected utility from not ratifying. Depending on the model parameters, one of the two expected utilities exceeds the other. It is thus optimal to play the pure strategy that leads to the higher expected utility. Only in the case, where the model parameters are such that

[47] It is unique up to a variation of the strategy after histories where it is not possible to achieve a sufficient number of ratifications. Suppose it is the turn of country $(N-i)$, $P$ countries have ratified before it and $Q$ have to ratify in total. A sufficient number of ratifications cannot be achieved, if too few countries have ratified so far, that is, if $P < Q - i - 1$.

both expected utilities are identical, does any mixed strategy belong to the set of optimal strategies. The same reasoning applies to all previous countries.

In analysing sequential ratification, I distinguish again between the two previous cases; the case where cooperation aims at the high types and where it aims at the low types. The outcome of the first case is largely the same as under simultaneous ratification, the optimal quota is as high as possible. Although I cannot determine the optimal quota in the second case, I find that the expected aggregate surplus is always higher under simultaneous than under sequential ratification.

**Case 1: Participation of High Types** This is the case where the benefit of local emission to both types of countries is smaller than the social damage of local emission, $\theta_H < N$. Again the optimal quota is 100% and it achieves the first best outcome, i.e. no emission by any country. Under a 100% quota each country is pivotal regardless of the order of play. The strategy to ratify regardless of the type and the history of the game is implementable if either $\theta_L > 1$ or if $\theta_L \leq 1$ and the condition of Proposition 1 that $\theta_H \leq 1 + (N-1)(1-p)$ is satisfied. In contrast to simultaneous ratifications, where there are multiple equilibria, this is the unique PBE. In case the low benefit types emit whenever the agreement fails, that is if $\theta_L > 1$, the intuition for an optimal quota of 100% is the same as in the simultaneous case. Countries, when ratifying, chose between a situation of no emission and a situation with full emission by all countries. Every country thus has an incentive to ratify. Contrary, if the low benefit types do not emit in any case, it has to hold that the benefits from emission do not outweigh the expected gains from ratification. That is, the condition of Proposition 1 that $\theta_H \leq 1+(N-1)(1-p)$ has to be met. This condition does not depend on the position in the order of play as the decisions of the previous countries do not reveal information. Effectively, the simultaneous and sequential specifications yield the same outcome.

Whenever the high type is large, i.e. $\theta_H > 1+(N-1)(1-p)$, then an agreement cannot help ameliorate the free-rider problem, just as in the simultaneous case. A 100% quota cannot induce cooperation among the high types. Lowering the quota reduces the expected gain of an agreement as fewer countries participate. Thus a lower quota is even less capable of inducing participation of the high types.

The result of the sequential game is therefore identical to the result in the simultaneous case. We either achieve the first best by implementing a 100% quota, or we cannot improve upon the situation at all.[48]

[48] The analysis of sequential ratification is closely related to sequential voting mechanisms. In particular, Dekel and Piccione (2000) show that in unanimity games, essentially the whole set of equilibria is the same in all sequential structures. There, sequential structures range from the one-period voting game, which would be the purely simultaneous case, over combined simultaneous and sequential structures to a purely sequential structure where each voter decides in a distinct period.

**Case 2: Participation of Low Types** In this case, high types never participate voluntarily in any agreement, as their individual benefits outweigh the social damage, $\theta_H \geq N$. The aggregate expected utility of the emission game is given by $U(0) = \sum_{i=0}^{N} \binom{N}{i} p^i (1-p)^{N-i} \left(i\theta_L + (N-i)\theta_H - N^2\right)$ if no contracting stage is preceding it.

A 100% ratification quota is clearly better than no agreement at all as it can achieve cooperation of all countries in the case where all countries are of low type. Each country is pivotal and the optimal strategy of a low type country is to ratify regardless of the history. The aggregate expected surplus is
$U(N) = \sum_{i=0}^{N-1} \binom{N}{i} p^i (1-p)^{N-i} \left(i\theta_L + (N-i)\theta_H - N^2\right)$. This is clearly better than no agreement as $U(N) - U(0) = -p^N N (\theta_L - N) \geq 0$.

Reducing the quota to $Q = N - 1$ implies that, on the one hand, one country free-rides surely, but on the other hand, the risk of a breakdown of the agreement is reduced. The optimal strategy for low types depends on the history of the game and on the position in the order of play. Whenever a country is pivotal,[49] the optimal strategy is to ratify. If it is not pivotal and the quota is not yet fulfilled, the country trades-off the gains from free-riding with the probability that the agreement fails due to its decision, and the associated loss. The probability that the agreement fails due to its decision depends on how many previous countries have ratified and how many more countries are to follow. If the benefits from emission are sufficiently large, the country takes the risk and tries to free-ride. Finally, if the quota is already satisfied, the optimal strategy is to abstain from ratification and free-ride. For a quota $Q = N - 1$, the optimal strategy for the $(N-i)^{\text{th}}$ country is given by

$$s^{N-i}(P|\theta_L, N-1) = \begin{cases} 1 & \begin{array}{l} \text{if } P \leq Q - i - 1 \text{ and } \theta_L \leq Q \text{ or} \\ \text{if } P = Q - i,\ i \neq 0 \text{ and } \theta_L \leq (1-p)\,Q \end{array} \\ 0 & \text{else} \end{cases}$$

for all $i \in \{0, ..., N-1\}$. $P$ signifies the number of countries that have already signed the agreement.[50] The proof of this strategy being the equilibrium strategy can be found in Appendix 3.B.1.

*Under the assumption that* $\theta_L \leq (1-p)\,Q$, the expected aggregate surplus increases with a smaller quota, if $p \leq \frac{N(N-1)}{N(N-1)+1}$. Otherwise, *under the assumption that* $\theta_L > (1-p)\,Q$, the expected aggregate surplus increases with a smaller quota, if $p \leq \frac{N-1}{N}$.[51] Both conditions are relatively mild, in particular if the number of countries $N$ is large. Thus, it is profitable

[49] Suppose $P$ countries have already ratified. Then, there need to be $Q - P$ more ratifications. A country $(N-i)$ is pivotal if the number of countries $i$ following that country equals the number of countries still needed to satisfy the ratification quota minus 1, i.e. $i = Q - P - 1$.

[50] The number of countries $P$ that have ratified before the $(N-i)^{\text{th}}$ country equals at most the number of countries preceding it, i.e. $P \leq N - i - 1 = Q - i$.

[51] For the derivation, see Appendix 3.B.1.

to reduce the quota from $Q = N$ to $Q = N - 1$ in most cases. This finding is intuitive as the gain from the reduction in the risk of breakdown of the agreement is high if the quota is large. Yet, the loss remains small, as a quota of $Q = N - 1$ allows only one country to free-ride.

A further reduction in the quota complicates the equilibrium analysis. Yet, countries still face the basic trade-off between risk of failure and free-riding. A complete characterisation of the PBE and therewith the optimal quota is cumbersome since it relies on many case distinctions, as the simple example of a $(N - 1)$-quota illustrates.[52] In what follows, I characterise two special cases. In the first case, benefits from emission for low types are very high such that every country's incentive to free-ride is large. Countries therefore only ratify if they are pivotal. In the other extreme, benefits from emission for low types are small and each country rather prefers to ratify than to risk the failure of the agreement. Proposition 6 characterises these two cases.

**Proposition 6** *Given a quota $Q$, the optimal strategy for a low type country at position $(N - i)$ in the order of ratification*

***Case 1)*** *if $Q \geq \theta_L \geq \overline{\theta}(p, Q) = (1 - p) Q$, is*

$$s^{N-i}(P|\theta_L, Q) = \begin{cases} 1 & \text{if } P \leq Q - i - 1 \\ 0 & \text{else,} \end{cases}$$

***Case 2)*** *if $\theta_L \leq \underline{\theta}(p, Q) = (1 - p)^{N-Q} Q$, is*

$$s^{N-i}(P|\theta_L, Q) = \begin{cases} 1 & \text{if } P \leq Q - 1 \\ 0 & \text{else.} \end{cases}$$

**Proof.** **ad 1)** Suppose $Q \geq \theta_L \geq (1 - p) Q$ and suppose every country follows the proposed strategy, a country $(N - i)$ has no incentive to deviate: After histories where so few countries have ratified, that it is not possible to fulfill the quota with the remaining countries anyway, i.e.$P < Q-i-1$, to ratify is among the optimal actions. Whenever the number of participating countries is such that exactly $i + 1$ ratifications are still needed, $P = Q - i - 1$, the country is pivotal and it prefers to ratify, as long as $\theta_L \leq Q$. Suppose the country is not pivotal and $P = Q - i + k$ with $k \in \{0, ..., i\}$, that is, $i - k$ ratifications are needed for satisfying the quota. If country $(N - i)$ ratifies, the following $k + 1$ countries do not ratify, regardless of their type. The countries thereafter are pivotal and ratify if of low type. The expected utility is $E(u(1)) = p^{i-k-1}Q + \left(1 - p^{i-k-1}\right)\theta_L - N$. If, however, the country $(N - i)$ does not ratify, then only the $k$ following countries do not ratify, regardless of their type. The expected utility of no ratification is $E(u(0)) = \theta_L - N + p^{i-k}Q$. The $(N - i)^{th}$ country has an incentive to deviate from the proposed strategy if $E(u(1)) - E(u(0)) > 0$. That is equivalent to $(1 - p) Q > \theta_L$. As by assumption $\theta_L \geq (1 - p) Q$, no country has an incentive to deviate from the proposed strategy. Finally, if the quota is already satisfied, $P \geq Q$, the

[52] It would be interesting to simulate the sequential equilibrium depending on the model parameters $\theta_L$, $p$, and $N$, and compare the resulting optimal quota with the simultaneous case.

country has no incentive to ratify.
**ad 2)** Suppose $\theta_L \leq (1-p)^{N-Q} Q$ and suppose every country follows the proposed strategy to ratify as long as the quota is not satisfied. A country $(N-i)$ has no incentive to deviate: Suppose $P = Q - 1 - k$ with $k \in \begin{cases} \{0, 1, ..., Q-1\} & \text{if } i \leq N-Q \\ \{Q-N+i, ..., Q-1\} & \text{if } i > N-Q \end{cases}$, that is $k+1$ ratifications are needed for satisfying the quota. If $k > i$, then it is not possible to fulfill the quota with the remaining countries, i.e. $P < Q - i - 1$, one optimal action is to ratify. If instead $k \leq i$, the agreement is feasible. Regardless of country $(N-i)$, all following low type countries ratify until the ratification quota is met. The utility of ratification is thus $E(u(1)) = \sum_{j=k}^{i} b(j|i,p) Q + \left(1 - \sum_{j=k}^{i} b(j|i,p)\right) \theta_L - N$, whereas the expected utility of no ratification is $E(u(0)) = \sum_{j=k+1}^{i} b(j|i,p) Q + \theta_L - N$. There is a positive incentive to deviate if $E(u(0)) - E(u(1)) > 0$, which is equivalent to

$$\theta_L > \frac{b(k|i,p)}{\sum_{j=k}^{i} b(j|i,p)} Q = \widetilde{\theta}(k,i). \tag{3.5}$$

In Appendix 3.B.2, we show that the following ordering holds

$$\widetilde{\theta}(k-1, i-1) \leq \widetilde{\theta}(k,i) \leq \widetilde{\theta}(k, i-1).$$

- The incentive to deviate for countries following the $Q^{th}$ country, that is, for countries with $i \leq N-Q$, is never positive. Equation (3.5) can never be satisfied as by assumption $\theta_L \leq b(0|N-Q) Q = \widetilde{\theta}(0, N-Q)$ and $\widetilde{\theta}(0, N-Q) \leq \widetilde{\theta}(k,i)$ for all $i \leq N-Q$ and $k \geq 0$.

- The incentive to deviate for countries before the $Q^{th}$ country, that is for countries with $i > N-Q$, is also never positive as $\theta_L \leq \widetilde{\theta}(0, N-Q) \leq \widetilde{\theta}(1, N-Q+1) \leq \widetilde{\theta}(2, N-Q+2) \leq ... \leq \widetilde{\theta}(Q-1, N-1)$.

If $P \geq Q$, the contract comes into force for sure and each country prefers to free-ride. ■

Clearly, the threshold levels $\underline{\theta}(p,Q)$ and $\overline{\theta}(p,Q)$ for the benefit parameter depend both on the probability of the low type as well as on the quota. The higher the quota, the less likely is case 1 and the more likely is case 2. The reverse holds for the probability of the low type. The higher $p$, the more likely we are in case 1 and the less likely we are in case 2.

In particular, the second case is interesting where, regardless of the history, each low type country ratifies as long as the quota is not yet satisfied. In this case, the simultaneous as well as sequential representation lead to exactly the same probability of success of the agreement with a given quota $Q \geq Q^*$ larger than the optimal quota of the simultaneous case. However, under the sequential representation, exactly $Q$ countries ratify, whereas with simultaneous ratification, most probably more than $Q$ countries ratify. Thus, the expected aggregate surplus is larger under simultaneous than under sequential ratification for any quota. Furthermore, for any given quota, the probability of success of the agreement is largest whenever we are in the equilibrium of case 2. Therefore, the expected aggregate surplus is larger with simultaneous rather than sequential ratification. The following proposition summarises this result.

**Proposition 7** *For a given quota $Q \geq Q^*$, the expected aggregate surplus is higher under simultaneous than under sequential ratification.*

The proof is straightforward and therefore omitted. In the simultaneous ratification game, every low type country ratifies, given the quota is larger than the optimal quota $Q \geq Q^*$. Turning from simultaneous to sequential play, some low type countries might abstain from ratification with the same quota. That lowers, on the one hand, the number of participating countries and, on the other, the probability of a success of the agreement. It remains to be noted that ex post some countries are better off in the sequential rather than the simultaneous game.[53]

## 3.5 Concluding remarks

The present model, proposes a three stage international bargaining game where countries first determine a ratification quota. Then, they decide whether to ratify and finally they decide over emission of a global pollutant. In a setting with incomplete information, two country types and a binary contribution to the provision, I examine the differences between simultaneous and sequential ratification. When the benefits from emission of both types are smaller than the social costs, the outcome in the simultaneous case is essentially identical to the sequential case. The optimal quota is 100% and achieves the first best. With the high type's benefits exceeding the social costs, I find that the optimal quota is as small as possible, if ratification is simultaneous. In the sequential ratification case, I cannot determine the optimal quota. However, I find that the aggregate expected surplus decreases with respect to the simultaneous case.

The crucial assumption driving the results of the model is the informational structure. The risk of failure of the agreement is introduced through the (costless) acquisition of new information concerning the benefits of a country. In reality, ratification processes differ from country to country.[54] They often involve the legislative body of a country to decide on the acceptance of the agreement. These processes take a substantial amount of time. The

[53]This is related to a result in an early version of the paper by Börgers (2004) that explores the effect of sequential voting. Under certain circumstances, Börgers shows that sequential voting weakly Pareto-dominates simultaneous voting. With sequential voting fewer agents incur the costs of voting. Hence, the public good is provided at lower total costs. The crucial difference to the ratification game is that in the voting setting, the "amount" of the public good does not change with the number of voters.

[54]For an overview of the different processes within Europe, see Stoiber and Thurner (2000).

outcrop of new information can realistically occur during that period of time. In the case of the Kyoto Protocol, the research group of the Intergovernmental Panel on Climate Change has published several special reports as well as a new Assessment Report since the Kyoto Conference in 1997, which constituted the start of the ratification period. Moreover, in the time period between the signature and the ratification of an agreement, the internal political situation of a country as well as the political and economic relationships to other countries can change. All this can influence the benefits accruing from emission. I do not model the national political processes leading to the ratification decision explicitly. These are definitely very important, but beyond the scope of the present paper.

A major restriction of the model is the abstraction from compliance problems, in particular, as compliance could depend on the number of countries that have ratified the agreement. The more countries join the agreement, the larger is potentially the pressure from these countries on non-complying members. Furthermore, I do not allow for transfer payments. Transfer payments from the low benefit to the high benefit countries could potentially induce all countries to participate in an agreement. However, low benefit countries can have an incentive to pretend to be of high type.[55] The assumption of no side-payments allows me to concentrate on the participation decision of each country.[56]

When introducing sequential ratification, further issues arise such as renegotiation and the order of ratification. The exogenously given order of ratification allocates bargaining power in favour of countries that are positioned later in the order. These countries might be able to exploit the ratification of previous countries. During the ratification process of the Kyoto Protocol, the case of Denmark suggests that countries do renegotiate with others that have already committed to ratification. The European Union's target was that all member states ratify until the World Summit of Sustainable Development in Johannesburg 2002. Denmark threatened not to do so, if its share of the entire union's reduction burden would not be lowered. An extension of the model would therefore consider renegotiation. Endogenising the order of ratifications, by allowing each country to chose its ratification time, represents another interesting possibility to extend the model.[57]

[55] Generally, incentive problems are tackled by the mechanism design literature. For a good survey article, see Moore (1992). Mechanism design focuses on whether there exists a mechanism that implements the efficient level of a public good. In contrast to this, I postulate a given institution and analyse the provision of the public good within that institution.

[56] For a recent paper that analyses the role of transfer schemes in international environmental agreements within the framework of coalition theory, see Carraro, Eyckmans, and Finus (2005).

[57] There is some literature on dynamic games of voluntary contributions to a public project, for example Marx and Matthews (2000) and the literature cited there. In their paper, Marx and Matthews assume that players have perfect information concerning the utility functions of every player. Furthermore, they neglect the impact of the minimum participation rule.

# Appendix

## 3.A Simultaneous ratification

### 3.A.1 Case 1: Mixed strategies

In case 1 where $\theta_L < 1$ and $\theta_H > 1 + (N-1)(1-p) = \underline{\theta}_H(N,p)$ and countries ratify simultaneously, the restriction to pure strategies is innocuous. To see this, suppose the quota is $Q = N - s$ for $s \in \{0, 1, ..., N-1\}$ and all high type countries ratify with a probability $q \in (0,1)$. For this to be an equilibrium, the expected utility of *no ratification* has to equal the expected utility of *ratification*. However, in the following, I illustrate that the expected utility of no ratification exceeds the expected utility of ratification for all $q \in (0,1)$. The expected utility of no ratification consists of a) the probability that the agreement succeeds times the benefit $\theta_H$ minus the sum of all emissions and b) the probability of a failure times the benefit $\theta_H$ minus the associated aggregate emissions, i.e.
$E(u(0|\theta_H, N-s)) = \sum_{h=1}^{s}\left(\sum_{P=0}^{N-h} b(P|N-1,p)\, b(N-P-h|N-1-P,q)(\theta_H - h)\right)$
$+\sum_{P=0}^{N-s} b(P|N-1,p)\sum_{j=0}^{N-s-P-1} b(j|N-1-P,q)(\theta_H + P - N)$. The expected utility of ratification is defined analogously and equals
$E(u(1|\theta_H, N-s)) = -\sum_{h=1}^{s+1}\sum_{P=0}^{N-h} b(P|N-1,p)\, b(N-1-h-P|N-1-P,q)(h-1)$
$+\sum_{P=0}^{N-s-1} b(P|N-1,p)\sum_{j=0}^{N-s-P-2} b(j|N-1-P,q)(\theta_H + P - N)$. The difference in expected utilities is given by

$$\begin{aligned} D(s) &= E(u(0|\theta_H, N-s)) - E(u(1|\theta_H, N-s)) \\ &= \sum_{h=1}^{s}\left(\sum_{P=0}^{N-h} b(P|N-1,p)\, b(N-P-h|N-1-P,q)(\theta_H - 1)\right) \\ &\quad + \sum_{P=0}^{N-s-1} b(P|N-1,p)\, b(N-s-P-1|N-1-P,q)(\theta_H + P - (N-s)). \end{aligned}$$

From the initial condition that

$$\begin{aligned} \theta_H &> \underline{\theta}_H(N,p) \\ &\leftrightarrow \sum_{P=0}^{N-1} b(P|N-1,p)(\theta_H + P - N) > 0, \end{aligned}$$

we know that the difference is positive for the starting value $s = 0$,
$D(0) = \sum_{P=0}^{N-1} b(P|N-1,p)\, b(N-P-1|N-1-P,q)(\theta_H + P - N) > 0$. Furthermore, I checked numerically that the first differences $D(s) - D(s-1)$ are positive. To do this, I

simulated

$$
\begin{aligned}
& D(s) - D(s-1) \\
= \; & \sum_{P=0}^{N-s-1} b(P|N-1,p)\, b(N-s-P-1|N-1-P,q) \left(\theta_H + P - N + \frac{s}{1-q}\right)
\end{aligned}
$$

over a parameter range of $N \in \{3,4,...,50\}$, $p,q \in \{0.05, 0.1, ..., 0.95\}$ and $\theta_H \in \{\underline{\theta}_H(N,p), ..., N\}$. The simulations show that the difference $D(s) - D(s-1)$ is positive, implying that $D(\cdot)$ increases in $s$. As $D(\cdot)$ starts with a positive value at $s = 0$, it is positive for all $s$. There is thus no probability $q \in (0,1)$ that equates the expected utility of no ratification and the expected utility of ratification.

### 3.A.2 Case 2: Derivation of the incentive function

In section 3.3, case 2, the proposed symmetric pure strategy is to ratify if of low type and to abstain from ratification otherwise. The incentive function (3.2) gives the incentive to deviate from this strategy for a low type country given all other countries follow it. On the one hand, the expected utility of *no ratification* is given by a) the benefit of emission $\theta_L$, plus b) the damage of global emission if the agreement gets binding, that is $P \geq Q$, times the probability that this happens, plus c) the damage if the agreement fails times the probability, i.e. $E(u(0|\theta_L, Q)) = \theta_L - \sum_{P=Q}^{N-1} prob(P|N-1)(N-P) - \sum_{P=0}^{Q-1} prob(P|N-1)N$. On the other hand, the expected utility of *ratification* is given by a) the damage of global emission if the agreement gets binding, that is $P \geq Q-1$, times the attached probability, plus b) the damage if the agreement fails minus the benefit of local emission times the probability, i.e. $E(u(1|\theta_L, Q)) = -\sum_{P=Q-1}^{N-1} prob(P|N-1)(N-1-P) - \sum_{P=0}^{Q-2} prob(P|N-1)(N-\theta_L)$. The probability that $P$ countries ratify out of the $N-1$ remaining countries is given by the binomial distribution

$$
b(P|N-1,p) = \binom{N-1}{P} p^P (1-p)^{N-1-P}.
$$

The incentive function is thus

$$
\begin{aligned}
D(\theta_L, Q) &= E(u(0|\theta_L, Q)) - E(u(1|\theta_L, Q)) \\
&= (\theta_L - 1) \sum_{P=Q}^{N-1} b(P|N-1,p) - (Q-\theta_L)\, b(Q-1|N-1,p).
\end{aligned}
$$

### 3.A.3 Case 2: Proofs

**Details to the proof of Proposition 2** Proposition 2 establishes that the optimal quota is the smallest integer that renders the incentive expression equal to naught. To proof this, we first show that, given the proposed strategy $s(\theta_i) = \begin{cases} 1 & \text{if } \theta_i = \theta_L \\ 0 & \text{else} \end{cases}$, the individual expected utility is decreasing in the quota $Q$. The individual expected utility is

$$\begin{aligned} U(Q) &= pE(u(\theta_L|Q)) + (1-p)E(u(\theta_H|Q)) \\ &= p\left(-\sum_{P=Q-1}^{N-1} b(P|N-1,p)(N-1-P) + (\theta_L - N)\sum_{P=0}^{Q-2} b(P|N-1,p)\right) \\ &\quad +(1-p)\left(\theta_H - \sum_{P=Q}^{N-1} b(P|N-1,p)(N-P) - N\sum_{P=0}^{Q-1} b(P|N-1,p)\right). \end{aligned}$$

Taking first differences yields

$$\begin{aligned} & U(Q) - U(Q-1) \\ = \;& p\left(b(Q-1|N-1,p)(N-Q) + (\theta_L - N)\, b(Q-1|N-1,p)\right) \\ & +(1-p)\left(b(Q|N-1,p)(N-Q) - Nb(Q|N-1,p)\right) \\ = \;& b(Q-1|N-1,p)\left(p(-Q+\theta_L) - (1-p)\frac{p}{1-p}\frac{N-Q}{Q}Q\right) \text{ with the use of (3.4)} \\ = \;& b(Q-1|N-1,p)\, p\,(\theta_L - N) < 0. \end{aligned}$$

This shows that the individual expected utility is decreasing in the quota $Q$.

**Proof of Proposition 4** Proposition 4 states that the optimal quota is increasing the probability of a low type $p$. To proof this, I first show that the probability of sufficient ratifications is increasing in the probability of the low type in the following way $\frac{\partial \sum_{P=Q}^{N-1} b(P|N-1,p)}{\partial p} = \binom{N-1}{Q} Qp^{Q-1}(1-p)^{N-1-Q} > 0$. This is shown by induction.

1. Suppose the above holds for $Q$, than for $Q-1$: $\frac{\partial \sum_{P=Q-1}^{N-1} b(P|N-1,p)}{\partial p} = \frac{\partial \sum_{P=Q}^{N-1} b(P|N-1,p)}{\partial p} + \frac{\partial b(Q-1|N-1,p)}{\partial p}$ $= \binom{N-1}{Q} Qp^{Q-1}(1-p)^{N-1-Q} +$ $\binom{N-1}{Q-1}\left((Q-1)p^{Q-2}(1-p)^{N-Q} - (N-Q)p^{Q-1}(1-p)^{N-Q-1}\right)$. After a couple of transformations and using (3.4) this yields the required result $\frac{\partial \sum_{P=Q-1}^{N-1} b(P|N-1,p)}{\partial p} = \binom{N-1}{Q-1}(Q-1)p^{Q-2}(1-p)^{N-Q}$.

2. The statement holds for the starting value of $Q = N-1$, $\frac{\partial \sum_{P=N-1}^{N-1} b(P|N-1,p)}{\partial p} = \frac{\partial p^{N-1}}{\partial p} = (N-1)\,p^{N-2}$.

Next, I determine the sign of the derivative of the incentive expression with respect to $p$: $\frac{\partial D(\theta_L,p,Q)}{\partial p} = (\theta_L - 1)\frac{\partial \sum_{P=Q}^{N-1} b(P|N-1,p)}{\partial p} - (Q-\theta_L)\frac{\partial b(Q-1|N-1,p)}{\partial p}$
$= (\theta_L - 1)\begin{pmatrix} N-1 \\ Q \end{pmatrix} Qp^{Q-1}(1-p)^{N-1-Q}$
$-(Q-\theta_L)\begin{pmatrix} N-1 \\ Q-1 \end{pmatrix}\left(p^{Q-2}(1-p)^{N-Q-1}\right)((Q-1)(1-p)-(N-Q)\,p)$
After a couple of transformations and the use of $\begin{pmatrix} N-1 \\ Q \end{pmatrix} Q = \begin{pmatrix} N-1 \\ Q-1 \end{pmatrix}(N-Q)$, this is equivalent to

$$\frac{\partial D(\theta_L, p, N)}{\partial p} = \frac{b(Q-1|N-1,p)}{(1-p)\,p}(Q-1)((N-\theta_L)\,p-(Q-\theta_L))\begin{cases} \geq 0 & \forall p \geq \frac{Q-\theta_L}{N-\theta_L} \\ < 0 & \forall p < \frac{Q-\theta_L}{N-\theta_L} \end{cases},$$

for all $Q \geq \theta_L$.

Furthermore, we know from the proof of Proposition 3 that $Q^* \leq p(N-\theta_L)+\theta_L = Q'$. This implies that $\frac{Q^*-\theta_L}{N-\theta_L} \leq \frac{p(N-\theta_L)+\theta_L-\theta_L}{N-\theta_L} = p$. Hence, around the optimal quota the incentive expression is increasing with $p$ which translates into the optimal quota itself being weakly increasing with $p$. ■

**Proof of Proposition 5** Proposition 5 says that the optimal quota is increasing in the number of countries $N$. To see this, I first show that the binomial distribution function with parameter $N+1$ first order stochastically dominates the distribution with parameter $N$ in the following way $\sum_{P=0}^{x}(b(P|N,p) - b(P|N+1,p)) = b(x|N,p)\,p > 0$ for all $x \in \{0,1,...,N\}$. This is shown by induction.

1. Suppose the above holds for $x$, then for $x+1$, $\sum_{P=0}^{x+1}(b(P|N,p) - b(P|N+1,p))$
$= b(x|N,p)\,p + b(x+1|N,p) - b(x+1|N+1,p)$
$= \begin{pmatrix} N \\ x \end{pmatrix} p^{x+1}(1-p)^{N-x} + \begin{pmatrix} N \\ x+1 \end{pmatrix} p^{x+1}(1-p)^{N-x-1} - \begin{pmatrix} N+1 \\ x+1 \end{pmatrix} p^{x+1}(1-p)^{N-x}$
$= p^{x+1}(1-p)^{N-x-1}\left(\left(\begin{pmatrix} N \\ x \end{pmatrix} - \begin{pmatrix} N+1 \\ x+1 \end{pmatrix}\right)(1-p) + \begin{pmatrix} N \\ x+1 \end{pmatrix}\right)$
$= \begin{pmatrix} N \\ x+1 \end{pmatrix} p^{x+2}(1-p)^{N-x-1} = b(x+1|N,p)\,p$ where we use (3.6).

The two binomial coefficients $\begin{pmatrix} N \\ x \end{pmatrix}, \begin{pmatrix} N+1 \\ x+1 \end{pmatrix}$ can be combined such that

$$\begin{pmatrix} N \\ x \end{pmatrix} - \begin{pmatrix} N+1 \\ x+1 \end{pmatrix} = - \begin{pmatrix} N \\ x+1 \end{pmatrix} \tag{3.6}$$

2. For the starting value $x = 0$, we know that $\sum_{P=0}^{0} \left( b\left(P|N,p\right) - b\left(P|N+1,p\right)\right) = \left(1-p\right)^{N} p = b\left(0|N,p\right)p$.

Next, I establish that the incentive function increases with $N$ around the optimal quota,

$$\begin{aligned} D\left(N+1\right) - D\left(N\right) &= \left(\theta_L - 1\right)\left(\sum_{P=Q}^{N} b\left(P|N,p\right) - \sum_{P=Q}^{N-1} b\left(P|N-1,p\right)\right) \\ &\quad - \left(Q-\theta_L\right)\left(b\left(Q-1|N,p\right) - b\left(Q-1|N-1,p\right)\right). \end{aligned} \tag{3.7}$$

Using the fact that the binomial distribution with $N$ first order stochastically dominates the binomial with $N-1$, we get that

$$\begin{aligned} & \sum_{P=Q}^{N} b\left(P|N,p\right) - \sum_{P=Q}^{N-1} b\left(P|N-1,p\right) \\ = & \sum_{P=0}^{Q-1} b\left(P|N-1,p\right) - \sum_{P=0}^{Q-1} b\left(P|N,p\right) = b\left(Q-1|N-1,p\right)p. \end{aligned}$$

Substituting this into equation (3.7), we get

$$\begin{aligned} D\left(N+1\right) - D\left(N\right) &= \left(\theta_L - 1\right) b\left(Q-1|N-1,p\right)p \\ &\quad - \left(Q-\theta_L\right)\left(b\left(Q-1|N,p\right) - b\left(Q-1|N-1,p\right)\right) \\ &= -b\left(Q-1|N-1,p\right)p - Q\left(b\left(Q-1|N,p\right) - b\left(Q-1|N-1,p\right)\right) \\ &\quad + \theta_L\left(b\left(Q-1|N,p\right) - \left(1-p\right)b\left(Q-1|N-1,p\right)\right) \\ &= \frac{b\left(Q-1|N,p\right)}{N\left(1-p\right)}\left(Q-1\right)\left(p\left(N+1-\theta_L\right) + \theta_L - Q\right). \end{aligned}$$

We therefore know that the incentive function increases for all small $Q$ and decreases thereafter, i.e.

$$D\left(N+1\right) - D\left(N\right) \begin{cases} > 0 & p\left(N+1-\theta_L\right) + \theta_L > Q \\ \leq 0 & \text{else} \end{cases}.$$

As we know, the optimal quota $Q^*$ is smaller than $Q' = p\left(N-\theta_L\right) + \theta_L$. Therefore, the incentive expression increases around the optimal quota and the optimal quota increases in the number of countries $N$. ■

### 3.A.4 Case 2: Mixed strategies

The following two figures report simulation results on the optimal mixing strategy $q(Q^*)$ for different values of the low type probability $p$. The x-axis depicts the number of countries $N$, while the y-axis shows the benefit parameter of the low type $\theta_L$. Panels 1-4 of Figure 3.4 show the results for small probabilities $p$ and panels 1-4 of Figure 3.5 for large probabilities $p$. Low type countries optimally ratify with a relatively large probability $q(Q^*)$. The smallest probability over the entire parameter range is $q(Q^*) = 0.76$ with $Q^* = 5$. It occurs at $N = 21$, $p = 0.1$ and $\theta_L = 4$. Again, I find that the corner solution is the rule rather than the exception. There is only an inner solution when the benefit parameter $\theta_L$ is relatively small.

Figure 3.4: Mixing probability, $p \in \{0.1, ..., 0.4\}$

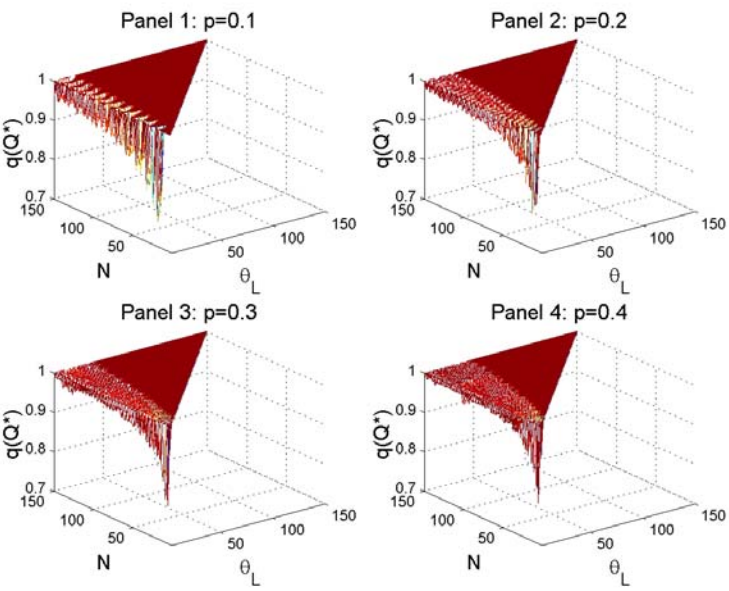
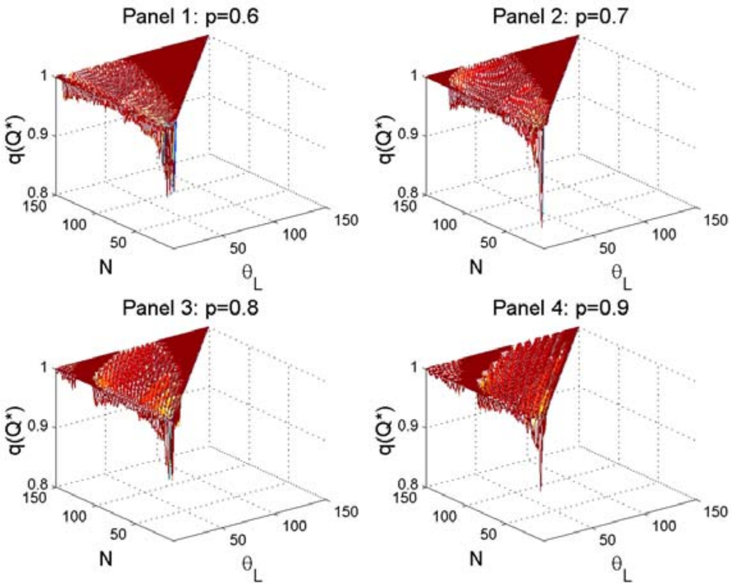


Figure 3.5: Mixing probability, $p \in \{0.6, ..., 0.9\}$




## 3.B Sequential ratification

### 3.B.1 Equilibrium with $Q = N - 1$

For a quota $Q = N - 1$, the proposed equilibrium strategy for the $(N-i)^{\text{th}}$ country is

$$s^{N-i}(P|\theta_L, N-1) = \begin{cases} 1 & \begin{array}{l} \text{if } P \leq Q - i - 1 \text{ and } \theta_L \leq Q \text{ or} \\ \text{if } P = Q - i,\ i \neq 0 \text{ and } \theta_L \leq (1-p)\,Q \end{array} \\ 0 & \text{else} \end{cases}$$

for all $i \in \{0, ..., N-1\}$. To see that this is indeed the optimal strategy, we distinguish between histories after which the country is pivotal, not pivotal or the quota is satisfied.

1) Suppose $P < Q - i - 1$. There are only $i$ countries to follow the $(N-i)^{\text{th}}$ country. Even if all $i+1$ countries ratify, the agreement does not get binding. It lacks at least one ratification. Therefore the strategy after this history is irrelevant. The proposed strategy belongs therefore to the optimal ones.

2) Suppose $P = Q - i - 1$. Then the $(N-i)^{\text{th}}$ country's decision is pivotal. As long as $\theta_L \leq N - 1 = Q$, it ratifies.

3) Suppose $P = Q - i$. For the last country $i = 0$, this implies that the quota is satisfied. The last country will therefore not ratify. For all other countries, the trade-off between free-riding and the increase in risk of contractual breakdown becomes relevant. Suppose

all countries follow the above strategy. If the $(N-i)^{\text{th}}$ country does not ratify, another $i$ ratifications are needed. All remaining $i$ countries ratify if they are of low type, as they are pivotal. The agreement gets binding with probability $p^i$. The expected utility of no ratification for the $(N-i)^{\text{th}}$ country is given by

$$E(u) = \theta_L - N + p^i Q. \tag{3.8}$$

If instead the $(N-i)^{\text{th}}$ country does ratify and only $i-1$ further ratifications are needed, then the optimal strategy of the remaining $i$ countries depends on the benefit parameter.

3.1) If $\theta_L \leq (1-p)\,Q$, then all countries, except the last, ratify whenever they are of low type. The probability of sufficient ratifications is $p^i + ip^{i-1}(1-p)$. The expected utility of ratification is given by

$$E(u) = \theta_L - N + \left(p^i + ip^{i-1}(1-p)\right)(Q - \theta_L). \tag{3.9}$$

The incentive to deviate from the proposed strategy for the $(N-i)^{\text{th}}$ country is given by the difference in expected utility (3.8)>(3.9)

$$\leftrightarrow \theta_L > \frac{i(1-p)}{(i-(i-1)\,p)}Q.$$

This condition contradicts the assumption that $\theta_L \leq (1-p)\,Q$. For all $i \geq 1$, it holds that $\frac{i(1-p)}{i-(i-1)p} \geq (1-p)$. Therefore, the $(N-i)^{\text{th}}$ country has no incentive to deviate from the proposed strategy to ratify.

3.2) If $\theta_L > (1-p)\,Q$, then the first country following the $(N-i)^{\text{th}}$ country does not ratify. All the remaining $i-1$ countries ratify, if they are of low type. As $i-1$ ratifications are still needed, the probability of sufficient ratifications is given by $p^{i-1}$. The expected utility of ratification is given by

$$E(u) = \left(1-p^{i-1}\right)\theta_L + p^{i-1}Q - N. \tag{3.10}$$

The incentive to deviate for the $(N-i)^{\text{th}}$ country is given by the difference in expected utility (3.10)>(3.8)

$$(1-p)\,Q > \theta_L.$$

This condition contradicts the assumption that $\theta_L > (1-p)\,Q$. Therefore the $(N-i)^{\text{th}}$ country has no incentive to deviate from the proposed strategy which is not to ratify.

The aggregate expected surplus *under the assumption that* $\theta_L \leq (1-p)\,Q$ is given by

$$\begin{aligned} U(N-1) &= p^N(\theta_L - N) + Np^{N-1}(1-p)(\theta_H - N) \\ &\quad + \sum_{P=0}^{N-2} \binom{N}{P} p^P (1-p)^{N-P} \left(P\theta_L + (N-P)\,\theta_H - N^2\right). \end{aligned}$$

It increases with a smaller quota if

$$\begin{aligned} & U(N-1) - U(N) \\ = {} & p^N(\theta_L - N) + Np^{N-1}(1-p)(\theta_H - N) \\ & + \sum_{P=0}^{N-2} \binom{N}{P} p^P (1-p)^{N-P} \left(P\theta_L + (N-P)\,\theta_H - N^2\right) \\ & - \left(\sum_{P=0}^{N-1} \binom{N}{P} p^P (1-p)^{N-P} \left(P\theta_L + (N-P)\,\theta_H - N^2\right)\right) \\ = {} & (N - \theta_L)\,p^{N-1}(N(N-1)(1-p) - p) \geq 0, \end{aligned}$$

which is equivalent to $p \leq \frac{N(N-1)}{N(N-1)+1}$. For a large number of countries $N$, this is a relatively mild condition. The aggregate expected surplus *under the assumption that* $\theta_L > (1-p)\,Q$ is given by

$$\begin{aligned} U(N-1) &= p^N(\theta_L - N) + p^{N-1}(1-p)(\theta_H - N) \\ &\quad + (N-1)\,p^{N-1}(1-p)\left((N-1)\,\theta_L + \theta_H - N^2\right) \\ &\quad + \sum_{P=0}^{N-2} \binom{N}{P} p^P (1-p)^{N-P} \left(P\theta_L + (N-P)\,\theta_H - N^2\right). \end{aligned}$$

Thus the expected aggregate surplus increases with a smaller quota, if

$$\begin{aligned} & U(N-1) - U(N) \\ = {} & p^N(\theta_L - N) + p^{N-1}(1-p)(\theta_H - N) \\ & + (N-1)\,p^{N-1}(1-p)\left((N-1)\,\theta_L + \theta_H - N^2\right) \\ & - \binom{N}{N-1} p^{N-1}(1-p)\left((N-1)\,\theta_L + \theta_H - N^2\right) \\ = {} & p^{N-1}(N - \theta_L)((1-p)(N-1) - p) \geq 0, \end{aligned}$$

which is equivalent to $p \leq \frac{N-1}{N}$. Again, for a large number of countries $N$, this is a mild condition.

### 3.B.2 Comparative statics on $\widetilde{\theta}(k,i)$

In the proof to Proposition 6, I derive a threshold level on the benefit parameter $\widetilde{\theta}(k,i) = \frac{b(k|i,p)}{\sum_{j=k}^{i} b(j|i,p)}$. Here, I show the ordering of the threshold levels to be

$$\widetilde{\theta}(k-1,i-1) \leq \widetilde{\theta}(k,i) \leq \widetilde{\theta}(k,i-1).$$

1. Show that $\widetilde{\theta}(k,i) \geq \widetilde{\theta}(k-1,i-1)$:

$$\begin{aligned}
& \frac{\binom{i}{k} p^k (1-p)^{i-k}}{\sum_{j=k}^{i} \binom{i}{j} p^j (1-p)^{i-j}} Q \geq \frac{\binom{i-1}{k-1} p^{k-1} (1-p)^{i-k}}{\sum_{j=k-1}^{i-1} \binom{i-1}{j} p^j (1-p)^{i-1-j}} Q \\
\leftrightarrow \quad & \frac{i}{k} \sum_{j=k}^{i} \binom{i-1}{j-1} p^j (1-p)^{i-j} \geq \sum_{j=k}^{i} \binom{i}{j} p^j (1-p)^{i-j} \\
\leftrightarrow \quad & \sum_{j=k}^{i} p^j (1-p)^{i-j} \binom{i-1}{j-1} \frac{i(j-k)}{kj} \geq 0
\end{aligned}$$

2. Show that $\widetilde{\theta}(k,i) \leq \widetilde{\theta}(k,i-1)$:

$$\begin{aligned}
& \frac{\binom{i}{k} p^k (1-p)^{i-k}}{\sum_{j=k}^{i} \binom{i}{j} p^j (1-p)^{i-j}} Q \leq \frac{\binom{i-1}{k} p^{k} (1-p)^{i-1-k}}{\sum_{j=k}^{i-1} \binom{i-1}{j} p^j (1-p)^{i-1-j}} Q \\
\leftrightarrow \quad & p^i \geq \sum_{j=k}^{i-1} p^j (1-p)^{i-j} \left( \frac{i}{i-k} \binom{i-1}{j} - \binom{i}{j} \right) \\
\leftrightarrow \quad & p^i \geq \sum_{j=k}^{i-1} \binom{i-1}{j} p^j (1-p)^{i-j} \left( \frac{i(k-j)}{(i-k)(i-j)} \right)
\end{aligned}$$

3. Show that $\widetilde{\theta}_L(k,i) \geq \widetilde{\theta}_L(k-1,i)$: Follows from 1 and 2.

# Bibliography

Babcock, L., and G. Loewenstein (1997): "Explaining Bargaining Impasse: The Role of Self-Serving Biases," *Journal of Economic Perspectives*, 11(1), 109–126.

Babcock, L., G. Loewenstein, S. Issacharoff, and C. Camerer (1995): "Biased Judgements of Fairness in Bargaining," *American Economic Review*, 85(5), 1337–1343.

Barrett, S. (1994): "Self-Enforcing International Environmental Agreements," *Oxford Economic Papers*, 46, 878–894.

——— (2003): *Environment and Statecraft: The Strategy of Environmental Treaty-Making.* Oxford University Press, New York.

Berg, J., J. Dickhaut, and K. McCabe (1995): "Trust, Reciprocity, and Social History," *Games and Economic Behavior*, 10(1), 122–142.

Black, J., M. D. Levi, and D. de Meza (1993): "Creating a Good Atmosphere: Minimum Participation for Tackling the 'Greenhouse Effect'," *Economica*, 60, 281–93.

Blount, S., and M. Bazerman (1996): "The Inconsistent Evaluation of Absolute versus Comparative Payoffs in Labor Supply and Bargaining," *Journal of Economic Behavior and Organization*, 30(2), 227–240.

Bolton, G. E., and A. Ockenfels (2000): "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review*, 90(1), 166–193.

Börgers, T. (2004): "Costly Voting," *American Economic Review*, 94(1), 57–66.

Brandts, J., and G. Charness (2000): "Hot Vs. Cold: Sequential Responses and Preference Stability in Experimental Games," *Experimental Economics*, 2(3), 227–238.

Brosig, J., J. Weimann, and C.-L. Yang (2003): "The Hot versus Cold Effect in a Simple Bargaining Experiment," *Experimental Economics*, 6(1), 75–90.

Buchan, N. R., R. T. A. Croson, and E. J. Johnson (2004): "When Do Fair Beliefs Influence Bargaining Behavior? Experimental Bargaining in Japan and the United States," *Journal of Consumer Research*, 31, 181–190.

CAMERER, C. F. (2003): *Behavioral Game Theory - Experiments in Strategic Interaction*. Princeton University Press, Princeton, New Jersey.

CARRARO, C., J. EYCKMANS, AND M. FINUS (2005): "Optimal Transfers and Participation Decisions in International Environmental Agreements," *FEEM Working Paper*, 50.2005.

CARRARO, C., C. MARCHIORI, AND S. OREFFICE (2003): "Endogenous Minimum Participation in International Environmental Treaties," *FEEM Working Paper*, 113.2003.

CARRARO, C., AND D. SINISCALCO (1993): "Strategies for the International Protection of the Environment," *Journal of Public Economics*, 52(3), 309–328.

CASON, T. N., AND V.-L. MUI (1998): "Social Influence on the Sequential Dictator Game," *Journal of Mathematical Psychology*, 42, 248–265.

CHARNESS, G., AND M. RABIN (2002): "Understanding Social Preference with Simple Tests," *Quarterly Journal of Economics*, 117(3), 817–869.

DAHL, G. D., AND M. R. RANSOM (1999): "Does Where You Stand Depend on Where You Sit? Tithing Donations and Self-Serving Bias," *American Economic Review*, 89, 703–727.

D'ASPREMONT, C., A. JACQUEMIN, J. J. GABSZEWICZ, AND J. A. WEYMARK (1983): "On the Stability of Collusive Price Leadership," *Canadian Journal of Economics*, 16(1), 17–25.

DEKEL, E., AND M. PICCIONE (2000): "Sequential Voting Procedures in Symmetric Binary Elections," *Journal of Political Economy*, 108(1), 34–55.

DIAMANTOUDI, E., AND E. S. SARTZETAKIS (2002): "Stable International Environmental Agreements: An Analytical Approach," *FEEM Working Paper*, 7.2002.

DUFWENBERG, M., AND U. GNEEZY (2000): "Measuring Beliefs in an Experimental Lost Wallet Game," *Games and Economic Behavior*, 30(2), 163–182.

FALK, A., E. FEHR, AND U. FISCHBACHER (2001): "Driving Forces of Informal Sanctions," *Institute for Empirical Research in Economics Working Paper*, 59.

FEHR, E., A. KLEIN, AND K. SCHMIDT (2001): "Fairness, Incentives and Contractual Incompleteness," *CESifo Working Papers*, 445.

FEHR, E., AND K. SCHMIDT (1999): "A Theory of Fairness, Competition and Cooperation," *Quarterly Journal of Economics*, 114, 817–868.

——— (2003): "Theories of Fairness and Reciprocity - Evidence and Economic Applications," in *Advances in Economics and Econometrics*, ed. by M. e. A. Dewatripont, pp. 208–257. Cambridge University Press, Cambridge, Mass.

Finus, M., and B. Rundshagen (2001): "Endogenous Coalition Formation in Global Pollution Control," *FEEM Working Paper*, 43.2001.

Fischbacher, U. (1999): "Z-Tree: Zurich Toolbox for Readymade Economic Experiments," *IEW Working paper*, 21.

Forsythe, R., J. L. Horowitz, N. E. Savin, and M. Sefton (1994): "Fairness in Simple Bargaining Experiments," *Games and Economic Behavior*, 6(3), 347–369.

Frohlich, N., and J. Oppenheimer (2001): "Some Doubts About Measuring Self-Interest Using Dictator Experiments: The Costs of Anonymity," *Journal of Economic Behavior and Organization*, 46, 271–290.

Frohlich, N., J. Oppenheimer, and A. Kurki (2004): "Modeling Other-Regarding Preferences and an Experimental Test," *Public Choice*, 119, 91–117.

Greene, W. H. (2000): *Econometric Analysis*. Prentice Hall International, Inc., New Jersey, 4. edn.

Greiner, B. (2004): "An Online Recruitment System for Economic Experiments," in *Forschung und Wissenschaftliches Rechnen 2003. GWDG Bericht 63*, ed. by V. M. Kurt Kremer, pp. 79–93. Ges. für Wiss. Datenverarbeitung, Göttingen.

Güth, W., S. Huck, and W. Müller (2001): "The Relevance of Equal Splits in Ultimatum Games," *Games and Economic Behavior*, 37, 161–169.

Hennig-Schmidt, H. (2002): "The Impact of Fairness on Decision Making - An Analysis of Different Video Experiments," in *Experimental Economics: Financial Markets, Auctions, and Decision Making*, ed. by F. Andersson, and H. Holm, pp. 185–210. Kluwer Academic Publishers, Boston, Dordrecht, London.

Hennig-Schmidt, H., Z.-Y. L., and C. Yang (2004): "Why People Reject Advantageous Offers - Non-Monotone Strategies in Ultimatum Bargaining," *Bonn Graduate School of Economics Discussion Paper*, 22/2004.

Hoffman, E., K. McCabe, K. Shachat, and V. Smith (1994): "Preferences, Property Rights, and Anonymity in Bargaining Games," *Games and Economic Behavior*, 7(3), 346–380.

Kagel, J. H., C. Kim, and D. Moser (1996): "Fairness in Ultimatum Games with Asymmetric Information and Asymmetric Payoffs," *Games and Economic Behaviour*, 13, 100–110.

Knez, M. J., and C. F. Camerer (1995): "Outside Options and Social Comparison in Three-Player Ultimatum Game Experiments," *Games and Economic Behavior*, 10, 65–94.

Kohnz, S. (2004): "Self-Serving Biases in Bargaining: Explaining Impasse," University of Munich, mimeo.

Konow, J. (2000): "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions," *American Economic Review*, 90, 1072–1091.

Kreps, D. M. (1990): "Corporate Culture and Economic Theory," in *Perspectives on Positive Political Economy*, ed. by J. Alt, and K. Shepsle. Cambridge University Press, Cambridge.

Marx, L., and S. Matthews (2000): "Dynamic Voluntary Contribution to a Public Project," *Review of Economic Studies*, 67, 327–358.

Messick, D., and K. Sentis (1979): "Fairness and Preference," *Journal of Experimental Social Psychology*, 15, 418–34.

Moore, J. (1992): "Implementation, Contracts, and Renegotiation in Environments with Complete Information," in *Advances in Economic Theory: Sixth World Congress of the Econometric Society, Vol.1*, ed. by J.-J. Laffront, Cambridge, UK. Cambridge University Press.

O'Donoghue, T., and M. Rabin (1999): "Doing It Now or Later," *American Economic Review*, 89(1), 103–124.

——— (2001): "Choice and Procrastination," *Quarterly Journal of Economics*, 161(1), 121–160.

Osborne, M. J., and A. Rubinstein (1994): *A Course in Game Theory*. MIT Press, Cambridge, Mass.

Oxoby, R. J., and K. N. McLeish (2004): "Sequential Decision and Strategy Vector Methods in Ultimatum Bargaining: Evidence on the Strength of Other-Regarding Behaviour," *Economics Letters*, 84, 399–405.

Palfrey, T. R., and H. Rosenthal (1984): "Participation and the Provision of Discrete Public Goods: A Strategic Analysis," *Journal of Public Economics*, 24(2), 171–193.

Roth, A. E. (1995): "Bargaining Experiments," in *Handbook of Experimental Economics*, ed. by J. Kagel, and A. E. Roth. Princeton University Press, Princeton, NJ.

Rutz, S. (2001): "Minimum Participation Rules and the Effectiveness of Multilateral Environmental Agreements," *Center for Economic Research Working Paper*, 01/22.

Schmitt, P. M. (2004): "On Perceptions of Fairness: The Role of Valuations, Outside Options, and Information in Ultimatum Bargaining Games," *Experimental Economics*, 7(1), 49–73.

SELTEN, R. (1967): "Die Strategiemethode Zur Erforschung Des Eingeschränkten Rationalen Verhaltens im Rahmen Eines Oligopolexperiments," in *Beiträge Zur Experimentellen Wirtschaftsforschung*, ed. by H. Sauermann, pp. 136–168. J.C.B. Mohr (Paul Siebeck), Tübingen.

SIEGEL, S., AND J. N. JOHN CASTELLAN (1988): *Nonparametric Statistics for the Behavorial Sciences*. McGraw-Hill, Inc., USA.

SMITH, A. (1759): *The Theory of Moral Sentiments*. A. Millar, London.

STOIBER, M., AND P. W. THURNER (2000): "Der Vergleich Von Ratifizierungsstrukturen der EU-Mitgliedsländer Für Intergouvernementale Verträge: Eine Anwendung Des Veto-Spieler Konzepts," *MZES Working Paper*, Nr. 27.

VAN HUYCK, J. B., R. C. BATTALIO, AND M. F. WALTERS (1995): "Commitment versus Discretion in the Peasant-Dictator Game," *Games and Economic Behavior*, 10(1), 143–170.

## Curriculum Vitae

| | |
|---|---|
| Februar 2006 | Ph.D. in economics (Dr. oec. publ.)<br>Ludwig-Maximilians-Universität, Munich |
| Sept. 2004 -<br>March 2005 | Marie-Curie Fellow<br>University College London |
| October 2002 -<br>Februar 2006 | Ph.D. program in economics<br>Munich Graduate School of Economics<br>at Ludwig-Maximilians-Universität, Munich |
| April 1996 -<br>Februar 2002 | Diplom in economics<br>Universität Mannheim |
| Sept. 2000 -<br>June 2001 | Affiliate student in economics<br>University College London |
| March -<br>June 1998 | Affiliate student in economics<br>Universidad de Sciencias Empresariales y<br>Sociales, Buenos Aires |
| 1987 to 1995 | Abitur<br>Kurfürst-Salentin-Gymnasium, Andernach |