Feedback Manipulation and Learning in Games

Feedback Manipulation and Learning in Games

This book was typeset by the author using LaTeX.

Feedback Manipulation and Learning in Games

Dissertation

to obtain the degree of Doctor at
Maastricht University,
on the authority of the Rector Magnificus,
Prof. dr. L.L.G. Soete,
in accordance with the decision of the Board of Deans,
to be defended in public
on Wednesday, October 7, 2015, at 12:00 hours

by

Aidas Masiliūnas

Supervisors:
Prof. Dr. Arno Riedl
Prof. Dr. J. Philipp Reiß, Karlsruhe Institute of Technology, Germany

Co-supervisor:
Dr. Friederike Mengel

Assessment Committee:
Prof. Dr. Rudolf Müller (chairman)
Prof. Dr. Jordi Brandts, Universitat Autònoma de Barcelona, Spain
Dr. Ronald Peeters
Prof. Dr. Martin Sefton, The University of Nottingham, United Kingdom

This research has been financially supported by the Graduate School of Economics (GSBE) at Maastricht University and by the European Union (grant PIEF 2009-235973)

Acknowledgements

First of all I would like to thank my supervisors, J. Philipp Reiss and Friederike Mengel, for providing guidance and help whenever necessary and at the same time giving me enough freedom to work on the topics that interest me. I have learnt a lot from you, not only about the immediate subject of the dissertation, but also about doing research and working in academia. "What did we learn from it?" is the first question that comes to my mind (in Philipp's voice) after hearing a presentation, reading a paper or starting a new project. I have also benefited from the visits in Karlsruhe and in Colchester, and I am very grateful for your hospitality.

I would also like to thank my promotor Arno Riedl for giving constructive feedback on the chapters of this thesis and on other ongoing projects. I thank the members of the assessment committee for taking time to read my thesis and for providing feedback.

While the first chapters for this thesis and the ideas for subsequent ones were developed in Maastricht, Colchester and Karlsruhe, the final touches were made when I was visiting the Institut d'Analisi Economica in Barcelona. I would like to thank Jordi Brandts, who was a wonderful host and a great role model. I would also like to thank GSBE for proving a travel grant that made this stay possible.

I would also like to thank fellow PhD students and faculty members in Maastricht, Karlsruhe, Colchester and Barcelona for the interesting discussions at the university and outside of it. Sometimes it was useful to hear different opinions about my thesis, and sometimes it was good to talk about something other than work. And thanks to those who kept me company while running, playing table tennis and board games, watching movies and doing other activities that made the PhD life much more enjoyable.

Most of this thesis was written in my A4.05 office, which I had a pleasure to share with Volker, Emel and Florian. I have also enjoyed working in wonderful libraries in Maastricht, Colchester and Barcelona, often by shamelessly

exploiting the lax access policy to take advantage of the internet, space and electricity.

Finally, I would like to thank my family and friends in Vilnius, who have been supporting me during the entire PhD process. The portion of the thesis that was written in Lithuania may be small, but the challenge to explain my research without using jargon has helped me think about the way to apply experimental findings. Some of these applications have been elaborated in the valorization addendum to this thesis.

Contents

1 Introduction **1**

2 Behavioural Variation in Tullock Contests **5**

2.1 Introduction 5
2.2 Experimental Design 9
2.3 Conjectures 12
2.4 Results: Reducing Complexity 14
2.4.1 Behavioural Variation 14
2.4.2 Explanatory Power of Nash Equilibrium 17
2.4.3 Cognitive Abilities and Equilibrium Deviations 21
2.5 Preference Based Explanations 22
2.5.1 Risk Preferences 23
2.5.2 Social Preferences 25
2.5.3 Heterogeneity 26
2.6 Strategic Uncertainty and Stability 27
2.7 Conclusion 29

Appendices **30**

2.A Additional results 30
2.A.1 Regression on the Explanatory Power of Nash Equilibrium 30
2.A.2 Deviations from Theoretical Predictions across Blocks . . 30
2.A.3 Determinants of Behaviour 31
2.A.4 Convergence and Dynamics 32
2.A.5 Treatments with no Strategic Uncertainty and no Additional Learning Opportunities 33
2.A.6 Within-Subject and Between-Subject Variability 34
2.A.7 Cognitive Abilities and Equilibrium Deviations 35
2.A.8 Summary Statistics 35

2.B Instructions 36
2.B.1 Instruction for Treatment **dD** 36
2.B.2 Instruction for Treatment **eT** 39
2.B.3 Instructions for the Questionnaire 42

3 Inefficient Lock-in with Sophisticated and Myopic Players 45
3.1 Introduction 45
3.2 Sophisticated Player Equilibrium 48
3.3 Sophisticated Player Equilibrium in a Critical Mass Game 52
3.3.1 Critical Mass Game 52
3.3.2 Choice Function of Myopic Players 53
3.3.3 Undominated Action Plans of Sophisticated Players . . . 54
3.3.4 Optimal Switching Period for Myopic Players 59
3.3.5 Payoffs of Sophisticated Players 62
3.3.6 Characterisation of Symmetric Sophisticated Player Equilibria 64
3.4 Summary and Comparative Statics 71
3.4.1 Planning Horizon of the Sophisticated Players 71
3.4.2 Player Composition 72
3.4.3 Length of the History of Inefficient Coordination 73
3.5 Conclusion 74

Appendices 76
3.A Proofs 76
3.A.1 Case 2 and Case 3 of Proposition 3 76
3.A.2 Proof of Lemmas 78
3.A.3 Comparative Statics 86

4 Overcoming Coordination Failure in a Critical Mass Game 91
4.1 Introduction 91
4.1.1 Related Literature 94
4.2 Experimental Design 96
4.2.1 Action Choice 96
4.2.2 Action Disclosure, Belief Elicitation and Feedback 97
4.2.3 Procedure 98
4.2.4 Theoretical Predictions 99
4.3 Experimental Results and Analysis 101

4.3.1 Is Coordination Failure More Often Overcome with Low Action Disclosure Costs? . . . 101
4.3.2 Is Adaptation Explained by Belief Learning Models? . . 104
4.3.3 Individual Heterogeneity in *LOW* . . . 113
4.4 Discussion . . . 117
4.5 Concluding Remarks . . . 119

Appendices **121**
4.A Figures . . . 121
4.B Tables . . . 123
4.C Alternative Explanations for the Treatment Difference . . . 125
4.D Belief Elicitation Task . . . 127
4.E Risk and Social Preferences . . . 128
4.E.1 Preliminaries . . . 128
4.E.2 Risk Neutral, Self-interested Players . . . 128
4.E.3 Risk Attitudes . . . 129
4.E.4 Social Preferences . . . 130
4.F SVO Slider Measure Tasks . . . 131
4.G Instructions . . . 132
4.H Screenshots . . . 139

5 Conclusion **143**

Addendum: Valorization **145**

Bibliography **158**

Biography **159**

Chapter 1

Introduction

Predictions about choices in games are typically made using a Nash equilibrium or one of its refinements, but the method by which players should reach an equilibrium is subject to debate. One possibility is that non-equilibrium outcomes are ruled out by deduction, but such explanation requires questionable assumptions of common knowledge of rationality, unlimited cognitive abilities and correct beliefs (see Mailath, 1998, or Perea, 2007, for a discussion). An alternative interpretation is that Nash equilibrium emerges as a long-run outcome of a dynamic process such as belief learning (Kalai and Lehrer, 1993) or replicator dynamics (Weibull, 1995). Convergence through learning requires weaker assumptions about rationality, but effective learning does require accurate feedback (Tversky and Kahneman, 1986): feedback about payoffs helps identify and avoid strategies that perform badly (reinforcement learning, e.g. Erev and Roth, 1998), while feedback about actions taken by others helps forming beliefs that are necessary for best-response dynamics (belief-based learning, e.g. Cheung and Friedman, 1997).

Chapters in this thesis investigate how manipulations that affect these types of feedback but not the set of equilibria influence the learning process and the long-run outcomes of a game. Chapter 2 reports the results of an experiment in which the reliability of feedback is manipulated by reducing the variability in opponent's actions and in the payoff function, leading to dramatically improved convergence rates and reduced behavioural variation in a contest game. A theoretical model presented in Chapter 3 specifies conditions under which sophisticated players would be willing to manipulate the feedback observed by myopic players by taking a suboptimal action in a repeated coordination game. An experiment in Chapter 4 varies the cost of action disclosure and shows that higher costs discourage action disclosure and lead to incomplete feedback that

prevents transitions to an efficient equilibrium.

Chapter 2 uses a Tullock contest as an example of a game in which game theory fails to make accurate predictions: Nash equilibrium is chosen only 7% of the time while dominated strategies are chosen more than half of the time. A common view is that the low explanatory power of a Nash equilibrium results from the failure of assumptions about risk-neutral and self-interested behavior (see a review in Sheremeta, 2013). However, deviations from theoretical predictions may be a result of complexities that make optimization difficult for boundedly rational players. Understanding the source of the discrepancy between choices and theoretical predictions would point to the correct way to address this problem. If the discrepancy arises because of non-standard preferences, the problem of low explanatory power could be solved by adding such preferences to the theory. But if the discrepancy results from bounded rationality, it may be desirable to introduce interventions that would enable players to learn to behave in an optimal way.

One factor that may limit the opportunities to learn in contests is noisy feedback arising from a stochastic payoff function. If there were no stochastic elements in the payoff function, dominated strategies would always provide low earnings and given enough time players should learn to avoid such strategies. In standard contests, however, payoffs are stochastic due to strategic uncertainty—changes in the actions chosen by other players—and due to non-strategic uncertainty generated by a randomization device. In a 2x2 design we reduce both types of uncertainty: strategic uncertainty is lowered by matching players to computers who play the same action for a certain number of rounds and uncertainty about the prize allocation is lowered by paying the expected value instead of playing a lottery. We find that when no uncertainty is present, the frequency of dominated actions decreases dramatically and the median response is almost always equal to the theoretical prediction. When either type of uncertainty is present, choices are very different from the theoretical prediction and dominated strategies are chosen more than half of the time.

Feedback may be manipulated by the experimenter, but it also depends on the actions taken by other players. If feedback affects behavior, a strategic player may therefore attempt to alter the future behavior of other group members by manipulating their observed feedback. A certain action could therefore be chosen just for the information it conveys to other group members rather than for the immediate profit it generates. **Chapter 3** and **chapter 4** use theory and experiments to test whether such strategic motives could explain deviations from an inefficient equilibrium in N-person critical mass games. In

these games non-equilibrium outcomes are observed much less frequently than in contests, as typically all groups rapidly coordinate on a common action. However, a social dilemma arises if play converges to an inefficient equilibrium: any unilateral deviation from an inefficient equilibrium would decrease deviant's immediate utility, even though a Pareto improvement could be achieved by collective action. Inefficient equilibrium persists because staying in it is optimal given the belief that others will stay too. If beliefs are the core of the problem, players who anticipate the belief formation process may maximize their earnings by deviating from an inefficient equilibrium because a decrease in immediate payoffs is more than compensated by the benefits of efficient coordination in the future. Chapter 3 shows that under certain conditions self-interested strategic players would be willing to deviate from an inefficient equilibrium. Chapter 4 tests whether players in an experiment are actually motivated by such strategic considerations and whether transitions to an efficient equilibrium are more frequent when strategic teaching is made easier by reducing the cost of action disclosure.

Chapter 3 proposes a solution concept that combines a learning model based on fictitious play and the concept of a Nash equilibrium by assuming that some players are myopic and update beliefs based on observed history while others are sophisticated and correctly anticipate the actions of all other players. The proposed solution concept is a combination of sophisticated player strategies that are optimal given the learning process of the myopic players and the strategies chosen by other sophisticated players. A combination of learning from history and strategic reasoning produces action paths in which sophisticated players find it optimal to use strategic teaching, and the existence of such action paths depends both on the history observed by myopic players and on the length of the reasoning horizon of sophisticated players.

Chapter 4 tests whether deviations from an inefficient equilibrium are motivated by strategic considerations and whether transitions to an efficient equilibrium are more likely if information about one's action can be provided at a low cost. In an experiment players could disclose their action to other group members by paying a cost that was varied across treatments. Players who only care about immediate payoffs would not be willing to pay the cost because disclosure provides no immediate benefits, but strategic players may do so if they expect that a disclosed action increases the chances of a transition to the efficient equilibrium. Data shows that many players are willing to pay to reveal their actions, especially when the costs are low, but only after choosing to deviate from an inefficient equilibrium. When disclosure costs are low, players

classified as more farsighted more often deviate from an inefficient equilibrium and disclose this action, providing further support for the strategic teaching hypothesis. Higher disclosure costs reduce the tendency to reveal actions, increasing strategic uncertainty and making transitions to an efficient equilibrium much less likely. In fact, no group moves to the efficient equilibrium when action disclosure costs are high, but half of the groups do so when the costs are low. Belief learning seems to be the most likely explanation for this treatment difference: stated beliefs generally move in accordance to the predictions of weighted fictitious play, but players who do not disclose their actions are perceived in a similar way to those who choose the inefficient action. Lack of feedback resulting from high action disclosure costs seems to introduce frictions into the learning process, reducing the frequency of transitions to the efficient equilibrium.

Chapter 2

Behavioural Variation in Tullock Contests[1]

2.1 Introduction

Many economic, political and social environments can be described as contests in which competing agents expend considerable resources (time, effort, money) in order to increase their chances of winning a "prize". Examples range from the competition for mates (Andersson and Iwasa, 1996), patents or research grants (Baye and Hoppe, 2003), to promotions or other relative reward schemes in firms (Chen, 2003), lobbying politicians (Baye et al., 1993), elections (Buchanan and Tullock, 1962), sports competitions (Szymanski, 2003), and ethnic conflicts (Esteban and Ray, 2011). Because of their many applications, these environments have attracted considerable attention in a wide range of fields, both in- and out-side of economics and there is a mature theoretical literature (for a survey see Konrad, 2009).

Experimental economists have tried to understand behaviour in contests empirically. The advantage of conducting experiments on contests (as opposed to empirical field studies) is that effort choices are observable and causal inferences can be drawn via treatment variations. One result that has emerged from this literature is that there is huge behavioural variation and Nash equilibrium (NE) has little explanatory power in typical Tullock contests (see Millner and Pratt, 1989, Potters et al., 1998, Sheremeta, 2010, among others). Figure 2.1 illustrates the behavioural variation in experimental Tullock contests and compares it to first- and second-price auctions.[2] Figure shows the cumulative distribution

[1] This chapter is based on the paper "Behavioural Variation in Tullock Contests" and is co-authored with Friederike Mengel and J. Philipp Reiß.

[2] Auctions and contests are both "competitive" allocation games, where everyone bids for a prize

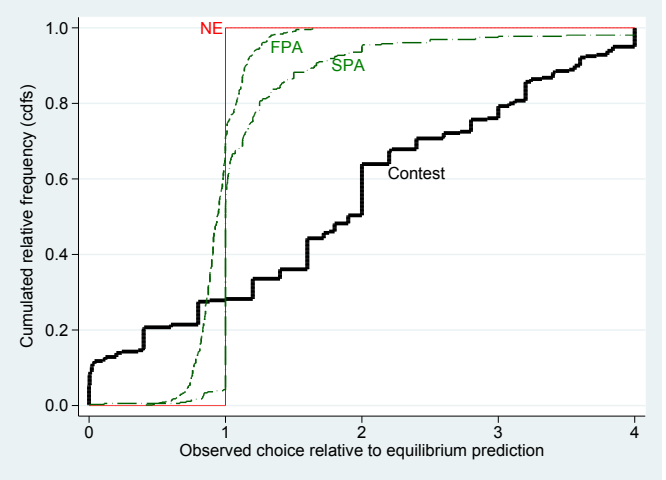


Figure 2.1: Cumulative Distribution of choices relative to Nash prediction in the contest (data from Abbink et al., 2010), first-price auction (data from Brosig and Reiss, 2007), second-price auction (data from Cooper and Fang, 2008) and according to Nash equilibrium.

of observed choices relative to the (risk neutral) Nash prediction for a typical Tullock contest experiment, a first-price auction (FPA), a second-price auction (SPA) and according to NE. According to theory, all the mass should be at 1, because all choices should equal the Nash prediction. Evidently, in the two auction formats (FPA and SPA), the cumulative distribution is pretty similar to theory. The Nash prediction clearly has something to say about the data here. In the contest, however, this pattern is completely different. Choices seem to have little to do with the unique pure strategy Nash equilibrium and investments are spread across the whole strategy space with no meaningful concentration around any specific value or range of values.

In this chapter we try to understand why behavioural variation is so large and why Nash equilibrium (understood more broadly than the risk neutral point prediction) has so little to say about the data in the standard contest. By contrast, most of the literature has focused on the so-called "overbidding phenomenon", the fact that *on average* investments are above the risk neutral Nash equilibrium. Explanations have focused on specific preferences or correlates of individual behaviour with specific forms of bounded rationality (see a survey by Sheremeta, 2013). Examples of preference based explanations include spiteful preferences, inequality aversion (Bartling et al., 2009) or the "joy of winning hypothesis" (Schmitt et al., 2004, Cason et al., 2010). The "over-

but there is one winner. Two differences are that in auctions there is incomplete information, while there is complete information in the contest we study. In addition, in auctions the prize allocation is deterministic, while in the typical contest it is non-deterministic.

bidding phenomenon" has also been explained by QRE (Sheremeta, 2011, Lim et al., 2014), distortion of probabilities (Baharad and Nitzan, 2008) or learning (Fallucchi et al., 2013). However, none of these studies explicitly addresses or explains the large observed behavioral variation. It has been conjectured, though, that heterogeneity in preferences or demographics might be one explanation (Sheremeta, 2013).

In this chapter we try to explain both the large amount of behavioural variation observed and the low explanatory power of Nash equilibrium in these games. We distinguish between two classes of alternative explanations: one based on preferences, the other on bounded rationality. Potentially, there are two sources of complexity that can inhibit the explanatory power of NE if agents are boundedly rational: (i) the difficulty to form correct beliefs and (ii) the difficulty to formulate best responses (even if beliefs were correct). In our experiment we vary these two sources of complexity systematically. If complexity is indeed the underlying reason for the large behavioural variation and the low explanatory power of NE in this game, then we should see (approximate) Nash behavior once both these sources of complexity are removed.[3]

In our benchmark treatment **dD** participants play a standard Tullock contest. Treatment **eD** exogenously manipulates the difficulty to formulate best responses. This treatment coincides with **dD** except for the fact that prize allocation is deterministic. In the light of ample evidence that people find it difficult to reduce uncertainty (Kahneman and Tversky, 1972), this should make it easier to formulate best responses. Treatment **dT** exogenously manipulates the difficulty of forming correct beliefs. It coincides with **dD**, but participants in this treatment face computer opponents who play pre-determined actions that are announced to participants before they make their choices. Hence there is no strategic uncertainty in this treatment and forming beliefs about the opponent's choices is trivial. Finally, treatment **eT** coincides with **dT**, but prize allocation is again deterministic. Hence in this treatment both best response formulation and belief formation should be easy or trivial, respectively.

Our main findings are as follows. As in previous experiments, investments in the standard contest (**dD**) are spread across the whole strategy space and Nash equilibrium has almost no explanatory power. When only one source of difficulty is removed (**eD** or **dT**), choices are still very different from the Nash prediction: more than half of the time players choose strategies that are strictly

[3]Since we do not observe preferences directly, it is hard to say what the "right" Nash equilibrium benchmark is. Virtually all of the existing literature uses the risk neutral NE as a benchmark (see survey below). We also use risk neutral expected utility maximizers as a benchmark, but, in addition, we also consider a wide range of other preferences. The classes of preferences we study include different risk preferences and social preferences.

dominated and investments choices are highly variable. However, when both best response and belief formation are easy or trivial, respectively (**eT**), players choose strategies that are very close to the prediction of the risk neutral model and behavioural variation is much lower. Bounded rationality can therefore explain the large behavioural variation and low explanatory power of Nash equilibrium in this game.

In order to understand possible interaction effects between our treatment manipulations and participants' (social) preferences, we also conducted a treatment that coincides with **eT**, but replaces computer players by humans. We show that our treatment rankings are unaffected by this change. We also show that all our treatment comparisons and conclusions remain valid if we allow for risk-aversion, risk seeking preferences or different types of social preferences.

We then ask whether heterogeneity in preferences can explain the large behavioural variation in the standard contest. To this end we conduct extensive simulations where we consider many possible population compositions of agents with differing risk, social and joy of winning preferences and show that none of these is able to recover the behavioural variation found in the standard contest (**dD**). We conclude that, while some preference based explanations can explain the so-called "overbidding phenomenon", they cannot explain the large behavioural variation typically found in these games. Reducing the complexity of the environment, on the other hand, eliminates the large behavioural variation and leads to behaviour that is very consistent with the predictions of the risk neutral model.

Complexity has been found to play an important role in other games and decision-problems. Grimm and Mengel (2012) show that players are able to learn the Nash equilibrium in normal form games with a unique pure strategy Nash equilibrium in situations of low complexity (few games and easy access to feedback), but not in situations with higher complexity (many games or difficult access to feedback). Huck et al. (2010) show that participants have more difficulty to form correct beliefs as the environment gets more complex. In decision problems Huck and Weizsäcker (1999) find that players are more likely to deviate from expected value maximization when choosing between a pair of lotteries if the task is more complex, as measured by the number of possible outcomes (see also Rabin and Weizsäcker, 2009).

The chapter is organized as follows. In section 2.2 we present the experimental design. Section 2.3 lists the conjectures that are tested in section 2.4, where we show how reducing complexity reduces behavioural variation and increases the explanatory power of Nash equilibrium. Preference-based explanations and

heterogeneity in preferences are discussed in section 2.5 and conclusions are drawn in section 2.7. The Appendix contains additional tables and figures as well as the experimental instructions.

2.2 Experimental Design

Our experimental design aims to understand whether bounded rationality could be the reason for the large behavioural variation and low explanatory power of Nash equilibrium in Tullock contests. In order to reach a Nash equilibrium, players need to do two things: (i) they need to be able to form correct beliefs and (ii) they need to be able to formulate best responses (to whatever beliefs they hold). Participants could learn to form correct beliefs or formulate best responses in a number of ways. Under some learning models they might even learn to play a Nash equilibrium without explicitly forming beliefs at all (e.g. reinforcement learning). How exactly participants do so is of secondary concern for us in this study. The question we ask is whether if we make it "easy enough" to formulate beliefs and best responses behavioural variation in this game will be (substantially) reduced and the explanatory power of Nash equilibrium increased. In a 2×2 factorial design we hence varied complexity along these two dimensions. Table 2.1 summarizes this treatment structure.

Treatment **dD** is our benchmark treatment and implements the standard contest as typically studied in the experimental literature. In treatment **eD**, we make it easy to formulate best responses, but keep the difficulty of forming correct beliefs. This reverses in treatment **dT**, where we keep the difficulty of formulating best responses and make it simple, in fact trivial, to formulate correct beliefs. In treatment **eT**, both sources of complexity are eliminated. Next, we describe these treatments in detail.

Table 2.1: Experimental design. Numbers in brackets indicate the total number of participants in each treatment. Participants in treatments **dD** and **eD** were allocated to matching groups of 6 participants, thus there are 9 independent observations in each of these treatments. In treatments **eT** and **dT** each participant is an independent observation. In **eT**$-h$ 24 participants were "dummy" players, so we use the choices of the remaining 24 players in our analysis.

	Belief Formation	
Best Response Formulation	Difficult	Trivial
Difficult	**dD** (54)	**dT** (44)
Easy	**eD** (54)	**eT** (44); **eT**$-h$ (48)

Treatment dD. In treatment **dD**, subjects participated in the standard Tullock contest. In our Tullock contest, two players compete for a commonly known prize of 16 Experimental Currency Units (ECU). In every round, participants received an endowment of 16 ECU. Participants could then invest an amount from this endowment, i.e. an amount from the action set $A_i = \{1, 2, ..., 16\}$ with typical element a_i $(i = 1, 2)$. Players' investments are sunk costs and the sets of feasible monetary payoffs are given by $\Pi_i = \{16 - a_i, 16 + 16 - a_i\}$. The contest success function that denotes the probability that player i receives the prize and, hence, the payoff of $(32 - a_i)$ ECU is given by $\rho_i(a_1, a_2) = \frac{a_i}{a_1+a_2}$. The experimental instructions can be found in Appendix 2.B.

The unique (risk neutral) Nash equilibrium in our contest game is given by $(a_1^*, a_2^*) = (4, 4)$. This is also the unique strategy profile surviving iterated elimination of strictly dominated strategies. All choices $a_i > 4$ are strictly dominated by $a_i = 4$ under the standard assumptions. Put differently, there are no beliefs that can support choices above the Nash choice of 4 for risk neutral expected utility maximizers. Even allowing for moderate degrees of risk aversion or risk seeking does not affect this property. To demonstrate this, Figure 2.6 illustrates the best response correspondence for a risk averse, risk neutral and risk seeking CRRA agent. These best response functions are very similar to one another and choices exceeding 4 are not rationalizable for any of these types. Therefore, we will follow the vast majority of the literature and use the risk neutral Nash equilibrium as a benchmark. We will be very careful, though, to account for heterogeneity in risk and social preferences when interpreting our results and we will investigate this issue more deeply in section 2.5.

Treatment eD. In treatment **eD** we wanted to make it easier for participants to formulate best responses (compared to **dD**). An obvious treatment manipulation would be to simply compute the best response for them. Remember, though, that we want to allow for the possibility that participants may have heterogeneous preferences. If we want to remain agnostic about their preferences, we cannot compute best responses "for them". In the light of ample evidence that people have difficulty to reduce uncertainty (e.g. Kahneman and Tversky, 1972, Peters et al., 2007), one treatment variation that should make it easier to formulate best responses is to eliminate random variables.[4] One way of doing so is to share the prize between both players according to their individual investment shares in total investment. Hence, instead of receiving $(32 - a_i)$ ECU with probability $\frac{a_i}{a_1+a_2}$ and $(16 - a_i)$ ECU with complementary

[4]In section 2.5.1, we analyze the possibility that this treatment variation interacts with risk preferences to generate differential results.

probability (as in **dD**), subjects received as the monetary round payment the hypothetical expected value:

$$\Pi_i = \frac{a_i}{a_1 + a_2}(32 - a_i) + (1 - \frac{a_i}{a_1 + a_2})(16 - a_i) = 16 - a_i + \frac{16a_i}{a_1 + a_2} \text{ ECU.}$$

In this treatment - given correct beliefs - formulating best responses should be easier, because participants do not have to reduce uncertainty. It is still difficult in this treatment, though, to form correct beliefs.

Treatment dT. Making it simpler (or even trivial) to form correct beliefs requires more elaborate design interventions. If one would like to maintain the simultaneous choice setting (which is desirable because of treatment comparisons), the only way is to let participants play against computers or against human players who are so restricted in their choice of strategy that they could almost be replaced by computers. In the experiment we did both. To keep the structure identical, all treatments (including the benchmark treatment **dD**) were divided into 4 successive blocks and each block was composed of 10 rounds of play.

Specifically, strategic uncertainty was removed by matching participants to computers who played a fixed strategy in every block.[5] Players were informed about the computer's choices ex ante and, hence, should hold deterministic and, moreover, correct beliefs about the opponent's play. The strategy adopted by the computer was held constant within each block of ten rounds and changed from one block to the other, thus, each player faced four different strategies. The first type of players faced a computer that played the sequence of strategies $\{1, 14, 11, 8\}$, the second type faced strategies $\{5, 10, 3, 16\}$ the third $\{9, 6, 15, 4\}$ and the fourth $\{13, 2, 7, 12\}$. Each sequence was allocated at random to the same number of players, hence there is an equal number of observations for each computer choice in $\{1, ..., 16\}$. Sequences were selected such that each participant faced the same average level of computer investment and some "high" as well as "low" investment levels.

Treatment eT. In treatment **eT** both sources of complexity were eliminated using the same procedures as adopted for treatments **dT** and **eD**, respectively. Since players were ex ante informed about the investment choice of the computer, they should hold correct beliefs in **eT** (just as in **dT**). The question is

[5] A similar procedure was used e.g. by Johnson et al. (2002), who match players to "robots" in three-round bargaining game. Robots were playing according to the subgame-perfect Nash equilibrium prediction and induced human players to choose strategies that are closer to the theoretical prediction.

whether - given correct beliefs - participants engage in best response behaviour. This should be easiest in **eT** where *both* sources of complexity are eliminated.

Treatment eT-*h*. Treatment **eT**-h coincides with treatment **eT** except that computers were replaced by human players. At the start of each ten-round block these human players were restricted to choosing strategies that prescribe the same action in each round of the block. Each player was randomly assigned a set from which this action could be chosen, where the four possible sets were $\{1, 14, 11, 8\}, \{5, 10, 3, 16\}, \{9, 6, 15, 4\}$ or $\{13, 2, 7, 12\}$. Each action had to be chosen at least once, therefore players were essentially choosing the order in which strategies are played, while in **eT** the computer played them in a fixed sequence not known to the human participants. In **eT**-h choices have payoff consequences to the other participant, contrary to the **eT** treatment, therefore treatment **eT**-h serves as a robustness check to understand whether replacing human subjects by computers affects our treatment comparison.

Other Details. To improve the understanding of the design and incentives, participants were facing an alternating pattern of 5 non-incentivized and 5 incentivized rounds. After all 40 rounds were completed, players took an incentivized numeracy test that measured the ability to understand and manipulate probabilities.[6] At the end of the experiment, one incentivized round was randomly chosen from each block and participants received the earnings from the chosen rounds as well as the payment for correct answers in the questionnaire. The experiment was conducted in March and September 2012 using z-Tree (Fischbacher, 2007) and ORSEE (Greiner, 2015) at the BEElab at Maastricht University. A total of 244 students participated in the experiment. The average duration of the experiment was 60 minutes and participants on average earned 15.15 euros.

2.3 Conjectures

Our first conjecture is that behavioral variation is lower when complexity is reduced by making it easier to formulate best responses and form correct beliefs.

Conjecture 1. *Behavioural variation is highest in **dD** and lowest in **eT**.*

[6]Questions for the numeracy test were taken from Peters et al. (2007). From the 15 questions in the original test we removed 7 questions that were found too simple (correctly answered by at least 80% of the population with higher education). We did so to increase incentives of answering the more complicated questions and to enable better differentiation based on performance. Subject were paid 1 ECU (0.25 euro) for a correct answer for each of the first 7 questions, and 2 ECU (=0.5 euro) for a correct answer to the final question that was the most difficult one.

We introduce our measure of behavioural variation further below (section 2.4). Before we do that, we formulate conjectures about the explanatory power of NE in our four treatments. Since we induce non-Nash choices by computerized players in the treatments with trivial belief formation, **dT** and **eT**, the percentage of choices which are Nash is not a reasonable measure. Instead we decompose Nash behaviour into the two components that motivated our design: (i) correct beliefs and (ii) best response behaviour. In the following we explain how our treatment variations impact each of these components.

We start by looking at the effect of making it easier to formulate best response. There is a large body of evidence that people have difficulty in comprehending probabilistic statements and in making decisions in probabilistic environments (Kahneman and Tversky, 1972). If NE has explanatory power, then we would expect that removing this uncertainty facilitates best response behaviour. As a consequence, we expect a larger share of choices to be consistent with best response behaviour to *some* beliefs when best response formulation is easy, i.e. in **eD** compared to **dD** and in **eT** compared to **dT**. What does it mean to be "consistent with best response behaviour"? It means that choices are rationalizable, i.e. that there *exists* a belief that would justify the player's choice. Since dominated choices are *not* consistent with best response behaviour, we expect to see less of these choices when best response formulation is easy, as in **eD** compared to **dD** and in **eT** compared to **dT**. Of course, what constitutes a dominated strategy depends on the agent's preferences. We follow the vast majority of the literature and assume a risk neutral expected utility maximizer as a benchmark, but we carefully look at risk and social preferences in section 2.5. There, we demonstrate that our results obtain for very large classes of preferences.

Conjecture 2. *A larger share of choices is undominated with easier best response formulation, i.e. in **eD** compared to **dD** and in **eT** compared to **dT**.*

Furthermore, if the difficulty in dealing with probabilistic statements is the main reason participants make suboptimal decisions in non-deterministic environments, then participants that are better at understanding probabilistic environments should do "better". In other words subjects with higher scores in the numeracy test should choose strategies consistent with best response behaviour more often than others (in treatments **eD** and **eT**). We address this issue in Section 2.4.3, where we present evidence from our questionnaire.

Since belief formation is trivial in treatments **dT** and **eT**, participants should have "correct" beliefs in these treatments. As a consequence, if NE has explanatory power, then we expect that the best response to the opponent's behaviour

is chosen more often when belief formation is trivial, that is in **dT** compared to **dD** and in **eT** compared to **eD**.

Conjecture 3. *A larger share of choices is a best response to the opponent's behaviour with trivial belief formation, i.e. in **dT** compared to **dD** and in **eT** compared to **eD**.*

2.4 Results: Reducing Complexity

In this section, we present our main results. We start by evaluating Conjecture 1 on behavioural variation and compare it across treatments (section 2.4.1). Subsequently, we focus on the explanatory power of Nash equilibrium regarding its two components, best response behaviour and correct beliefs, and evaluate Conjectures 2 and 3 (section 2.4.2). Unless explicitly stated otherwise, throughout the chapter we analyze data generated in the incentivized rounds during the second half of the experiment. The reason is that we want to focus on mature behaviour and eliminate behavioural variation that would disappear after some learning has occurred.

2.4.1 Behavioural Variation

Figure 2.2 is the analogue of Figure 2.1 and compares treatments **dD** and **eT** in terms of behavioural variation. The figure plots the empirical cumulative distribution functions (CDF) of choices a_i divided by the equilibrium prediction along with the theoretically predicted CDF. Theoretically, hence, all the observations should yield a degenerate CDF where all the CDF mass cumulates at one, because choices should equal the equilibrium prediction in each period. What is the equilibrium prediction for both treatments? As discussed before, in treatment **dD** (i) the risk neutral Nash equilibrium prescribes an investment of $a_i^* = 4$ ECU and (ii) the best response of a risk neutral agent to the empirical distribution of choices in **dD** is also $a_i^* = 4$. In treatment **eT**, however, the theoretical prediction is different. Since the computer's investment level is known before making a choice, the prediction is simply the best response to that predetermined investment level. The resulting best responses range from 1 to 4 depending on the computer's choice, noting that all choices exceeding 4 are strictly dominated.

Figure 2.2 shows that our benchmark treatment **dD** produces results for the Tullock contest that are "standard" in the sense of high behavioural variation and choices completely disconnected from the theoretical (Nash) prediction.

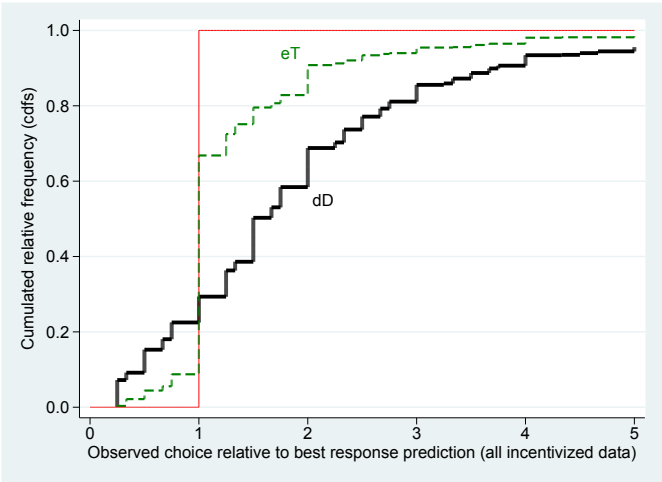


Figure 2.2: Cumulative distributions of ratio variable, where investment choices are divided by equilibrium prediction, in theory and treatments **dD** as well as **eT**.

There does not seem to be more CDF mass at 1, where the actual choice equals the theory prediction, than elsewhere. Note also that all the CDF mass to the right of 1 stems from dominated choices for a risk neutral and risk averse (see section 2.5.1) agents. In treatment **eT**, on the other hand, the theoretical prediction clearly has something to say about the data with most CDF mass concentrated at 1.

We use two measures previously used in the literature to obtain more conventional measures of behavioural variation. The first measure is the standard deviation of choices. The second measure is entropy, which evaluates the stochastic variation of a random variable that can assume a finite set of values (Shannon, 1948). Entropy has been used, e.g. by Bednar et al. (2011), to evaluate behavioural variation in normal-form games. While standard deviation is a very common measure, the advantage of using entropy is that it yields a measure of the average unpredictability of a random variable that captures the amount of information needed to describe a distribution. If a is the random variable of investments and $p_i = P(a = a_i)$ is its probability density function for all possible strategies $a_i = 1, 2, \ldots, 16$, entropy is computed as:

$$H = -\sum_{i=1,2,\ldots,16} p_i \log_2(p_i)$$

It is common to use a logarithm to base 2, so that entropy can be interpreted as the total number of bits needed to describe the data. For our strategy space with 16 possible choices, the entropy measure can take values from 0 (if a single strategy is always chosen) to 4 (if all strategies are chosen with equal

frequency).[7]

To compare behavioural variation across treatments, we compute our measures (entropy and standard deviation) by treatment conditional on the strategy chosen by the opponent. Aggregate entropy in each treatment is then computed as the weighted average of the conditional entropy levels and the weights are determined by the frequency with which each of these strategies are chosen in the standard contest (**dD**). Conditioning on opponents' strategies ensures that the behavioural variation is zero in all treatments if all participants always choose best responses to correct beliefs. Without this conditioning, there would be artificial variation in treatments **dT** and **eT**, which is merely due to the computer changing strategies between blocks, which in some cases changes the best response.[8]

Table 2.2: Behavioral variation across treatments. Data from incentivized rounds in the the second half of the experiment.

	dD	**eD**	**dT**	**eT**
Entropy	3.22	2.79	2.45	1.50
Std. dev.	3.28	2.56	3.15	1.16

Both measures are summarized in Table 2.2. Both entropy and standard deviation are highest in the standard contest (**dD**) and lowest when complexity is lowest (**eT**). The two measures disagree on ranking the intermediate treatments, where either best response formulation or belief formation is simplified (but not both): entropy is higher in **eD** than in **dT**, but standard deviation ranks them the other way round. Table 2.14 in Appendix 2.A.6 shows that both within and between subject variability contribute to the high amount of behavioural variability observed in **dD** and that eliminating uncertainty (as in **eT**) reduces both types of variability.

We conduct Wilcoxon ranksum (Mann-Whitney) tests to check for the statistical significance of these differences. To account for the dependency of observations within matching groups we compute entropy (standard deviation) separately for each matching group and then compare the distribution of the results using two-sided ranksum tests. Since in treatments **dT** and **eT** each participant is an independent observation, we can either compare matching group averages with individual levels or form artificial matching groups in the **dD** and **eD** treatments and then compare matching groups. We do both and re-

[7] To accommodate the case that a strategy is never played, we follow the literature and assume that $0 \cdot \log(0) = 0$.

[8] Despite this fact - even if entropy and standard deviation are computed by simply pooling all the observations in a treatment, treatments are ranked similarly. The only exception would be treatment **dT** which in that case exhibits larger behavioural variation than **dD**.

port the lower/higher p-value, whichever is more relevant. For entropy, we find that the difference between **dD** and **eT** is marginally significant even when the comparison is performed on a matching group level ($p < 0.0544$). Hence, reducing both dimensions of complexity significantly reduces behavioural variation. The difference between **eD** and **eT** is also significant ($p < 0.0210$), indicating that when best response formulation is easy, making belief formation trivial significantly reduces behavioural variation.

In terms of the standard deviation we again find significant differences between **dD** and **eT** ($p < 0.0005$) and **dT** and **eT** ($p < 0.0047$). Overall we find clear support for Conjecture 1 and conclude that reducing complexity along the two dimensions in our design does reduce behavioural variation.

Result 1 *Behavioral variation is lowest in **eT**, where best response formulation is easy and belief formation trivial, among all treatments and it is lower in **eD** and **dT**, where one dimension of complexity is reduced, compared to **dD**.*

2.4.2 Explanatory Power of Nash Equilibrium

Now that we have established that making best response formulation easy and belief formation trivial reduces behavioral variation substantially, we examine whether the explanatory power of Nash equilibrium responds accordingly. We evaluate this conjecture in detail in this section. We follow the vast majority of the literature and focus on the benchmark prediction for risk neutral agents. In Sections 2.5.1 and 2.5.2 we relax this assumption and explore different preference specifications, in particular different risk preferences and social preferences.

Table 2.3: Indicators measuring the explanatory power of Nash equilibrium across treatments. Data from the incentivized rounds in the second half of the experiment.

	dD	**eD**	**dT**	**eT**
$P(a = NE)$	7.04%	13.33%	-	-
$P(a = BR)$	7.04%	12.04%	22.50%	65.23%
$P(\lvert a - NE\rvert \leqslant 1)$	25.74%	32.78%	-	-
$P(\lvert a - BR\rvert \leqslant 1)$	26.30%	31.30%	47.95%	83.64%
$P(a > 4)$	60.19%	62.78%	51.36%	16.14%

Table 2.3 shows summary statistics on different measures of the explanatory power of Nash equilibrium. Let us first focus on the benchmark treatment **dD**. In this treatment, we observe only 7 percent of choices equal to the Nash prediction. To get a sense of how little this is, note that a player choosing uniformly

at random between the possible investment levels $a_i \in \{1, ..., 16\}$ would play the Nash choice 6.25 percent of the time. The explanatory power of NE is not much improved when we look at the percentage of choices that are "close" to the Nash choices ($a_i \in \{3, 4, 5\}$). This percentage is barely above 25 percent. Note again that a random player would hit these numbers in 18.75 percent of the cases. Moreover, approximately 60 percent of all choices in **dD** are strictly dominated ($a > 4$). The top-left panel of Figure 2.3 illustrates these results, where the dashed line separates dominated ($a > 4$) from undominated choices ($a \leqslant 4$). Our results in the benchmark treatment **dD** hence display the low explanatory power of NE typically found in Tullock contest data (Abbink et al., 2010, Sheremeta, 2013). We now study how it changes when we reduce the difficulty of (i) best response formulation and (ii) belief formation.

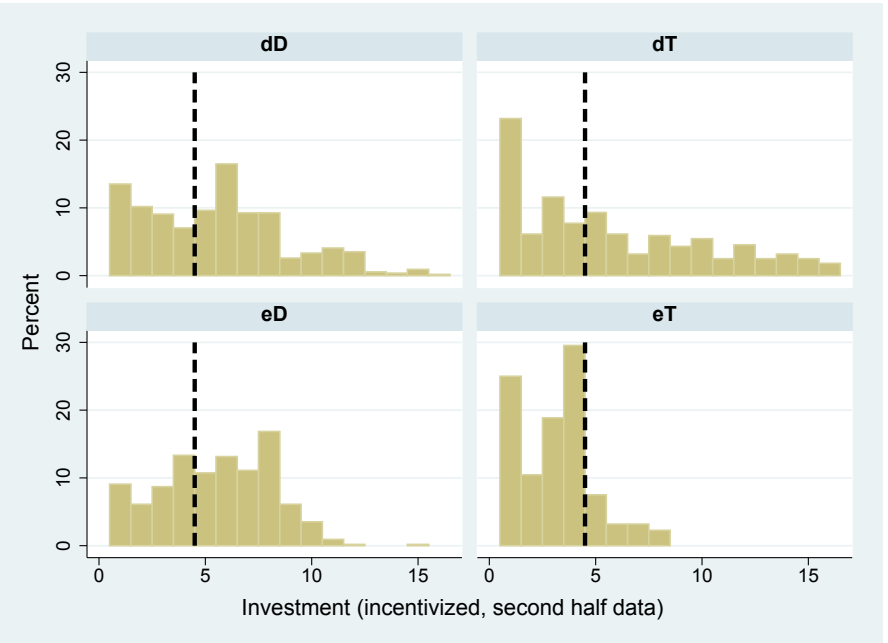


Figure 2.3: Distribution of investment choices by treatment in the incentivized rounds of the second half of the experiment (rounds 26-30 and 36-40). Vertical dashed lines separate undominated ($\leqslant 4$) from dominated choices (> 4).

"Easy best responses" (Conjecture 2) Let us first see whether and how the explanatory power of NE is increased by making it easier to formulate best responses. As outlined in section 2.3, we expect participants to choose strategies that are consistent with best responses to *any belief*, i.e. undominated strategies, more often when best response formulation is easier, i.e. in **eD** compared to **dD** and in **eT** compared to **dT**. For risk neutral (and moderately risk

averse) agents, investment levels below or equal 4 are undominated. Table 2.3 shows the percentage of dominated choices by treatment. If belief formation is difficult, then making best response formulation easy does not improve the explanatory power of NE, as the comparison of treatments **dD** and **eD** shows. In fact the percentage of strictly dominated strategies is even slightly higher under **eD** compared to **dD**, though this difference is not statistically significant (see Table 2.9 in Appendix 2.A.1). Overall, the explanatory power of Nash equilibrium is very low in both treatments. This is also illustrated by Figure 2.3 (top-left vs bottom-left panel).

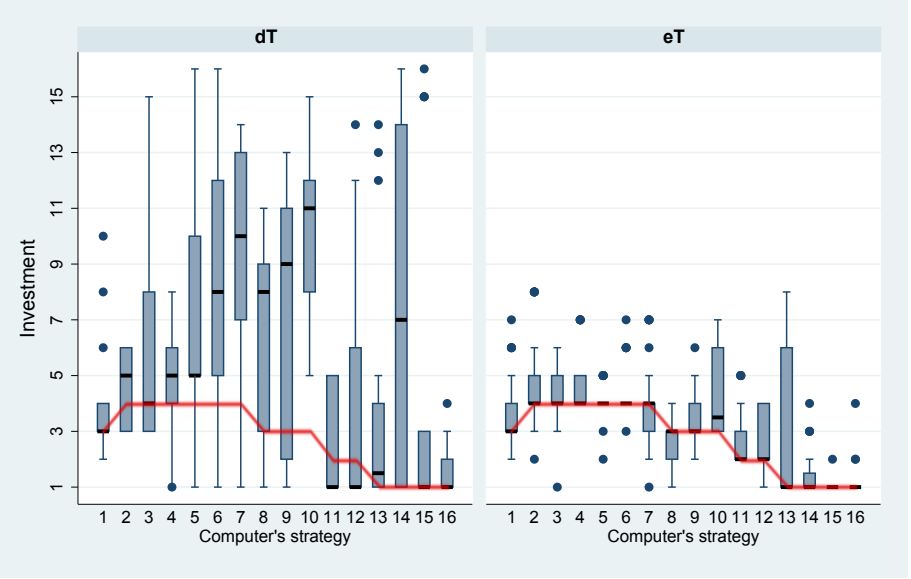


Figure 2.4: Theoretical best response line and the empirical distribution of investments, conditional on the strategy played by the computer. Box plots represent the medians, differences between 25th and 75th percentiles and lower and upper adjacent values.

If belief formation is trivial, on the other hand, behavior in a treatment in which best responding is easier is closer to theoretical predictions. Comparing treatments **dT** and **eT**, we find that easier best response formulation substantially improves the explanatory power (Figure 2.3, top-right vs bottom-right panel). The frequency of strictly dominated choices decreases from 51.36% in **dT** to 16.14% in **eT** (Table 2.3). This difference is highly statistically significant ($p < 0.0001$) according to a simple logit regression reported in Table 2.9 in Appendix 2.A.1. Hence making it easier for participants to formulate best responses reduces the frequency of dominated choices only if it is trivial to form correct beliefs.

Figure 2.4 illustrates this effect in more detail. The figure plots the quartiles

of responses to each of the strategies chosen by the computer. The difference between the first and the third quartiles is much larger in treatment **dT**, especially for large investment values chosen by the computer. In treatment **eT** median investments track the best response function well.[9] Table 2.10 in Appendix 2.A shows that differences between treatments do not disappear with learning. The difference between **dD** and **eD** is insignificant in every block except for the first one where it is marginally significant, while the difference between **dT** and **eT** is highly significant throughout the experiment.

To sum up, there is a complementary effect. Making it easier to formulate best responses reduces the frequency of dominated choices if it is trivial to form correct beliefs.

Result 2 *Making best response formulation easy, increases the explanatory power of NE if belief formation is trivial. Dominated actions are chosen less frequently in* ***eT*** *compared to* ***dT****.*

"Trivial Beliefs" (Conjecture 3) If beliefs are correct, then the explanatory power of Nash equilibrium is reflected in the share of choices which are best responses to the opponent's behaviour. Recall, that in the treatments with trivial belief formation the opponent's behaviour is announced, giving rise to correct beliefs. For comparison to the treatments with difficult belief formation we report the frequency of choices that are best response to actual opponents' behaviour in the latter.[10]

Consistently with Conjecture 3, Table 2.3 shows that the share of best responses increases as belief formation becomes trivial irrespective of whether best response formulation is difficult or easy. While less than 10% of participants best-respond to their opponent's choices in **dD**, this number more than doubles in **dT** and reaches 22.50%. Table 2.9 in Appendix 2.A.1 shows that this treatment difference is significant ($p < 0.0001$). The effect is even stronger if we compare **eD** and **eT**: the share of choices which are best responses to the opponent's behaviour increases from 12% to more than 65%, a difference that is statistically significant ($p < 0.0001$, Table 2.9). 83.64% of choices in **eT** deviate from the best response by 1 ECU or less.[11]

[9] Additional evidence for this effect can be found in Table 2.11 in Appendix 2.A that compares the median investments for each choice of the opponent across treatments **dT** and **eT**. It can be seen that the median investment level tends to be closer to the best response in **eT** and that the treatment difference is particularly large for situations where the opponent chooses a very high investment level (5 and above)

[10] Results are qualitatively the same if we consider best responses to population averages, instead.

[11] **eT** is also the treatment where most learning is observed (see Table 2.10 in Appendix 2.A). The change in average equilibrium deviations decreases by 63% from the first to the fourth block.

Result 3 *Easier belief formation increases the explanatory power of NE: more choices are best responses to opponent's behaviour in **eT** compared to **eD** and in **dT** compared to **dD**.*

It should be noted that the difference between **eT** and **eD** is larger than the difference between **dT** and **dD**, i.e. making belief formation trivial increases the explanatory power of NE more if best response formation is easy ($\chi^2 : p < 0.0023$, see Appendix 2.A.1). Also recall that making best response formation easy is effective only if belief formation is trivial (no significant difference between **eD** and **dD**). Hence both sources of complexity compound to decrease the explanatory power of NE. Removing both sources of complexity leads to very high consistency with best response behavior and a consequently high explanatory power of Nash equilibrium. These results suggest that bounded rationality is a key force behind the typically low explanatory power of NE in Tullock contests.

2.4.3 Cognitive Abilities and Equilibrium Deviations

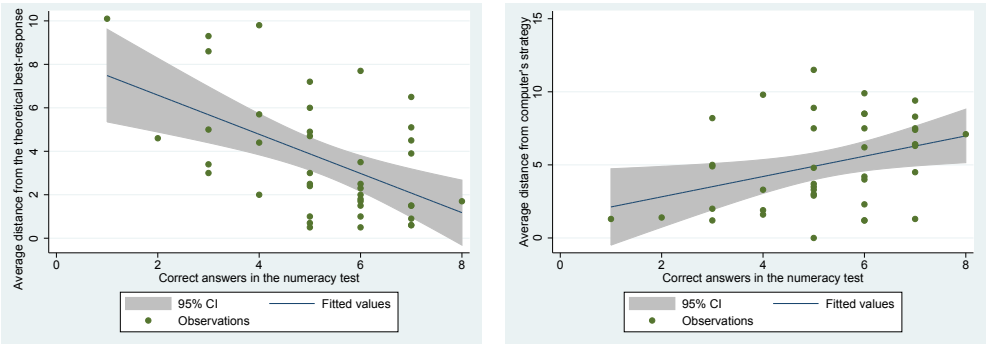
In the previous subsection we have established that while behavioural variation is huge in **dD** and significantly reduced in **eT**, there remains a substantial amount of behavioural variation in **dT**. This shows that even if belief formation is trivial, participants still find it difficult to formulate best responses with probabilistic prize allocation. Since in **eT** (where prize allocation is deterministic) this variation disappears, the difficulty in forming best responses seems to be related to (some) participants' difficulty in dealing with probabilistic choice situations.

In this subsection we evaluate this hypothesis using data from our post-play questionnaire on cognitive ability and risk numeracy.[12] If complexity is indeed one of the main reasons why participants' choices are not consistent with best response behaviour in the experiment, then one might conjecture that those with better scores in a risk numeracy test should have less difficulty in dealing with this complexity and hence should be closer to best response behaviour.

As argued above, we expect risk numeracy to matter particularly in treatment **dT**, where belief formation is trivial, but best response formation difficult. It is not clear why cognitive ability (as measured in numeracy tests) should help to form correct beliefs, i.e. explain variation in **dD** and **eD**. In treatment **eT**, on the other hand, where both best response formation and belief formation are easy we have seen that behavioural variation is low. We focus hence on

[12]The entire questionnaire, as well as summary statistics on answers can be found in Appendix 2.B.3.

dT. Figure 2.5 plots the score in the numeracy test (higher score means higher numeracy) against the average distance between a participant's choice and the theoretical prediction (panel a) and the computer strategy (panel b).



(a) Distance to BR (b) Distance to Computer Strategy

Figure 2.5: Relationship between the performance in a numeracy test and absolute difference between investments and computer's strategy/best response in **dT**, averaged by subjects. Incentivized rounds from the second half of the experiment.

Figure 2.5 shows data points as well as predicted values from an OLS regression and the associated 95 percent confidence interval. It can be seen that participants with higher scores in the numeracy test on average choose investments that are closer to the theoretical best-response. They are also less prone to imitating the computer strategy, as Figure 2.5(b) suggests, but the statistical significance here is somewhat lower. The underlying regression results can be found in table 2.15, in Appendix 2.A.7. Both results provide further support to our conjecture that the difficulty of participants to make decisions in uncertain environments is the main cause for the low explanatory power of Nash equilibrium in the Tullock contest.

Result 4 *In treatment **dT** participants who do better in the numeracy test best respond more often and imitate less often than those who do worse.*

2.5 Preference Based Explanations

In this section we first ask whether our treatment comparisons and conclusions remain valid under alternative assumptions on preferences. We first look at risk preferences (section 2.5.1) and then at social preferences (section 2.5.2). Afterwards, in section 2.5.3 we ask whether heterogeneity in preferences can explain the large behavioural variation observed in the standard contest.

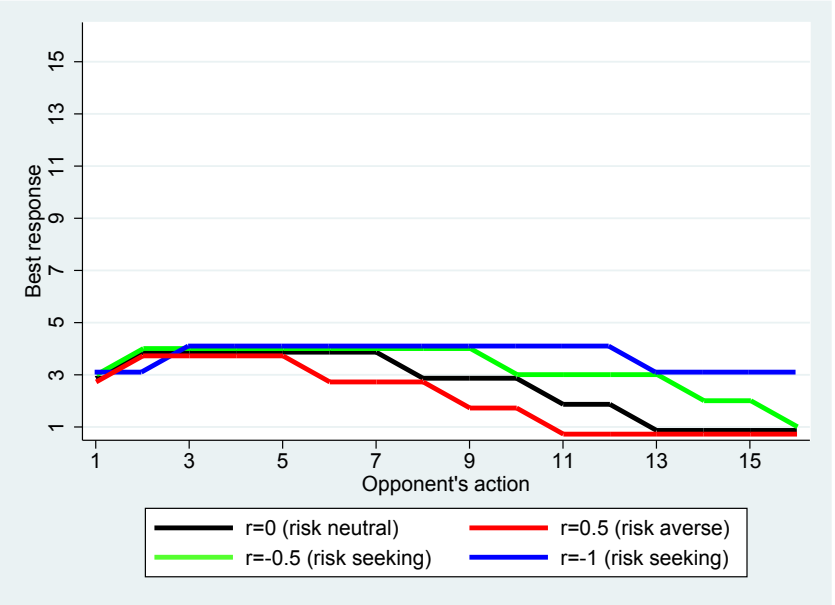


Figure 2.6: Best response correspondence for an agent with CRRA preferences $U(x) = \frac{x^{1-r}}{1-r}$ with parameter $r \in \{-1, -0.5, 0, 0.5\}$.

2.5.1 Risk Preferences

In this subsection we show that our results and interpretation are robust to considering different risk preferences. First, note that results on behavioural variation are independent of our participant's preferences. In this section, we hence focus on our results from section 2.4.2 (on the explanatory power of Nash equilibrium) and show that they are robust when different risk preferences are considered.

The first question to ask is how risk aversion (or risk seeking behaviour) affects the theoretical predictions in the standard contest. Figure 2.6 shows the best response function for a risk seeking (parameter $r = -0.5$), risk neutral ($r = 0$) and risk averse ($r = 0.5$) CRRA agent. The figure shows that the best response to choices below or equal to four are almost identical for these three types. It further shows that all choices above 4 are never a best response for any of these types. In addition, the more risk averse an agent is, the more the best response function shifts downwards, i.e. the lower his best response is. Hence choices above the Nash level of 4 cannot be explained by risk aversion (see also Hillman and Katz, 1984, or Abbink et al., 2010). Table 2.4 shows that the percentage of choices consistent with best response behaviour are much higher in **eT** compared to all other treatments also for risk averse and risk loving agents and they remain very low (barely above uniformly random) in treatment **dD**.

Table 2.4: Percentage of choices that are best responses to the opponent's choice (**dT** and **eT**) or the empirical distribution of choices (**dD** and **eD**) for risk averse (CRRA with $r = 0.5$)/risk neutral/risk seeking (CRRA with $r = -0.5$) players.

Treatment	**dD**	**eD**	**dT**	**eT**
Risk neutrality	7.04%	12.04%	22.50%	65.23%
Risk aversion ($r = 0.5$)	7.41%	10.74%	27.50%	48.64%
Risk seeking ($r = -0.5$)	7.59%	13.52%	11.14%	40.68%

Further, risk aversion and in particular also risk seeking preferences are even less consistent with evidence in **eT**, suggesting that risk aversion did not play a major role in the experiment. It has also been shown that loss aversion or S-shape probability weighting cannot explain such choices (Baharad and Nitzan, 2008). This also means that the results shown in Table 2.3 would remain unchanged if moderately risk averse or risk loving CRRA agents are considered.

A different question is whether, in the presence of risk averse agents, our treatment comparisons are distorted by making prize allocation deterministic, as this reduces risk for our participants. In particular a risk averse CRRA agent might respond to the risk associated with probabilistic prize allocation in treatments **dD** and **dT**, but behave as if he was risk neutral in **eT**, because all risk has been eliminated there. The preceding discussion suggests that - given moderate degrees of risk aversion - this should not make too much of a difference. In particular, Figure 2.6 shows that under CRRA utility risk averse participants invest *more* than risk neutral participants if the opponent invests very little and *less* than risk neutral participants if the opponent invests a lot. Since in treatments **dT** and **eT**, participants know their opponents' investments, we can use this to test for such distortions. If risk aversion plays a significant role in explaining the differences between these treatments, then we should expect that investments are higher in **dT** compared to **eT** for low investment choices of the opponent and lower for high investment choices of the opponent. Figure 2.4 shows that this is not the case. Mostly behavioural variation is higher under **dT** compared to **eT**, but if at all investments in **dT** are lower for low investments and higher for high investments of the opponent compared to **eT**. Hence we can rule out this type of distortion of our treatment comparisons.

We conclude that the results summarized in Table 2.3 as well as treatment rankings are robust to considering moderate degrees of risk aversion or risk seeking behaviour.

2.5.2 Social Preferences

Similarly, we ask whether our conclusions drawn so far, and in particular our treatment comparisons, are valid if our participants had social preferences, such as inequity aversion, joy of winning or reciprocity. For example, some participants might have preferences over the outcomes of other participants in treatments **dD** and **eD**, but not in treatments **dT** and **eT**, where choices have no payoff consequences for other participants. To address this potential concern we ran an additional treatment (**eT**-h) that coincides with **eT**, but where computers are replaced by human opponents. We were interested in seeing whether this change reverses any of our treatment rankings and therefore affects any of the conclusions we drew.

Table 2.5: Behavioral variation across treatments. Data from incentivized rounds in the the second half of the experiment.

	dD	**eD**	**dT**	**eT**-h
Entropy	3.22	2.79	2.45	1.13
sd	3.28	2.56	3.15	0.91

Table 2.5 corresponds to Table 2.2, but treatment **eT** has been replaced by treatment **eT**-h. It can be seen that the conclusions are the same. Irrespective of the measure (entropy or standard deviation) behavioural variation is much lower in **eT**-h compared to any of the other treatments.

Table 2.6: Indicators measuring the explanatory power of Nash equilibrium across treatments. Data from the incentivized rounds in the the second half of the experiment.

	dD	**eD**	**dT**	**eT-h**
$P(x = NE)$	7.04%	13.33%	-	-
$P(x = BR)$	7.04%	12.04%	22.50%	50.42%
$P(\lvert x - NE\rvert \leqslant 1)$	25.74%	32.78%	-	-
$P(\lvert x - BR\rvert \leqslant 1)$	26.30%	31.30%	47.95%	74.58%
$P(x > 4)$	60.19%	62.78%	51.36%	23.33%

Table 2.6 corresponds to Table 2.3, but again treatment **eT** has been replaced by treatment **eT**-h. Again our conclusions are robust. The share of choices corresponding to best response behaviour or being "close" to best response behaviour is much larger in **eT**-h compared to any of the other treatments. Hence, while some participants may well have social preferences we conclude that (i) homogeneous social preferences cannot explain the large (within subject) variation in behaviour we observe and (ii) our treatment rankings and conclusions are not affected by allowing for this possibility. In the next subsection we ask whether *heterogeneity* in social or other preferences can explain the large behavioural variation we observe.

2.5.3 Heterogeneity

In this subsection we report the results of simulations based on a number of different populations with heterogeneous preferences. Since in our experiment (treatments **dD** and **eD**) we had matching groups of 6 players, we form arbitrary populations consisting of 6 players with different preferences. The preference types we consider are the following

- τ_1: risk neutral expected utility maximizers.
- τ_2: risk averse CRRA agents with CRRA parameter $r = 0.5$ (see Appendix 2.A)
- τ_3: risk seeking CRRA agents with CRRA parameter $r = -0.5$ (see Appendix 2.A)
- τ_4: social preferences: Charness-Rabin $(\rho, \sigma) = (0.4, 0)$ (Charness and Rabin, 2002)
- τ_5: social preferences: Charness-Rabin $(\rho, \sigma) = (0.8, 0.1)$ (Charness and Rabin, 2002)
- τ_6: joy of winning preferences: additional utility of 8 for winning (Sheremeta, 2013)

Types $\tau_1 - \tau_3$ vary risk preferences. Types $\tau_4 - \tau_5$ vary social preferences as in Charness and Rabin (2002). Type τ_4 reflects the parameters estimated from their experiments and type τ_5 has the same parameters but multiplied by two. Type τ_6 receives an additional utility of 8 if s/he wins the contest. These type of "joy of winning" preferences have received a lot of attention in the contest literature to explain the so-called "overbidding phenomenon" (see Sheremeta, 2013). An additional utility of 8 reflects a 50% increase in the perceived value of the prize.

We simulated different population compositions of these types for 40 periods. Before the first period we randomly matched agents into pairs. In period 1 actions were drawn at random from $\{1, ..., 16\}$. In all subsequent periods agents played a myopic best response to the action of their previous match. To compute the myopic best response for the Charness-Rabin types τ_4 and τ_5 we assume that the prize is shared proportionally as in treatment **eD**. All simulations were run 100 times to account for path dependence caused by random matching and initial actions. We report the minimal, maximal, mean and median entropy across these 100 runs.

Table 2.7: Minimal, maximal, median and mean entropy across 100 runs of simulated populations. Entropy in treatments **dD** and **eD** was calculated separately for each matching group using incentivized rounds in second half of experiment.

composition $(\tau_1, ..., \tau_6)$	min	max	median	mean
$(6, 0, 0, 0, 0, 0)$	0.217	0.530	0.391	0.379
$(0, 6, 0, 0, 0, 0)$	0.255	0.538	0.430	0.425
$(0, 0, 6, 0, 0, 0)$	0.147	0.429	0.321	0.315
$(0, 0, 0, 6, 0, 0)$	0.806	2.104	1.624	1.656
$(0, 0, 0, 0, 6, 0)$	0.754	2.076	1.613	1.584
$(0, 0, 0, 0, 0, 6)$	0.213	0.426	0.333	0.328
$(0, 2, 0, 2, 2, 0)$	0.320	0.974	0.710	0.694
$(1, 1, 1, 1, 1, 1)$	1.158	1.649	1.415	1.409
$(1, 2, 0, 1, 1, 1)$	1.339	1.695	1.550	1.543
$(1, 2, 0, 2, 1, 0)$	0.327	0.541	0.549	0.758
$(2, 0, 0, 2, 1, 1)$	1.099	1.587	1.330	1.344
$(2, 0, 0, 2, 2, 0)$	0.406	0.952	0.690	0.683
$(2, 1, 1, 1, 1, 0)$	0.252	0.628	0.444	0.447
$(2, 2, 0, 1, 1, 0)$	0.262	0.638	0.464	0.464
$(3, 0, 0, 2, 1, 0)$	0.247	0.790	0.547	0.536
$(3, 0, 0, 3, 0, 0)$	0.192	0.801	0.522	0.522
$(3, 1, 0, 1, 1, 0)$	0.264	0.652	0.473	0.466
$(3, 2, 0, 1, 0, 0)$	0.255	0.597	0.425	0.423
$(3, 3, 0, 0, 0, 0)$	0.186	0.525	0.412	0.405
dD	2.284	3.441	2.860	2.941
eD	2.014	3.165	2.568	2.568

Table 2.7 shows the results of our simulations. The highest levels of entropy are reached in homogeneous populations of agents with Charness-Rabin preferences because such preferences lead to multiple equilibria and therefore matching groups converge to different equilibria. However, even in these populations the average levels of entropy observed (1.656 or 1.584, respectively) fall well short of the high levels of entropy observed in treatments **dD** and **eD**. We conclude that heterogeneity in preferences cannot explain the high behavioural variation observed in treatments **dD** and **eD**.

2.6 Strategic Uncertainty and Stability

In treatments **dT** and **eT** players face no strategic uncertainty in the current round, but they also know that the computer will keep playing the same action in all remaining rounds of a block. This stability of opponent's choices provides players with additional learning opportunities that may further decrease the gap between the theoretical predictions and choices. To disentangle the effects of strategic uncertainty and stability of choices we ran two additional treatments in which players faced no strategic uncertainty in the current round, but did face uncertainty about opponent's future choices. We denote these two

treatments by **dT**-*var* and **eT**-*var* ("var" stands for "variable"). In these treatments the computer was programmed to play the history of choices observed by participants in **dD**, therefore each participant in **dT**-*var* and **eT**-*var* had a counterpart in **dD** who faced the same sequence of choices but was not informed about opponent's choice in the current round.[13]

Table 2.8: Indicators measuring the explanatory power of Nash equilibrium across treatments. Data from the incentivized rounds in the second half of the experiment.

	dD	**eD**	**dT**-*var*	**eT**-*var*	**dT**	**eT**
$P(a = NE)$	7.04%	13.33%	-	-	-	-
$P(a = BR)$	7.04%	12.04%	7.59%	25.37%	22.50%	65.23%
$P(\lvert a - NE\rvert \leqslant 1)$	25.74%	32.78%	-	-	-	-
$P(\lvert a - BR\rvert \leqslant 1)$	26.30%	31.30%	25.00%	51.85%	47.95%	83.64%
$P(a > 4)$	60.19%	62.78%	62.96%	47.04%	51.36%	16.14%

The explanatory power of NE is displayed in Table 2.8, which reproduces Table 2.3 but also includes the two additional treatments. Additional materials on behavioral variation and best response behavior have been placed in Appendix 2.A.5 (Table 2.13, Figure 2.8 and Figure 2.9). We find that if best response formulation is difficult, removing strategic uncertainty without providing additional opportunities to learn does not increase the explanatory power of a Nash equilibrium: the tendency to best-respond or choose dominated strategies is almost identical in treatments **dT**-*var* and **dD** (Table 2.8), while behavioral variation is even slightly higher in **dT**-*var* (Table 2.13). The increase in explanatory power of NE is therefore almost entirely due to the additional opportunities to learn. The difference between top-left and bottom-left panels in Figure 2.8 further shows that stability is particularly helpful in finding optimal responses to small investment levels.

When best response formulation is easy, both lower strategic uncertainty and higher stability of choices increase the explanatory power of theoretical predictions. Treatment **eT**-*var* has a higher explanatory power than **eD** (Table 2.8), but behavioral variation is almost identical (Table 2.13) and even when strategic uncertainty is not present dominated strategies are still chosen almost 50% of the time. These results are consistent with the finding that the theoretical best response is rarely chosen in sequential contests (see Fonseca, 2009). The explanatory power is sufficiently high only if players have additional opportunities to learn, as can be seen from the difference between **eT**-*var* and **eD** in terms of the explanatory power (Table 2.8) and behavioral variation (Table 2.13). Figure 2.8 demonstrates that the distribution of responses is more con-

[13] 54 participants took part in each of these treatments and all participants were informed that computer choices were determined by the choices of participants in a previous experiment.

centrated and the median response is closer to the theoretical prediction when player have additional opportunities to learn.

Finally, the comparison between **eT**-*var* and **dT**-*var* shows that making belief formulation easier increases the explanatory power of NE if belief formation is trivial (Result 2) even when players have no additional learning opportunities. In **eT**-*var* the theoretical best response is chosen more frequently and dominated strategies are chosen less frequently (Table 2.8) and behavioral variation is lower (Table 2.13 and Figure 2.8).

2.7 Conclusion

We conducted an experiment to understand the reasons behind the large behavioural variation and the low explanatory power of Nash equilibrium typically found in Tullock contests. Across treatments we vary the difficulty of (i) forming correct beliefs and (ii) formulating best responses. In the treatment where both belief formation and best response formulation are "easy", behavioural variation is substantially lower and the explanatory power of Nash equilibrium much higher. Via additional treatments and several simulations we show that heterogeneity in preferences cannot explain the large behavioural variation found in the standard contest. We conclude that bounded rationality rather than preference heterogeneity is the reason for the typically large behavioural variation in experimental Tullock contests.

Appendix 2

2.A Additional results

2.A.1 Regression on the Explanatory Power of Nash Equilibrium

Table 2.9 shows the results of logit regressions where the dependent variable equals 1 if players make an undominated choice (in columns 1 and 2) or if players best-respond to the opponent's choice (in columns 3 and 4) and the independent variables are the treatment dummies. Results illustrate the statistical significance of the differences seen in Table 2.3. A χ^2 test shows that the difference between **dD** and **dT** is smaller than the difference between **eD** and **eT** ($p < 0.0023$).

Table 2.9: Logit regressions for the treatment effects. The first column compares the frequency of undominated choices in treatments **dD** and **eD**, the second column compares **dT** to **eT**. The third column compares the share of choices that are best responses to the opponent's behaviour in **dD** and **dT** and the fourth column compares **eD** to **eT**. Standard errors clustered at matching group level are reported in parentheses.

	$\Pr(a \leqslant 5)$	$\Pr(a \leqslant 5)$	$\Pr(a \in BR(a\_i))$	$\Pr(a \in BR(a\_i))$
Easy best response	−0.109	1.702***	No	Yes
	(0.392)	(0.328)		
Trivial belief formation	No	Yes	1.344***	2.617***
			(0.234)	(0.416)
Constant	-0.413**	-0.054	-2.581***	-1.988***
	(0.216)	(0.198)	(0.130)	(0.355)
Observations	1080	880	980	980
Baseline	**dD**	**dT**	**dD**	**eD**

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.A.2 Deviations from Theoretical Predictions across Blocks

Table 2.10 shows how much choices differ from theoretical best responses in each block, how big the effect of removing *either* source of uncertainty is and whether these effects are statistically significant in each block. The statistical comparison between **dD** and **eD** treatments was done using matching groups as independent observations, while the comparison between **dT** and **eT** was done using individual players as independent observations. If individual players are used for the first comparison instead, results do not change and p-values are above 0.1 in every block. Student's t test yields the similar results, except that the difference between means of treatments **dD** and **eD** is marginally significant in blocks 2, 3 and 4, with p-values respectively 0.0633, 0.0415 and 0.0186.

Table 2.10: Comparison of mean equilibrium deviations across treatments in every block. Two-tailed Mann-Whitney U test p-values of pairwise comparisons are displayed in the right panel

Treatment	**dD**	**eD**	**dT**	**eT**
First block	3.77	3.04	4.45	1.59
Second block	3.04	2.60	4.41	1.15
Third block	2.86	2.53	3.69	0.74
Fourth block	2.90	2.43	3.23	0.60
Total	3.14	2.65	3.95	1.02

Comparison	**dD** vs **eD**	**dT** vs **eT**
First block	0.0094	0.0000
Second block	0.3008	0.0000
Third block	0.1050	0.0000
Fourth block	0.1224	0.0000
Total	0.0014	0.0000

Table 2.11 shows the median investment for each strategy played by the computer in treatments **dT** and **eT**. In treatment **dT** strong deviations from best response behaviour are not due to outliers but are a systematic pattern. The table also shows that in **eT**median choices are largely consistent with best response behaviour.

Table 2.11: Theoretical best responses and the median actual investments for each strategy played by the computer.

Strategy	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Theoretical BR	3	4	4	4	4	4	4	3	3	3	2	2	1	1	1	1
eT (last 20)	3	4	4	4	4	4	4	3	3	3.5	2	2	1	1	1	1
dT (last 20)	3	5	4	5	5	8	10	8	9	11	1	1	1.5	7	1	1

2.A.3 Determinants of Behaviour

Table 2.12 shows the results of a regression looking at what factors induce a player to deviate from the theoretical prediction. Among the explanatory variables we include personal characteristics: the total number of correct answers in a numeracy test and gender as well as variables on the history of play: opponent's investment in the previous period (treatments **dD** and **eD**), strategy chosen by the computer (treatments **dT** and **eT**), a binary variable indicating whether the agent won in the previous period and the inverse of the period variable.

Results of the regression are presented in Table 2.12. Players who did better in the numeracy test on average behave more in line with the theoretical prediction in treatment **dT** (see section 2.4.3). No significant gender effect is observed. In all treatments behaviour tends to move towards the equilibrium over time, as seen from the positive coefficient on the inverse period variable, but the learning effect is strongest in treatment **eT**.

Table 2.12: Random effects estimation, using data from all incentivized rounds. The dependent variable is the absolute difference between the best response prediction and the investment level. Standard errors are clustered on subject level.

	dD	eD	dT	eT
Numeracy score	-0.142	-0.133	-0.648***	-0.214*
Female	-0.128	-0.582	0.0426	0.112
1/period	6.370*	5.777**	6.800	7.808***
Won in the previous round	-0.103		0.973***	
Opponent's action in the previous round	0.0598*	0.0584**		
Computer's investment			0.331***	-0.0136
Constant	3.217**	3.021***	3.815***	1.754***
Observations	980	1000	820	820

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.A.4 Convergence and Dynamics

Figure 2.7 shows the frequency with which participants switch actions over time in the non-incentivized and incentivized periods. In treatments **dD** and **eD** there seems to be little learning over time. Participants switch their choices across periods with a probability of around 0.6. This number barely declines over time (with the exception of the first 5 periods, where it drops from 0.8 to 0.6) and there is no discernible difference between incentivized and non-incentivized periods. There seems to be almost no learning in these treatments and participants remain unsure about what to choose until the end of the 40 periods.

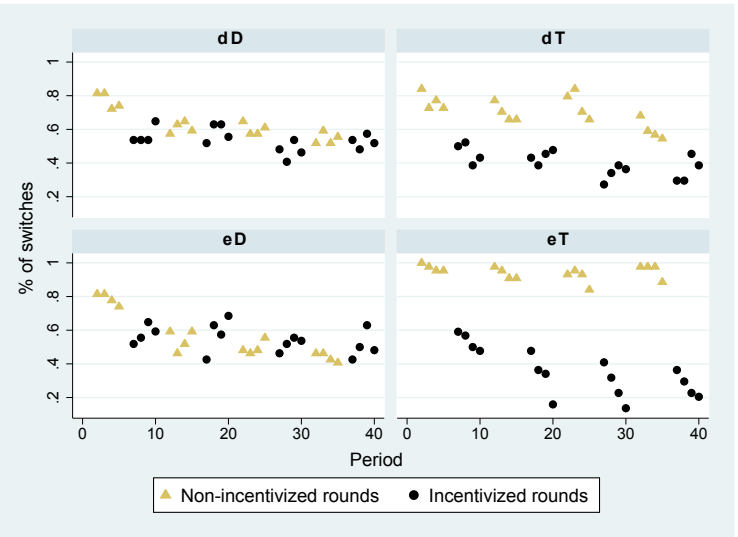


Figure 2.7: Relative frequency of changes in investments from one round to the other, by treatment.

The picture looks much different in the treatments **dT** and **eT**. In the incentivized rounds, participants on average switch their choice with a probability of 0.4 (**dT**) or 0.3 (**eT**). Switching probabilities decrease over time, both within and across blocks and there is a clear difference between incentivized and non-incentivized rounds. Participants seem to be using non-incentivized rounds to experiment and learn about best responses and then apply these in incentivized rounds. The difference between treatments with difficult and trivial belief formation is highly statistically significant (one-sided t-test $p < 0.0001$), so is the difference between **dT** and **eT** (one-sided t-test on incentivized rounds only $p = 0.0392$.

2.A.5 Treatments with no Strategic Uncertainty and no Additional Learning Opportunities

Figure 2.8: Theoretical best response and empirical distribution of investments, conditional on the strategy played by the computer. Panels at the top reproduce figure 2.4.

Table 2.13: Behavioural variation across treatments, using data from incentivized rounds in the the second half of the experiment. Behavioural variation differs from that in table 2.2 because here it is calculated on pooled data instead of conditioning on opponent's actions.

	dD	**eD**	**dT**-*var*	**eT**-*var*
Entropy	3.51	3.28	3.55	3.25
Std. dev.	3.30	2.61	3.84	2.58

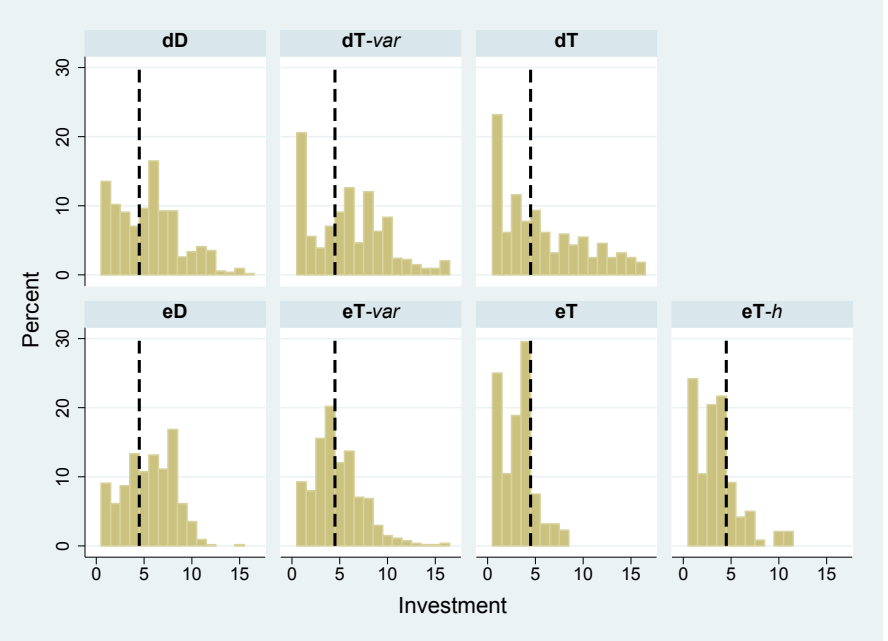


Figure 2.9: Distribution of investment choices from all treatments in the incentivized rounds of the second half of the experiment (rounds 26-30 and 36-40). Vertical dashed lines separate undominated ($\leqslant 4$) from dominated choices (> 4).

2.A.6 Within-Subject and Between-Subject Variability

Table 2.14: Within-subject and between-subject variability by treatment. Variables: mean squares (MSQ) and sum of squares (SSQ) within subjects and between subjects. Variability is computed conditional on opponent's investment and then averaged using the frequency of each strategy chosen in the standard contest as weights. Data from incentivized rounds in the second half of the experiment.

Treatment	MSQ Within	MSQ Between	SSQ Within	SSQ Between
dD	3.96	17.34	1.65	9.17
eD	2.27	11.16	1.14	5.55
dT	2.74	49.15	2.28	8.28
eT	0.36	7.71	0.30	1.25

Table 2.14 decomposes the behavioural variation (standard deviation) into within and between subject variation. The table reports both sum of squares and mean of squares, each decomposed into between and within subject variation. The table shows that both within-subject and between-subject variability is lower in **eD** compared to **dD** and lower in **eT** compared to **dT**. Hence making it easier to formulate best responses lowers both within and between subject variability irrespective of whether belief formation is difficult or trivial. Furthermore, within subject variation is lower in **eT** compared to **eD**, but the

difference between **dT** and **dD** is not significant. Overall, **eT** has the lowest level of both types of variability whereas **dD** and **dT** are most variable.

2.A.7 Cognitive Abilities and Equilibrium Deviations

Table 2.15: Random effects estimation in **dT** treatment. In the first model the dependent variable is the absolute difference between the strategy used by the computer and the investment level, in the second model the dependent variable is the absolute difference between the theoretical best-response and investment level. Standard errors clustered on subject level. Data from all incentivized rounds.

	Distance to computer strategy	Distance to BR
Numeracy score	0.607*	-0.843***
	(2.56)	(-3.71)
Age	0.318	-0.191
	(1.47)	(-1.31)
Female	0.683	-0.457
	(0.75)	(-0.67)
Constant	-5.253	12.38***
	(-1.16)	(3.65)
Observations	410	410

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.A.8 Summary Statistics

Table 2.16: Summary statistics from the questionnaire (mean and range). Numeracy score counts the number of correct answers in the numeracy test.

	dD	**eD**	**dT**	**eT**
Age	21.86	21.57	21.70	21.85
	(18,41)	(18,26)	(18,26)	(19,30)
Female	0.36	0.53	0.45	0.63
	(0,1)	(0,1)	(0,1)	(0,1)
Numeracy score	5.20	5.22	5.11	5.27
	(1,7)	(0,8)	(1,8)	(0,8)

2.B Instructions

2.B.1 Instruction for Treatment dD

INSTRUCTIONS

Welcome to the experiment. Please read the instructions carefully. They are identical for all the participants with whom you will interact during this experiment.

If you have any questions please raise your hand. One of the experimenters will come to you and answer your questions. From now on communication with other participants is not allowed. If you do not conform to these rules you will be excluded from the experiment with no payment. Please do also switch off your mobile phone at this moment.

In this experiment you can earn some money. How much you earn depends on your decisions and the decisions of the other participants. During the experiment we will refer to ECU (Experimental Currency Unit) instead of Euro. The total amount of ECU that you will have earned during the experiment will be converted into Euro at the end of the experiment and paid to you in cash confidentially. The conversion rate that will be used to convert your ECU earnings into your Euro cash payment will be shown to you on the screen at the beginning of the experiment.

The Experiment

The main part of this experiment consists of 4 blocks with 10 rounds in each block. In each block, the first five rounds are for practice only so that you can experiment without affecting your cash earnings from this experiment. At the end of the experiment, one round out of the final five rounds in each block (that is one round of rounds 6-10 in each block) will be randomly selected and the sum of your round incomes in these selected rounds will be converted into euros and paid to you in cash.

The Task

The task is the same in each of the 40 rounds. At the beginning of each round the computer will randomly match you with another participant in this room. You will not know who the other participant is, and the other participant you are matched with is likely to change every round. The other participant that is matched with you will receive the same information and will face exactly the same task. In each round each participant will receive an endowment of 16 ECU. The endowment can be used to purchase “tokens”. Each token costs 1

ECU so that you can purchase up to 16 of these tokens. You have to buy at least one token. Any part of your endowment that you do not spend on tokens will be added to your round income.

After you and the other participant have chosen how many tokens to buy, only one of you will receive the extra 16 ECU. The probability that you will receive the extra 16 ECU depends on the number of tokens that you buy and the number of tokens that the other participant buys. More precisely, the probability that you receive the extra 16 ECU is given by:

$$\text{Probability of receiving the extra 16 ECU} = \frac{\text{Number of tokens you bought}}{\text{Number of tokens you bought} + \text{Number of tokens the other participant bought}} \times 100\%$$

For example, if both of you have purchased the same number of tokens, the probability that each of you will receive the extra 16 ECU is 50%. Note that either you or the other participant will always receive the extra 16 ECU.

Whether you or the other participant will receive the extra 16 ECU will be determined by a random draw by the computer according to the probabilities given by the number of tokens bought by you and by the other participant. Then the computer will compute your round income based on the number of tokens that you bought and whether you have received the extra 16 ECU or not. Once the round is over, you will be informed about the number of tokens bought by you and by the other participant, the probability to receive the extra 16 ECU, the outcome of who receives the extra 16 ECU and your round income, and that of the other participant.

- If you receive the extra 16 ECU, your round income will be:

  Round Income = 16 ECU − Number of purchased tokens + 16 ECU

- If you do not receive the extra 16 ECU, your total earnings in the round will be:

  Round Income = 16 ECU − Number of purchased tokens

At the end of the experiment, four rounds will be randomly selected for payment. More precisely, the first round for payment will be randomly selected from rounds 6-10, the second round for payment from rounds 16-20, the third from rounds 26-30 and the fourth one from rounds 36-40. Outcomes in all other rounds will not influence your final earnings, but you will not know which rounds will be selected until the end of the experiment.

After the 40 rounds we will ask you to fill in a short questionnaire. The questionnaire will have two parts, each of which will be explained on the screen before you start answering the questions. After the first part of the questionnaire you will be informed about all of your round incomes as well as about the four rounds that were randomly selected for payment (see Figure 1 on the next page). After the second part of the questionnaire you will be informed about your final earnings in euro. You will receive these earnings in cash and in private at the end of the experiment. Please stay seated until we ask you to come to receive the earnings.

If you have any further questions, please raise your hand now.

Summary

The structure of the experiment is as follows:

- The main part of the experiment consisting of 40 rounds.
- Questionnaire, part 1. After answering these questions you will be informed about your final earnings in ECU.
- Questionnaire, part 2. Once you have completed it, you will be informed about your final cash earnings in euros.
- Please stay seated until the experimenter asks you to come and receive the earnings.

2.B.2 Instruction for Treatment eT

INSTRUCTIONS

Welcome to the experiment. Please read the instructions carefully. They are identical for all the participants with whom you will interact during this experiment.

If you have any questions please raise your hand. One of the experimenters will come to you and answer your questions. From now on communication with other participants is not allowed. If you do not conform to these rules you will be excluded from the experiment with no payment. Please do also switch off your mobile phone at this moment.

In this experiment you can earn some money. How much you earn depends on your decisions and the decisions of the other participants. During the experiment we will refer to ECU (Experimental Currency Unit) instead of Euro. The total amount of ECU that you will have earned during the experiment will be converted into Euro at the end of the experiment and paid to you in cash confidentially. The conversion rate that will be used to convert your ECU earnings into your Euro cash payment will be shown to you on the screen at the beginning of the experiment.

The Experiment

The main part of this experiment consists of 4 blocks with 10 rounds in each block. In each block, the first five rounds are for practice only so that you can experiment without affecting your cash earnings from this experiment. At the end of the experiment, one round out of the final five rounds in each block (that is one round of rounds 6-10 in each block) will be randomly selected and the sum of your round incomes in these selected rounds will be converted into euros and paid to you in cash.

The Task

The task is the same in each of the 40 rounds. In each round you will receive an endowment of 16 ECU. The endowment can be used to purchase "tokens". Each token costs 1 ECU so that you can purchase up to 16 of these tokens. You have to buy at least one token. Any part of your endowment that you do not spend on tokens will be added to your round income.

In every round you will be matched to a computerized participant (computer). The computer will buy a certain number of tokens (between 1 and 16), and this number is pre-determined before the start of the experiment. The number of

tokens bought by the computer in any round will be announced on the screen before you make your buying decision for that round. This amount will be the same in each round of a block, but will change from one block to another. That means that the computer will buy the same number of tokens in rounds 1-10, 11-20, 21-30 and 31-40.

After you have chosen how many tokens to buy, you and the computer will receive an extra 16 ECU. The share of the extra 16 ECU that you receive depends on the number of tokens that you buy and the number of tokens that are bought by the computer. More precisely, the share of the extra 16 ECU is given by:

$$\text{Share of the extra 16 ECU} = \frac{\text{Number of tokens you bought}}{\text{Number of tokens you bought} + \text{Number of tokens the other participant bought}} \times 100\%$$

For example, if you purchase the same number of tokens as the computer, the share of 16 ECU will be equal to 50%, meaning that you will receive 8 ECU.

Your round income will be computed based on the number of tokens that you bought and the share of the extra 16 ECU that you have received. Once the round is over, you will be informed about the number of tokens bought by you and by the computer, the share of the extra 16 ECU, your round income, and that of the computer. Information about the number of tokens and the share of 16 ECU allocated to you and to the other participant will also be represented visually.

- Your round income will be:

  Round Income = 16 ECU − Number of purchased tokens + (Share of the extra 16 ECU)*16 ECU

At the end of the experiment, four rounds will be randomly selected for payment. More precisely, the first round for payment will be randomly selected from rounds 6-10, the second round for payment from rounds 16-20, the third from rounds 26-30 and the fourth one from rounds 36-40. Outcomes in all other rounds will not influence your final earnings, but you will not know which rounds will be selected until the end of the experiment.

After the 40 rounds we will ask you to fill in a short questionnaire. The questionnaire will have two parts, each of which will be explained on the screen before you start answering the questions. After the first part of the questionnaire you will be informed about all of your round incomes as well as about

the four rounds that were randomly selected for payment (see Figure 1 on the next page). After the second part of the questionnaire you will be informed about your final earnings in euro. You will receive these earnings in cash and in private at the end of the experiment. Please stay seated until we ask you to come to receive the earnings.

If you have any further questions, please raise your hand now.

Summary

- The main part of the experiment consisting of 40 rounds.
- Questionnaire, part 1. After answering these questions you will be informed about your final earnings in ECU.
- Questionnaire, part 2. Once you have completed it, you will be informed about your final cash earnings in euros.
- Please stay seated until the experimenter asks you to come and receive the earnings.

Your income in the payoff-relevant rounds

Block 1		Block 2		Block 3		Block 4	
Round	Income	Round	Income	Round	Income	Round	Income
6	22.00	16	25.00	26	28.00	36	24.00
7	31.00	17	4.00	27	25.00	37	17.00
8	26.00	18	31.00	28	25.00	38	4.00
9	10.00	19	20.00	29	7.00	39	4.00
10	11.00	20	26.00	30	26.00	40	26.00

The rounds that were randomly selected for payment are:

10, 19, 26 and 39.

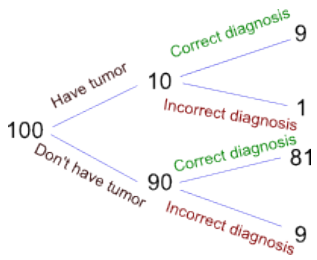
Figure 2.10: A screenshot of the final profit display. Round incomes in this screenshot were randomly generated. [Picture used in the instructions of all treatments]

2.B.3 Instructions for the Questionnaire

Apart from questions about age, gender, bachelor programme and previous experience in the lab, the following 8 questions were asked in the numeracy questionnaire. These questions were the ones with the lowest number of correct responses (less than 80% among respondents with higher education) among the 15 questions used in Peters et al. (2007).

- If Person A's chance of getting a disease is 1 in 100 in 10 years, and person B's risk is double that of A, what is B's risk?
- Imagine that you are taking a class and your chances of being asked a question in class are 1% during the first week of class and double each week hereafter (i.e., you would have a 2% chance in Week 2, a 4% chance in Week 3, an 8% chance in Week 4). What is the probability that you will be asked a question in class during Week 7? (in %)
- Suppose that 1 out of every 10,000 doctors in a certain region is infected with the SARS virus; in the same region, 20 out of every 100 people in a particular at-risk population also are infected with the virus. A test for the virus gives a positive result in 99% of those who are infected and in 1% of those who are not infected. A randomly selected doctor and a randomly selected person in the at-risk population in this region both test positive for the disease. Who is more likely to actually have the disease?
- In the Acme Publishing Sweepstakes, the chance of winning a car is 1 in 1,000. What percentage of tickets of Acme Publishing Sweepstakes wins a car?
- Imagine that we roll a fair, six-sided die 1,000 times. Out of 1000 rolls, how many times do you think the die would come up even (2, 4, or 6)?
- In the Big Bucks Lottery, the chances of winning a 10 prize are 1%. What is your best guess about how many people would win a 10 prize if 1000 people each buy a single ticket from Big Bucks?
- The chance of getting a viral infection is 0.0005. Out of 10,000 people, about how many of them are expected to get infected?
- Suppose you have a close friend who has a lump in her breast and must have a mammogram. Of 100 women like her, 10 of them actually have a malignant tumor and 90 of them do not. Of the 10 women who actually have a tumor, the mammogram indicates correctly that 9 of them have a

tumor and indicates incorrectly that 1 of them does not. Of the 90 women who do not have a tumor, the mammogram indicates correctly that 81 of them do not have a tumor and indicates incorrectly that 9 of them do have a tumor. The graph below summarizes all of this information. Imagine that your friend tests positive (as if she had a tumor), what is the likelihood that she actually has a tumor?



Chapter 3

Inefficient Lock-in with Sophisticated and Myopic Players

3.1 Introduction

This chapter as well as chapter 4 switch focus from a game of conflict to a game of coordination. The problem faced by game theory in coordination games is quite different from the problem addressed in chapter 2: in contests Nash equilibrium could not explain experimental data, but in coordination games the set of subgame perfect Nash equilibria is so large that it leads to vacuous prediction about the choice path. Any repetition of a stage game Nash equilibria could be supported by some subgame perfect Nash equilibrium, but even miscoordination can occur at the start of the game if it is followed by efficient coordination. One reason for the multiplicity of equilibria is the lack of history dependence. As an example, consider figure 3.1 that represents two stages of a repeated game between players 1 and 2. Subgames starting at nodes $1b$ and $1c$ for player 1 are identical[1], therefore if there is an equilibrium that supports an action for player A in node $1b$, there will also be an equilibrium that supports this action in node $1c$. Nash equilibrium requires mutually consistent beliefs and actions but places no restrictions on how beliefs should depend on observed history. However, even though expecting the same action to be played is just as rational as expecting a different action (Goodman, 1983), there is robust experimental evidence that choices and beliefs do depend on past play, especially in games with multiple stable states (Van Huyck et al., 1990; Romero, 2011).

[1]Except for the accumulated earnings that play no role under the standard assumptions of risk neutrality and selfishness.

We use this finding to place additional restrictions on the belief formation process and develop a solution concept that depends on past play and refines the predictions of a subgame perfect Nash equilibrium.

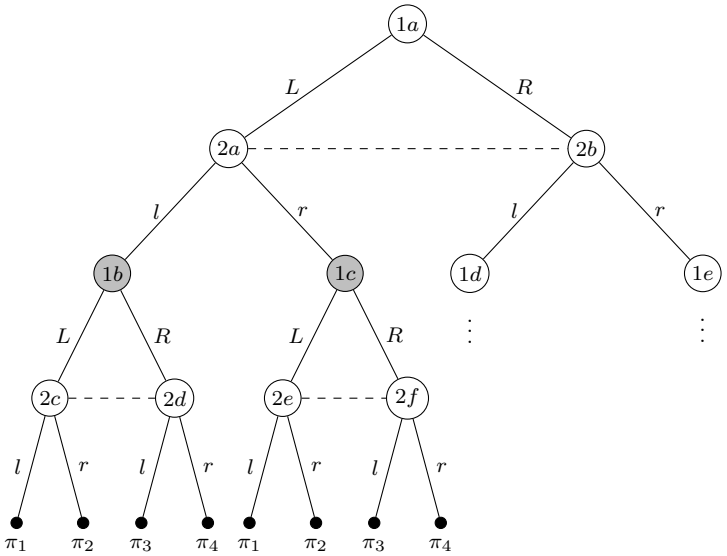


Figure 3.1: Two stages of a repeated two player game, where the first number indicates the player to whom the node belongs. End nodes display payoffs from the second stage.

Instead of using a solution concept such as a subgame perfect Nash equilibrium we could use a learning model which makes predictions about the path of play based on outcomes in previous rounds. However, in such learning models (see Fudenberg and Levine, 1998, Camerer, 2003, chapter 6) choices are determined only by observed history, ignoring the structure of upcoming rounds. In this chapter the belief formation assumed in learning models is combined with an equilibrium concept to define a solution concept that takes into account both the observed history and the structure of future rounds.

Players in our model are assumed to be either "myopic" or "sophisticated". Myopic players behave as predicted by adaptive learning models: they form beliefs about the actions of other players, update them based on observed history (we use a weighted fictitious play model by Cheung and Friedman, 1997) and choose a myopic best response. Sophisticated players have a certain planning horizon and compare payoffs of action plans that prescribe an action for each point in time within this planning horizon. We also assume that sophisticated players anticipate the learning process of myopic players and know about other sophisticated players, therefore our solution concept requires action plans of sophisticated players to be mutual best responses to each other.

One of the reasons for developing such a solution concept is its ability to make predictions following a particular history of choices. Game theory and experimental economics typically abstract from the experience that players might have prior to the game, but in many real life situations decisions are made repeatedly and choices are sensitive to conventions that have been established by past actions. Such situations could be modelled using a coordination game with multiple stable states in which play can converge to an inefficient state (Arthur, 1989), establishing a history that makes subsequent efficient coordination hard. It is important to have a theory that could explain how transitions to an efficient state depend on game parameters and history of play, but existing models are not able to do that. An adaptive learning model with a deterministic choice rule predicts that an inefficient state would be implemented in all rounds because of a positive feedback loop between beliefs and actions. Subgame perfect Nash equilibrium could support any common action and the history of inefficient coordination would play no role. A model presented in this chapter combines the two approaches and predicts that a transition from an inefficient to the efficient state can occur if certain conditions are satisfied, while transitions in the opposite direction never occur.

Our model also predicts that some players may deviate from an inefficient state, but none will deviate from the efficient one, therefore the efficient state is absorbing and there is a unique point in time when play transitions from the inefficient to the efficient state. For sophisticated players action paths that prescribe a switch from an efficient to an inefficient action are dominated, therefore sophisticated players will switch to the efficient action at most once. We calculate how such action plans of sophisticated players affect the switching period of myopic players, and how the latter affects sophisticated player payoffs. This mapping from sophisticated player action plans to payoffs is then used to determine the combinations of action plans that are mutual best responses to each other.

Depending on the parameters of the game, three types of equilibria may exist: in a "teaching equilibrium" sophisticated players switch to the efficient action at the start of the game, and myopic players switch later. In an "interior equilibrium" sophisticated players initially play the inefficient action, but switch to the efficient one and are subsequently followed by myopic players. In a "delay" equilibrium all sophisticated players choose the inefficient action for the entire duration of the game, and myopic players never switch. Point predictions cannot be made because of the multiplicy of equilibria, therefore instead we show how the speed of transition and the types of equilibria that exist re-

spond to changes in game parameters. We find that as the planning horizon of sophisticated players increases, the teaching and the interior equilibria are more likely to exist, while the delay equilibrium is less likely. A longer history of inefficient coordination makes teaching equilibrium less likely and delays transitions. The effect of player composition is ambiguous: on one hand, a larger number of sophisticated players leads to a faster transition and higher profits in the teaching and interior equilibria, reducing incentives to completely stop teaching. On the other hand, as the number of sophisticated players grows, one player's actions have a smaller effect on myopic players, increasing incentives to delay teaching and leading to a potential breakdown of a teaching equilibrium.

Several other studies have extended the adaptive learning model with sophistication in different ways. Camerer et al. (2002a) and Camerer et al. (2002b) propose a sophisticated experience-weighted attraction (EWA) model in which some players are adaptive and learn using adaptive EWA, while others are sophisticated, anticipate how adaptive players learn and use strategic teaching. While conceptually this chapter is similar to the model of sophisticated EWA, we develop a solution concept that can be used to make ex-ante predictions about the path of play in the game, while the parameters of sophisticated EWA can be estimated only ex-post. Ellison (1997) models a population of adaptive players, learning according to fictitious play, repeatedly matched in pairs to play a binary choice coordination game. Adding one rational player to the population of adaptive players can change the outcome from coordination on the inefficient equilibrium to coordination on the efficient one, as long as the number of players is fixed and the rational player is patient enough. Acemoglu and Jackson (2015) develop an overlapping generations model that shows how a social norm of low cooperation can be overturned by a single forward-looking player. Schipper (2011) uses an optimal control model with two players and shows how a strategic player can control an adaptive player in repeated games with strategic substitutes or strategic complements. Mengel (2014) studies adaptive players who are also forward-looking and finds that in two-player coordination games the efficient equilibrium may be stochastically stable, in contrast to the the case with only adaptive players.

3.2 Sophisticated Player Equilibrium

Consider n players, indexed by $i \in N \equiv \{1, 2, \ldots, n\}$, who play a repeated game in continuous time by choosing an action from a stage game action space $\{A, B\}$. We denote the time at which the game starts by 0, the duration of the remaining

game by $\bar{T}$ and the duration of observed history by T', with $\bar{T}, T' \in (0, \infty)$. We implement the history of inefficient coordination by assuming that prior to time 0 only action B has been chosen.

We assume two types of players: m players are myopic and $n-m$ players are sophisticated. Throughout the chapter we will index sophisticated players by $s \in S$ and myopic players by $i \in N \backslash S$. The two types of players follow different choice rules, respectively denoted by a_i and a_s, which prescribe an action for each moment in time. We will refer to a_i as a *choice function* and to a_s as an *action plan*. Denote the action of player i at time t by $a_i(t)$ and the action of player s by $a_s(t)$, where action A is coded as 1 and action B is coded as 0. Denote the combination of actions of all players except i by $a_{-i}(t) = \times_{j \in N \backslash \{i\}} a_j(t)$, with $a_{-i}(t) \in A_{-i}$, and denote the combination of actions of all sophisticated players except s by $a_{-s}(t) = \times_{j \in S \backslash \{s\}} a_j(t)$, with $a_{-s}(t) \in A_{-s}$. The payoff flow for player i at time t is $\pi_i(a_i(t), \times_{j \in N \backslash \{i\}} a_j(t))$. Similarly, denote the combination of choice functions of all myopic players except i by $a_{-i} = \times_{j \in \{N \backslash S\} \backslash \{i\}} a_j$ and the combination of action plans for all sophisticated players except s by $a_{-s} = \times_{j \in S \backslash \{s\}} a_j$.

The difference between a choice function for myopic players and an action plan for sophisticated players lies in how these functions are determined: choices of myopic players are determined by the history of play while the choices of sophisticated players must be optimal given the choices of all other players. Before specifying these two function we first have to define the beliefs and expected payoffs of myopic players.

Belief of a myopic player is a probability assigned to the event that a randomly chosen other group member chooses action A. Denote the belief of player i at time t by $x_i(t)$. Belief formation is assumed to follow a one parameter weighted fictitious play model,[2] proposed by Cheung and Friedman (1997). The original weighted fictitious play model is specified for two player games and we extend it to N-person games by assuming that a joint distribution of choices is used to form beliefs about the actions of group members, but players do not distinguish between the identities of others.[3] Beliefs are therefore homogeneous

[2]Fictitious play corresponds to Bayesian updating of the probability that any group member will choose A, using a Dirichlet prior and assuming that the choice of each group member was independently drawn from the distribution about which players are learning.

[3]There are several other ways how weighted fictitious play could be extended to N-person games. One way could be to assume that players form beliefs about the joint distribution of the actions of all others and update it using observed aggregate feedback: for example, Crawford (1995) assumes that players form beliefs and observe feedback about an order statistic of all the choices. Another way is to assume that separate beliefs are formed about every other player j based on the empirical distribution of j's choices (e.g. Monderer and Shapley, 1996). We combine the two approaches by assuming that players use the joint distribution of choices to form beliefs about the action of any opponent, but do not distinguish between their identities.

(Rapoport, 1985; Rapoport and Eshed-Levy, 1989): a single belief is formed about the probability that any other player will choose A. The fictitious play rule used to calculate myopic player beliefs is as follows:

$$x_i(t) = \frac{\int_{k=0}^{t} \gamma^k \sum_{j \in N \setminus \{i\}} \frac{a_j(t-k)}{n-1} \, \mathrm{d}k}{\int_{k=0}^{t+T'} \gamma^k \, \mathrm{d}k} \tag{3.1}$$

The integral in the numerator measures the weighted length of time in which action A has been observed, determined by the action plans of other group members. Observations prior to time 0 play no role because we assume that prior to time 0 only action B has been observed. The γ parameter measures the rate at which old observations are forgotten. We assume that $\gamma \in (0, 1)$, where values close to 1 indicate that all past observations are given similar weights, while values close to 0 indicate that only the most recent experience is taken into account.

Expected payoffs of myopic players associated with each pure action are determined by beliefs $x_i(t)$, which are used to assign a probability to each action profile of other group members:

$$\begin{aligned} E\pi_i(a, x_i(t)) &= \sum_{a_{-i} \in A_{-i}} [Pr(a_{-i}(t) = a_{-i} | x_i(t)) \times \pi_i(a, a_{-i})] = \\ &= \sum_{a_{-i} \in A_{-i}} [x_i(t)^{(\sum a_{-i})} (1 - x_i(t))^{(n-1-\sum a_{-i})} \times \pi_i(a, a_{-i})], \quad \forall a \in \{1, 0\} \end{aligned} \tag{3.2}$$

Choice function $a_i(t, \times_{s \in S} a_s)$ prescribes an action for a myopic player i at any point in time $t \in [0, \bar{T}]$, conditional on the profile of action plans chosen by sophisticated players, $\times_{s \in S} a_s$. We assume that myopic players choose the action that maximizes immediate expected utility and ties are broken in favour of action A:

$$a_i(t, \times_{s \in S} a_s) = \begin{cases} 1 & \text{if } E\pi_i(1, x_i(t)) \geqslant E\pi_i(0, x_i(t)) \\ 0 & \text{otherwise} \end{cases} \tag{3.3}$$

Action plans chosen by sophisticated players, $\times_{s \in S} a_s$, are explicitly included in the choice function to make it transparent that myopic player actions can be affected by sophisticated players. Note that the choice function depends only on the current round payoffs and beliefs, which are determined by observed history, therefore it is possible to anticipate myopic player choices at any history.

Sophisticated players anticipate the learning process of myopic players and are also farsighted, thus at time 0 they choose an action plan for the

interval $[0, T]$, where T is the length of the planning horizon of sophisticated players.

Action plan a_s prescribes an action for a sophisticated player s at any point in time $t \in [0, T]$. Denote the set of all action plans by A_s. The action plan is assumed to be an open-loop strategy, which depends only on time and not on observed history. Sophisticated players face no strategic uncertainty about the actions of myopic players, but they do face uncertainty about the actions of other sophisticated players. Payoffs associated with an action plan a_s depend on the vector of action plans of other sophisticated players, a_{-s}, and on the choices of myopic players, whose choice function $a_i(t, a_s \times a_{-s})$ also depends on the action plans of all sophisticated players. The total payoff that a sophisticated player expects to earn over the period of length T is calculated as follows:

$$\Pi(a_s, a_{-s}, a_i(\cdot, a_s \times a_{-s})) = \int_0^T \pi[a_s(t), a_{-s}(t) \times a_i(t, a_s \times a_{-s})]dt \tag{3.4}$$

Since sophisticated players choose action plans and face no strategic uncertainty about the actions of myopic players, the game can be reduced to a static game between sophisticated players. Theoretical predictions in static games are typically made using a Nash equilibrium, so we follow the convention and require that sophisticated players choose action plans that are mutual best responses to each other.

Definition 1. *A combination of action plans* $\times_{s \in S} a_s^*$ *is a symmetric* ***sophisticated player equilibrium*** *if for each player* $s \in S$, a_s^* *satisfies*

$$\begin{aligned} \Pi(a_s^*, a_{-s}^*, a_i(\cdot, a_s^*, a_{-s}^*)) \geqslant \Pi(a_s, a_{-s}^*, a_i(\cdot, a_s, a_{-s}^*)), \quad & \forall a_s \in A_s \\ \text{and } a_s^* = a_j^*, \quad & \forall s, j \in S \end{aligned} \tag{3.5}$$

and $a_i(\cdot, a_s, a_{-s})$ *is defined in* (3.3).

If there were no myopic players, equation (3.5) would reduce to the standard Nash equilibrium. If all players were myopic, equation (3.5) would not apply, and the choices of all players would be calculated using the belief learning model. We will look at an intermediate case where both myopic and sophisticated players are present.

In the remainder of the chapter we will characterize the symmetric sophisticated player equilibria for a repeated N-person critical mass coordination game.

3.3 Sophisticated Player Equilibrium in a Critical Mass Game

We are interested in determining conditions under which an inefficient convention could be replaced by an efficient one. One way how such a transition could take place is by strategic choice: sophisticated players could attempt to teach other players to play according to the efficient convention. To determine conditions under which such strategic teaching is possible we will characterize symmetric sophisticated player equilibria following lock-in to an inefficient state.

3.3.1 Critical Mass Game

Recall that we defined a sophisticated player equilibrium for a class of games with n players and an action space $\{A, B\}$. A special class of such games is a critical mass game, in which payoffs of each player depend on their action, $a_i(t)$ and on the total number of other group members who chose action A at time t, denoted by $r(a_{-i}, t) = \sum_{j \in N \setminus \{i\}} a_j(t)$, with $r(a_{-i}, t) \in \{0, 1, \ldots, n-1\}$. The payoff flow for player i at time t is defined as follows:

$$\pi(a_i(t), a_{-i}(t)) = \begin{cases} H & \text{if } r(a_{-i}, t) \geqslant \theta \text{ and } a_i(t) = 1 \\ 0 & \text{if } r(a_{-i}, t) < \theta \text{ and } a_i(t) = 1 \\ M & \text{if } r(a_{-i}, t) \geqslant \theta \text{ and } a_i(t) = 0 \\ L & \text{if } r(a_{-i}, t) < \theta \text{ and } a_i(t) = 0 \end{cases} \tag{3.6}$$

To have a coordination game, we assume that $H > M$ and $L > 0$. The coordination requirement is determined by an exogenous threshold θ: action A generates a larger payoff than B if and only if at least θ other group members choose A. There are two stable states[4] in pure strategies if one point in time is considered in isolation: in the first stable state all players choose A and in the second one all players choose B. We assume that states are Pareto-ranked and define coordination on A as an efficient state by assuming that $H > L$. Finally, we assume that $M \geqslant L$, so that players who choose B also prefer a situation in which the threshold has been exceeded.

Assumption 1: H > M ⩾ L > 0.

We assume that there are at least 2 sophisticated players so that an equilibrium could be defined using equation 3.5. We also assume that the number of myopic players is sufficiently large to implement the efficient state, and the number of sophisticated players is small enough so that sophisticated players

[4] We will use the term "state" rather than "equilibrium" when referring to a Nash equilibrium in a stage game to avoid confusion with the sophisticated player equilibrium.

on their own could not implement the efficient state. If the latter condition was not satisfied, a sophisticated player equilibrium would reduce to the standard Nash equilibrium because sophisticated players would not need to take into account the learning process of myopic players.

Assumption 2: $2 \leqslant n - m < \theta \leqslant m$.

3.3.2 Choice Function of Myopic Players

Myopic players form beliefs about the actions of other players and choose an action that maximizes immediate payoffs. In this subsection we specify the choice function $a_i(t, a_s)$ that prescribes an action for player i at time t when sophisticated players are choosing action plans $\times_{s \in S} a_s$ (for brevity, we will omit the subscript under the product sign).

Proposition 1. *Suppose that in a game with payoffs defined by* (3.6) *at time t myopic player i holds beliefs* $x_i(t)$. *Then the choice function from* (3.3) *simplifies to:*

$$a_i(t, \times a_s) = \begin{cases} 1 & \text{if } x_i(t) \geqslant I^{-1}_{\frac{L}{L+H-M}}(\theta, n-\theta) \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

where I^{-1} *is the inverse of an incomplete regularized beta function.*

Proof.

From (3.3), action A is chosen if the expected payoff of A at time t exceeds the expected payoff of B:

$$a_i(t) = 1 \Leftrightarrow E\pi(1, x_i(t)) \geqslant E\pi(0, x_i(t)) \tag{3.8}$$

In a critical mass game payoff depends only on the chosen action and on whether the number of other group members who chose A exceeds θ. Denote the subjective probability assigned to the latter event by $Pr[r(a_{-i}, t) \geqslant \theta | x_i(t)]$. Then expected payoffs in equation (3.2) can be defined as:

$$\begin{aligned} E\pi(1, x_i(t)) &= 0 \times (1 - Pr[r(a_{-i}, t) \geqslant \theta | x_i(t)]) + H \times Pr[r(a_{-i}, t) \geqslant \theta | x_i(t)] \\ E\pi(0, x_i(t)) &= L \times (1 - Pr[r(a_{-i}, t) \geqslant \theta | x_i(t)]) + M \times Pr[r(a_{-i}, t) \geqslant \theta | x_i(t)] \end{aligned} \tag{3.9}$$

The subjective probability that the threshold will be exceeded is calculated by adding the probabilities assigned to all action profiles of other players in which more than θ players choose A:

$$Pr[r(a_{-i}, t) \geqslant \theta | x_i(t)] = \sum_{k=\theta}^{n-1} (x_i(t))^k (1 - x_i(t))^{n-1-k} \binom{n-1}{k} \tag{3.10}$$

Use equations (3.9) and (3.10) to rewrite (3.8) the following way:

$$a_i(t) = 1 \Leftrightarrow \sum_{k=\theta}^{n-1} (x_i(t))^k (1 - x_i(t))^{n-1-k} \binom{n-1}{k} \geqslant \frac{L}{L + H - M} \tag{3.11}$$

Notation in (3.11) is simplified using the definition of an incomplete regularized beta function:[5]

$$a_i(t) = 1 \Leftrightarrow I_{x(t)}(\theta, n - \theta) \geqslant \frac{L}{L + H - M} \tag{3.12}$$

Taking the inverse of (3.12) and substituting into (3.3) leads to the desired expression:

$$a_i(t, \times a_s) = \begin{cases} 1 & \text{if } x_i(t) \geqslant I^{-1}_{\frac{L}{L+H-M}}(\theta, n - \theta) \\ 0 & \text{otherwise} \end{cases}$$

■

Proposition 1 states that a myopic player chooses A instead of B if his probabilistic belief exceeds $I^{-1}_{\frac{L}{L+H-M}}(\theta, n - \theta)$, a threshold value that depends only on the game parameters. For brevity, we will refer to this threshold value by I^{-1}. We should note that the properties of inverse regularized beta functions imply that I^{-1} is increasing in L, M and θ, but decreasing in H and n.

Proposition 1 shows that myopic player actions can be determined by comparing their beliefs to a threshold value that is fixed in a given game. Once myopic player actions are know, Assumption 2 ensures that the efficient state is implemented if and only if all myopic players choose A. The next section will simplify the payoff calculation even further by showing that to know the payoff flow it is sufficient to know the first time when myopic player beliefs exceed the threshold value.

3.3.3 Undominated Action Plans of Sophisticated Players

This section shows that although sophisticated players could use action plans that prescribe many switches from one action to the other, undominated action plans must prescribe at most one switch from action B to action A and no switches from action A to action B. The sophisticated player action space can therefore be restricted to a set of real numbers that denote a switching time from A to B.

[5] An incomplete regularized beta function is defined as $I_c(a, b) = \sum_{k=a}^{a+b-1} c^k (1-c)^{a+b-1-k} \binom{a+b-1}{k}$. The function is well defined because $\frac{L}{L+H-M} \in (0, 1)$, from Assumption 1.

Definition 2. *Denote by U_s (for "undominated") the set of action plan profiles in which no sophisticated player is choosing strictly dominated action plans:*

$$U_s = \{\times_{s\in S} a_s \in A_s | \nexists a'_s : \Pi[a'_s, a_{-s}, a_i(\cdot, a'_s \times a_{-s})] > \Pi[a_s, a_{-s}, a_i(\cdot, a_s \times a_{-s})]\}$$

An action profile will be called *dominated* if it is not in set U_s, that is if in this action profile at least one sophisticated player is choosing a dominated action plan.

We will show that the set of undominated action plans cannot contain any strategies that prescribe a switch from A to B, but the proof requires two additional supporting lemmas.

Lemma 1. *If two action plans of the sophisticated player prescribe the same action, the payoff flow is higher for the action plan with which myopic player beliefs are higher:*

$$\pi[a'_s(t), a_{-s}(t) \times a_i(t, a'_s \times a_{-s})] \geqslant \pi[a_s(t), a_{-s}(t) \times a_i(t, a_s \times a_{-s})]$$
$$\textit{if} \quad x(t)' \geqslant x(t) \quad \textit{and} \quad a'_s(t) = a_s(t)$$

where $x(t)'$ is the belief held by myopic players if the sophisticated player uses action plan a'_s and $x(t)$ is the belief if the sophisticated player uses action plan a_s.

Proof: see Appendix 3.A.2.

Lemma 1 shows that sophisticated players can only benefit from myopic players assigning a higher probability to others choosing A. The proof is based on an observation that the tendency for myopic players to choose A is increasing in their beliefs and sophisticated player payoffs are increasing in the number of players who choose action A.

Definition 3. *Denote by AB_M the set of action plan profiles for sophisticated players with which myopic players switch from A to B:*

$$\begin{aligned} AB_M = \{\times_{s\in S} a_s \in A_s | \exists t_1, t_2 \in [0, T] : \quad & t_1 < t_2 \\ & a_i(t_1, \times_{s\in S} a_s) = 1 \\ & a_i(t_2, \times_{s\in S} a_s) = 0\} \end{aligned}$$

Lemma 2. *All action plan profiles for sophisticated players with which myopic players switch from A to B are strictly dominated:*

$$AB_M \cap U_s = \varnothing$$

Proof: see Appendix 3.A.2.

The intuition of Lemma 2 is straightforward: if myopic players ever switch to an efficient action A, the participation threshold will be exceeded as long as sophisticated players continue choosing action A. Consequently, sophisticated players who would choose B would lower their earnings. However, note that the proof rests on Assumption 2, which says that the number of myopic players exceeds the participation threshold. If this assumption did not hold, an argument about dominance could not be made because other sophisticated players may prevent efficient coordination by switching to B, which would make switching to B optimal.

Definition 4. *Denote by AB_S the set of action plan profiles for sophisticated players with which at least one sophisticated player switches from A to B:*

$$AB_S = \{\times_{s\in S} a_s \in A_s | \exists t_1, t_2 \in [0,T], s \in S: \quad \begin{aligned} & t_1 < t_2 \\ & a_s(t_1) = 1 \\ & a_s(t_2) = 0\} \end{aligned}$$

Proposition 2. *Action plan profiles for sophisticated players that prescribe a switch from A to B for at least one sophisticated player are dominated:*

$$AB_S \cap U_s = \varnothing$$

Proof.

Take an action plan profile $\times_{s\in S} a_s \in AB_S$. We will show that in this action profile at least one sophisticated player must be choosing an action plan that is dominated.

If $\times_{s\in S} a_s \in AB_M$, at least one sophisticated player must be choosing a dominated action plan, from Lemma 2, and the proof would be completed. Alternatively, assume that $\times_{s\in S} a_s \in \{AB_S \backslash AB_M\}$. By the definition of AB_S, there must be a sophisticated player whose action plan prescribes a switch from A to B; denote the action plan of this player by $\widetilde{a}_s$ and denote the switching time prescribed by $\widetilde{a}_s$ by t'. Then there must be some small ϵ such that $\widetilde{a}_s(t) = 1$ if $t \in [t' - \epsilon, t')$ and $\widetilde{a}_s(t) = 0$ if $t \in [t', t' + \epsilon]$. Since $\times_{s\in S} a_s \notin AB_M$, myopic players switch from B to A at most once, thus their choices can be described by a number $\hat{t}(\widetilde{a}_s)$ that identifies this switching time: B is chosen in the interval $[0, \hat{t}(\widetilde{a}_s))$ and A is chosen in the interval $[\hat{t}(\widetilde{a}_s), T]$.

First, suppose that $t' \geqslant \hat{t}(\widetilde{a}_s)$, then myopic players would be choosing A at any time $t \geqslant t'$. Assumption 2 implies that the threshold will be exceeded at any such point in time, therefore an action plan $\widetilde{a}_s$ is dominated by an action

plan that prescribes A at each point in time $t \geqslant \hat{t}(\widetilde{a}_s)$. Next, suppose that $t' < \hat{t}(\widetilde{a}_s)$ and $\hat{t}(\widetilde{a}_s) > T$. Then myopic players will choose B for the entire period that is taken into account by the sophisticated player, thus action plan $\widetilde{a}_s$ will be dominated by an action plan that prescribes B at all times.

Alternatively, suppose that $\hat{t}(\widetilde{a}_s) > t'$ and $\hat{t}(\widetilde{a}_s) \leqslant T$ (see an illustration in figure 3.2). Choose ϵ to be sufficiently small to satisfy $\hat{t}(\widetilde{a}_s) > t' + \epsilon$. Then for any $\widetilde{a}_s$ construct an action plan a'_s the following way:

$$a'_s(t) = \begin{cases} \widetilde{a}_s(t) & \text{if } t \in [0, t' - \epsilon) \cup (t' + \epsilon, T] \\ 0 & \text{if } t \in [t' - \epsilon, t'] \\ 1 & \text{if } t \in (t', t' + \epsilon] \end{cases}$$

In other words, a'_s is constructed by taking $\widetilde{a}_s$ and swapping choices prescribed in the interval $(t' - \epsilon, t')$ with choices prescribed in the interval $(t', t' + \epsilon)$. We will show that $\widetilde{a}_s$ is dominated by a'_s.

The comparison of payoff flows generated by these two action plans is shown in figure 3.2. In the interval $[0, t + \epsilon)$ the sum of payoff flows is the same for both action plans ($\pi_1 + \pi_2 + \pi_3$ in figure 3.2). Payoffs are equal because with both action plans myopic players choose B in this entire interval (both $\hat{t}(a'_s)$ and $\hat{t}(\widetilde{a}_s)$ exceed $t' + \epsilon$), therefore the participation threshold is never exceeded. Action plan $\widetilde{a}_s$ prescribes A for the same duration of time as a'_s, therefore the sum of payoffs in the interval $[0, t' + \epsilon)$ would be the same for both action plans.

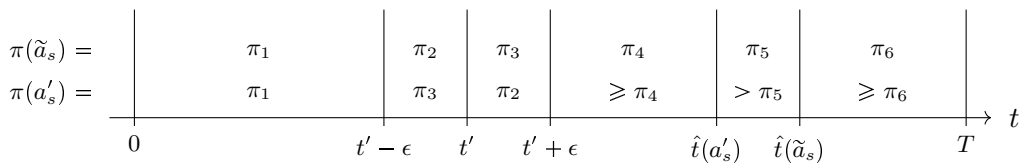


Figure 3.2: Payoff flows generated by action plans $\widetilde{a}_s$ and a'_s for the case $\hat{t}(\widetilde{a}_s) > t'$ and $\hat{t}(\widetilde{a}_s) \leqslant T$.

In the interval $[t' + \epsilon, T]$ the sum of payoffs generated by a'_s is strictly higher than that of $\widetilde{a}_s$. Since $\widetilde{a}_s(t) = a'_s(t)$, $\forall t \in (t' + \epsilon, T]$, any payoff difference between the two action plans in this interval must be due to the choices of myopic players. From equation 3.1, $x_i(t)$ would be the same under $\widetilde{a}_s(t)$ as under $a_s(t)'$ if γ was equal to 1. But as $\gamma \in (0, 1)$, older observations receive less weight and therefore myopic player beliefs would be strictly higher following a'_s than following $\widetilde{a}_s$ at any time $t \in (t' + \epsilon, T]$. Then Lemma 1 implies that the payoff flow is always weakly higher for a'_s at any time in the interval $[t' + \epsilon, T]$. To get strict dominance, note that $\hat{t}(a'_s) < \hat{t}(\widetilde{a}_s)$, for the following reasons. Since $\hat{t}(a'_s) \in (t' + \epsilon, T]$ and $x(t)$ is continuous, the switching period $\hat{t}(a'_s)$ must satisfy $x'_i(\hat{t}(a'_s)) = I^{-1}$. But since $\widetilde{x}'_i(t) < x'_i(t)$, $\forall t \in (t' + \epsilon, T]$, it must also

hold that $\widetilde{x}_i(\hat{t}(a'_s)) < x'_i(\hat{t}(a'_s)) = I^{-1}$. Consequently, the intersection of beliefs $\widetilde{x}_i(t)$ and belief threshold I^{-1} must occur strictly later, so that $\hat{t}(a'_s) < \hat{t}(\widetilde{a}_s)$. In the interval $(\hat{t}(a'_s), \hat{t}(\widetilde{a}_s))$ action plan $\widetilde{a}_s$ provides a flow of payoffs of at most L, while a'_s provides a payoff of H because more than θ players are choosing A.

The comparison of payoff flows associated with action plans $\widetilde{a}_s$ and a'_s is shown in figure 3.2. The sum of payoff flows generated by a'_s will be strictly higher than the sum of payoff flows generated by $\widetilde{a}_s$, therefore action plan $\widetilde{a}_s$ that prescribes switching from A to B is strictly dominated by another action plan a'_s.

■

The intuition of the proof is as follows: suppose that a sophisticated player switches from A to B. If myopic players switch from B to A at the same time or earlier, a sophisticated player would do better by always playing A instead. If myopic players never switch to A, there would be no incentive to play A in the first place. If myopic players switch at some time after the sophisticated player, the sophisticated player can strictly increase the earnings by teaching less at the start of the game and teaching more later.[6] Doing so would not reduce the payoffs prior to the switch, but would strictly decrease the switching time of the myopic players, because weighted fictitious play puts more weight on recent experience. Consequently, whenever sophisticated players are considering teaching for some period of time, they would be better off concentrating all the teaching just before the predicted switch of myopic players, thus a switch from A to B would never occur.

This section has shown that if sophisticated players do not choose dominated action plans, both myopic and sophisticated players will switch from B to A at most once, thus in the equilibrium the path of choices for either type can be described by a scalar indicating the switching time.

Each action path of myopic players that can be induced by undominated action paths of sophisticated players has the following structure:

$$a_i(t, \times a_s) = \begin{cases} 0 & \text{if } t \in [0, \hat{t}(\times a_s)) \\ 1 & \text{if } t \in [\hat{t}(\times a_s), T] \end{cases} \qquad \forall a_s \in U_s$$

Define $\hat{t} \in (0, \infty)$ as the **switching period** of myopic players. Note that $\hat{t} > 0$ because equation 3.1 implies that $x_i(0) = 0$, thus B is chosen at time 0.

Each undominated action plan for sophisticated players has the following structure:

$$a_s(t) = \begin{cases} 0 & \text{if } t \in [0, y_s) \\ 1 & \text{if } t \in [y_s, T] \end{cases} \qquad \forall a_s \in U_s$$

[6] By "teaching" we mean choosing action A to induce myopic players to choose A in the future.

Define $y_s \in [0, T]$ as a **strategy** for player s.

In the next section we will specify how the switching period of myopic players depends on the strategies of sophisticated players.

3.3.4 Optimal Switching Period for Myopic Players

The characterization of symmetric sophisticated player equilibria requires information about payoffs in an equilibrium and payoffs from potential deviations: in the first case all $(n - m)$ sophisticated players choose the same strategy, in the second case $(n - m - 1)$ sophisticated players choose one strategy and one player chooses a different one. Denote the strategy of one sophisticated player by $y_s = y$ and the strategy of other $n - m - 1$ sophisticated players by $y_j = \bar{y}$, for all $j \in \{S \backslash s\}$. Sophisticated player payoffs are determined by the switching period of myopic players, thus we first specify function $\hat{t}(y, \bar{y})$ that shows how the myopic player switching period depends on y and $\bar{y}$.

There are three cases to consider. In the first case, $\hat{t}(y, \bar{y}) > \max\{y, \bar{y}\}$, so that myopic players observe no other players choosing A from time 0 to time $\min\{y, \bar{y}\}$, a fraction of $\frac{n-m}{n-1}$ others choosing A from time $\max\{y, \bar{y}\}$ to $\hat{t}(y, \bar{y})$ and either a fraction of $\frac{1}{n-1}$ others choosing A from time $\bar{y}$ to time y (if $y > \bar{y}$) or a fraction of $\frac{n-m-1}{n-1}$ others choosing A from time y to $\bar{y}$ (if $\bar{y} > y$). Feedback observed by myopic players in this case is illustrated in figure 3.3.

In the second case, $y < \hat{t}(y, \bar{y}) < \bar{y}$. This will be true only if $\frac{1}{n-1} > I^{-1}$, that is if myopic players would switch to A after observing only one player choosing A. In this case each myopic player will observe no others choosing A from time 0 to y and a fraction of $\frac{1}{n-1}$ others choosing A from time y to $\hat{t}(y, \bar{y})$.

In the third case, $\bar{y} < \hat{t}(y, \bar{y}) < y$. Then each myopic player will observe no others choosing A from time 0 to $\bar{y}$ and a fraction of $\frac{n-m-1}{n-1}$ others choosing A from time $\bar{y}$ to $\hat{t}(y, \bar{y})$.

It is never possible that $\hat{t}(y, \bar{y}) < \min\{y, \bar{y}\}$ because at time $t \in [0, \min\{y, \bar{y}\})$ myopic players observe no others choosing A and therefore always choose B.

Proposition 3. *The switching period of myopic players is:*

$$\hat{t}(y, \bar{y}) = \begin{cases} \hat{t}_2(y) & \textit{if } y < \hat{t}_2(y) \leqslant \bar{y} & \textit{and} & \frac{1}{n-1} > I^{-1} \\ \hat{t}_3(\bar{y}) & \textit{if } \bar{y} < \hat{t}_3(\bar{y}) \leqslant y & \textit{and} & \frac{n-m-1}{n-1} > I^{-1} \\ \hat{t}_1(y, \bar{y}) & \textit{if } \max\{y, \bar{y}\} < \hat{t}_1(y, \bar{y}) & \textit{and} & \frac{n-m}{n-1} > I^{-1} \\ \infty & \textit{otherwise} \end{cases} \tag{3.13}$$

such that

$$\hat{t}_1(y,\bar{y}) = \frac{\log(\frac{n-m}{n-1} - I^{-1}) - \log(\gamma^{-\bar{y}}\frac{n-m-1}{n-1} + \gamma^{-y}\frac{1}{n-1} - \gamma^{T'}I^{-1})}{\log(\gamma)} \tag{3.14}$$

$$\hat{t}_2(y) = \frac{\log(\frac{1}{n-1} - I^{-1}) - \log(\gamma^{-y}\frac{1}{n-1} - \gamma^{T'}I^{-1})}{\log(\gamma)} \tag{3.15}$$

$$\hat{t}_3(\bar{y}) = \frac{\log(\frac{n-m-1}{n-1} - I^{-1}) - \log(\gamma^{-\bar{y}}\frac{n-m-1}{n-1} - \gamma^{T'}I^{-1})}{\log(\gamma)} \tag{3.16}$$

where y is the strategy of one sophisticated player and $\bar{y}$ is the strategy of other $(n-m-1)$ sophisticated players.

It is never possible that more than one condition of 3.13 is satisfied because $\hat{t}_1(y,\bar{y}) \leqslant \hat{t}_2(y)$ and $\hat{t}_1(y,\bar{y}) \leqslant \hat{t}_3(y)$ (see Lemma 10 in Appendix 3.A).

Proof.

Case 1: $\hat{t}(y,\bar{y}) > \max\{y,\bar{y}\}$

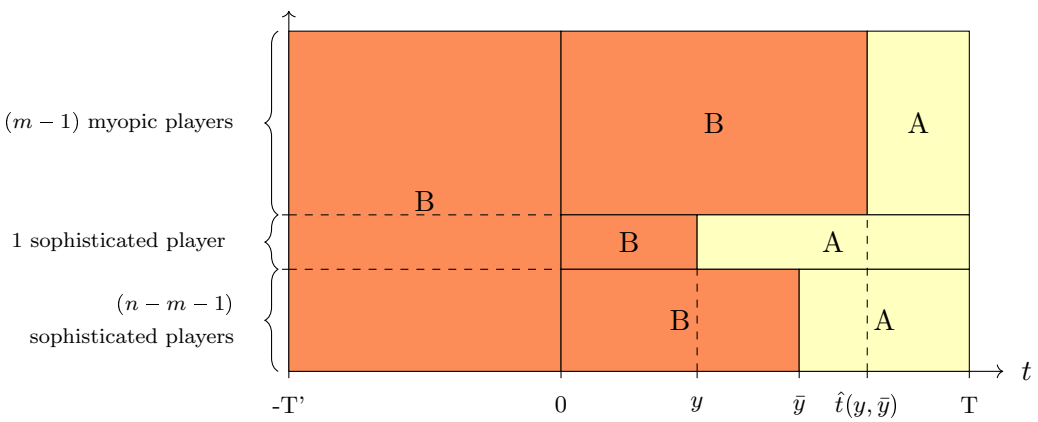


Figure 3.3: Illustration of the feedback observed by a single myopic player in the first case, where $\hat{t}(y,\bar{y}) > \max\{y,\bar{y}\}$. In this example $\bar{y} > y$. Vertical axis shows the fraction of other players choosing A or B, horizontal axis shows the passage of time. The first sophisticated player switches from B to A at time y, other $(n-m-1)$ sophisticated players switch at time $\bar{y}$ and myopic players switch at time $\hat{t}(y,\bar{y})$

Recall that beliefs of myopic players are calculated using weighted fictitious play from equation 3.1. If sophisticated players are using strategies y and $\bar{y}$, myopic player beliefs at any time $t \in (\max\{y,\bar{y}\}, \hat{t}(y,\bar{y})]$ will be calculated using the following rule:

$$x_i(t) = \frac{\int_{k=0}^{t-\bar{y}} \gamma^k (\frac{n-m-1}{n-1}) \mathrm{d}k + \int_{k=0}^{t-y} \gamma^k (\frac{1}{n-1}) \mathrm{d}k}{\int_{k=0}^{t+T'} \gamma^k \mathrm{d}k} =$$
$$= \frac{(\gamma^{t-\bar{y}} - 1)(\frac{n-m-1}{n-1}) + (\gamma^{t-y} - 1)(\frac{1}{n-1})}{\gamma^{t+T'} - 1}$$

Expressions in the numerator correspond to the history observed by a myopic player at time $t \in (\max\{y, \bar{y}\}, \hat{t}(y, \bar{y})]$: $(n - m - 1)$ sophisticated players are observed choosing A for a period of $t - \bar{y}$ and one sophisticated player is observed choosing A for a period of $t - y$. This feedback is illustrated in figure 3.3. The denominator measures the length of the entire history, including the T' rounds of inefficient coordination.

From Proposition 1, myopic players will choose A at time t if $x_i(t) \geqslant I^{-1}$:

$$a_i(t) = 1 \quad \Leftrightarrow \quad \frac{(\gamma^{t-\bar{y}} - 1)(\frac{n-m-1}{n-1}) + (\gamma^{t-y} - 1)(\frac{1}{n-1})}{\gamma^{t+T'} - 1} \geqslant I^{-1} \quad \Leftrightarrow$$
$$\gamma^{t+T'}(\gamma^{-\bar{y}-T'} \frac{n-m-1}{n-1} + \gamma^{-y-T'} \frac{1}{n-1} - I^{-1}) \leqslant \frac{n-m}{n-1} - I^{-1} \tag{3.17}$$

If $\frac{n-m}{n-1} - I^{-1} \leqslant 0$, equation (3.17) is never satisfied because of the following relationship that contradicts (3.17):

$$\gamma^{t+T'}(\gamma^{-\bar{y}-T'} \tfrac{n-m-1}{n-1} + \gamma^{-y-T'} \tfrac{1}{n-1} - I^{-1}) > \gamma^{t+T'}(\tfrac{n-m}{n-1} - I^{-1}) \geqslant \tfrac{n-m}{n-1} - I^{-1} \tag{3.18}$$

The first inequality holds because $\gamma^{-\bar{y}-T'} > 1$ and $\gamma^{-y-T'} > 1$ and the second inequality holds because $\gamma^{t+T'} < 1$ and $\frac{n-m}{n-1} - I^{-1} \leqslant 0$. But (3.18) contradicts (3.17), therefore if $\frac{n-m}{n-1} - I^{-1} \leqslant 0$, equation (3.17) is never satisfied and myopic players would choose B at any time t.

Alternatively, if $\frac{n-m}{n-1} - I^{-1} > 0$, condition (3.17) can be expressed the following way:

$$\gamma^{-t} \geqslant \frac{\gamma^{-\bar{y}} \frac{n-m-1}{n-1} + \gamma^{-y} \frac{1}{n-1} - \gamma^{T'} I^{-1}}{\frac{n-m}{n-1} - I^{-1}} \tag{3.19}$$

The left-hand side of (3.19) is strictly increasing in t and unbounded for any $\gamma \in (0, 1)$, so (3.19) will be satisfied for some t, although not necessarily with $t \leqslant T$. Equation (3.19) is not satisfied for $t = 0$ because the RHS of (3.19) is always strictly larger than 1 (RHS is increasing in both y and $\bar{y}$, but RHS > 1 even if $y = \bar{y} = 0$ because $\frac{n-m}{n-1} - \gamma^{T'} I^{-1} > \frac{n-m}{n-1} - I^{-1}$) and $\gamma^{-t} < 1$. Consequently, (3.19) must be satisfied with equality at a unique value

of t, which we denote by $\hat{t}_1(y, \bar{y})$, with $\hat{t}_1(y, \bar{y}) \in (0, \infty)$. This value is the first moment in time at which myopic players are indifferent between choosing A and B, thus it is exactly the switching period which we were looking for. To get an expression for $\hat{t}_1(y, \bar{y})$, require (3.19) to be satisfied with equality and rearrange the following way:

$$\hat{t}_1(y, \bar{y}) = \frac{\log(\frac{n-m}{n-1} - I^{-1}) - \log(\gamma^{-\bar{y}} \frac{n-m-1}{n-1} + \gamma^{-y} \frac{1}{n-1} - \gamma^{T'} I^{-1})}{\log(\gamma)} \tag{3.20}$$

Of course, $\hat{t}(y, \bar{y})$ can be calculated using (3.20) only if $\frac{n-m}{n-1} - I^{-1} > 0$, otherwise myopic players would always play B. The precise characterization of the switching period if case 1 is applicable is as follows:

$$\hat{t}(y, \bar{y}) = \begin{cases} \hat{t}_1(y, \bar{y}) & \text{if } \frac{n-m}{n-1} - I^{-1} > 0 \\ \infty & \text{otherwise} \end{cases} \tag{3.21}$$

Note that it is not required that $\hat{t}_1(y, \bar{y}) \leqslant T$, therefore it is possible that the planning horizon of a sophisticated player is too short to take re-coordination into account.

Case 2: $y < \hat{t} < \bar{y}$

Case 3: $\bar{y} < \hat{t} < y$

Proofs for Case 2 and Case 3 are in Appendix 3.A.1.

■

Lemma 3. $\frac{\partial \hat{t}_2(y)}{\partial y} > 1$.

Proof: see Appendix 3.A.2.

Lemma 3 implies that if $\hat{t}_2(0) < T$, it would be optimal for all sophisticated players to choose $y = 0$: increasing y by an amount of ε would increase the payoffs by εL, because of a longer delay, but would simultaneously decrease the payoffs by more than εH because of the longer switching period of myopic players.

3.3.5 Payoffs of Sophisticated Players

Proposition 3 shows how $\hat{t}(y, \bar{y})$, the switching period of myopic players, depends on sophisticated player strategies, if one player is using strategy y and all other players are using strategies $\bar{y}$. Proposition 4 will show how this specification can be used to calculate the sum of payoffs received by sophisticated players over the period that is taken into consideration.

Proposition 4. *If a sophisticated player s uses strategy $y_s = y$ and other sophisticated players use strategies $y_{-s} = \bar{y}$, total payoff received by player s over period $[0, T]$ is $\Pi(y, \bar{y})$ such that:*

$$\Pi(y,\bar{y}) = \begin{cases} \Pi_1 = yL + (T - \hat{t}_1(y,\bar{y}))H & \text{if } \hat{t}_1(y,\bar{y}) \leqslant T,\ \hat{t}_2(y) \geqslant \bar{y},\ \hat{t}_3(\bar{y}) \geqslant y \quad (3.22a) \\ \Pi_2 = yL + (T - \hat{t}_2(y))H & \text{if } \hat{t}_2(y) < \bar{y} \quad (3.22b) \\ \Pi_3 = yL + (T - y)H & \text{if } \hat{t}_3(\bar{y}) < y \quad (3.22c) \\ \Pi_4 = yL & \text{if } \hat{t}_1(y,\bar{y}) > T \quad (3.22d) \end{cases}$$

where $\hat{t}_1(y, \bar{y})$, $\hat{t}_2(y)$ and $\hat{t}_3(\bar{y})$ are specified in Proposition 3.

Proof.

The payoff function depends on the switching period of myopic players, which is determined by one of the four equations in condition (3.13). Each possibility is shown in figure 3.4. Consider panel (a), which illustrates a situation where all sophisticated players switch to A first[7], and myopic players follow later, therefore their switching time is calculated as $\hat{t}_1(y, \bar{y})$. The participation threshold is not exceeded at any time prior to $\hat{t}_1(y, \bar{y})$ and is exceeded afterwards, therefore the payoff flow of a sophisticated player is L prior to time y, 0 between time y and $\hat{t}_1(y, \bar{y})$ and H afterwards. The sum of payoffs in this case would be equal to $\Pi_1(y, \bar{y}) = yL + (T - \hat{t}_1(y, \bar{y}))H$. Panel (a), however, applies only if myopic players switch after all sophisticated ones, that is if $\hat{t}_2(y) \geqslant \bar{y}$ and $\hat{t}_3(\bar{y}) \geqslant y$, and if switching occurs prior to time T.

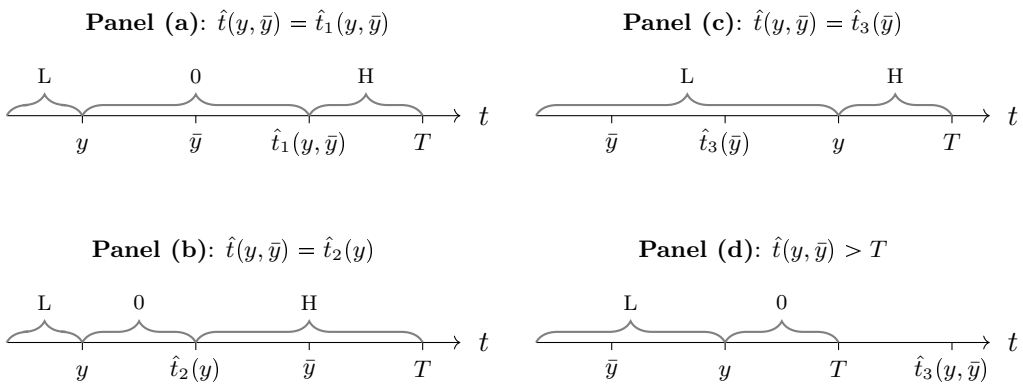


Figure 3.4: Stage game payoffs for every possible case. Panel numbering corresponds to equations in (3.22).

Another possibility is that myopic players switch after observing only one sophisticated player switching to A, a case illustrated in panel (b). Then the

[7]Panel (a) illustrates the situation with $y < \bar{y}$, but the payoff calculation for $y \geqslant \bar{y}$ would be equivalent.

sophisticated player will receive a payoff flow equal to L at any time prior to y, a flow of 0 between time y and $\hat{t}_2(y)$ and a flow of H between $\hat{t}_2(y)$ and T. The sum of payoffs in this case would be equal to $\Pi_2(y, \bar{y}) = yL + (T - \hat{t}_2(y))H$. Panel (b) applies only if $\hat{t}_2(y) < \bar{y}$.

In a similar way, $(n - m - 1)$ sophisticated players may switch first, followed by myopic players and then by a single sophisticated player, illustrated in panel (c). Sophisticated player would receive L until time y, and would receive H afterwards. The sum of payoffs would therefore be equal to $\Pi_3(y, \bar{y}) = yL + (T - y)H$. Panel (c) applies only if $\hat{t}_3(\bar{y}) < y$.

Finally, myopic players may never switch to A, as illustrated in panel (d). In this case the sophisticated player would receive L until time y, and 0 afterwards, thus the total payoff would be $\Pi_4(y, \bar{y}) = yL$.

■

3.3.6 Characterisation of Symmetric Sophisticated Player Equilibria

Payoffs for each strategy of player s and the strategies of other sophisticated players are specified in (3.22). This specification transforms a repeated game into a static game played by sophisticated players, who are able to perfectly anticipate the choice path of myopic players. To make theoretical predictions, we can use the standard solution concept for static games – a Nash equilibrium – which requires mutual best responses for each player.

Proposition 2 shows that undominated action plans for sophisticated players can be identified by a strategy that identifies a switching time. We will therefore use the definition from (3.5) to call a combination of *strategies* (y^*, y^*) a symmetric sophisticated player equilibrium if it satisfies:

$$\Pi(y^*, \bar{y}^*) \geqslant \Pi(y, \bar{y}^*), \quad \forall y \in [0, T] \quad \text{and } y^* = \bar{y}^* \tag{3.23}$$

We will look at the existence of three types of equilibria: interior solutions with $y = \bar{y} \in (0, T)$, a corner solution with $y = \bar{y} = 0$ and a corner solution with $y = \bar{y} = T$. For each type we will determine the conditions under which an equilibrium exists, and the speed of transition to an efficient state.

3.3.6.1 Interior Sophisticated Player Equilibria

In this section we will derive the existence conditions for an interior equilibrium and show how the speed of transition to the efficient equilibrium depends on the game parameters.

Proposition 5. *A combination of strategies (y^*, y^*) with $y^* \in (0, T)$ is a sophisticated player equilibrium ("interior equilibrium") if and only if conditions **I1**, **I2**, **I3** and **I4** are satisfied:*

$$\hat{t}_1(y^*, y^*) < T, \qquad \textbf{(I1)}$$

$$\frac{\log(\frac{n-m-H/L}{I^{-1}(n-1)})}{\log(\gamma)} - T' > 0, \qquad \textbf{(I2)}$$

$$\hat{t}_1(y^*, y^*) - y^* L/H \leqslant \hat{t}_2(0), \qquad \textbf{(I3)}$$

$$\hat{t}_1(y^*, y^*) - y^* L/H \leqslant T(1 - L/H), \qquad \textbf{(I4)}$$

where equilibrium strategies are calculated by

$$y^* = \frac{\log(\frac{n-m-H/L}{I^{-1}(n-1)})}{\log(\gamma)} - T'$$

Proof.

The structure of the proof is shown in figure 3.5. First we need to specify the equilibrium payoffs. If condition **I1** holds, condition (3.22a) will hold as well, from Lemma 10, therefore $\Pi(y^*, y^*) = \Pi_1(y^*, y^*)$. If **I1** does not hold, $\Pi(y^*, y^*) = \Pi_4(y^*, y^*) = y^* L$, and an interior equilibrium will not exist because there is a profitable deviation to a strategy $y = T$ that provides a payoff of TL. Condition **I1** is therefore the first necessary condition for the existence of an interior equilibrium, and we will show that it is also jointly sufficient, together with conditions **I2**, **I3** and **I4**. These proofs are given in additional lemmas. Lemma 4 shows that equilibrium payoffs exceed deviation payoffs if and only if equilibrium payoffs exceed the payoffs of two endpoints, 0 and T, and the payoffs of 'neighboring' strategies, calculated by $\Pi_1(y, y^*)$. Lemma 5, 6 and 7 derive the conditions under which there are no profitable deviations for each case.

Lemma 4.

$$\Pi_1(y^*, y^*) \geqslant \Pi(y, y^*), \forall y \in [0, T] \quad \Leftrightarrow \quad \begin{cases} \Pi_1(y^*, y^*) \geqslant \Pi_2(0, y^*) \\ \Pi_1(y^*, y^*) \geqslant \Pi_1(y, y^*), \forall y \in [y'', y'] \\ \Pi_1(y^*, y^*) \geqslant \Pi_4(T, y^*) \end{cases}$$

If $\hat{t}_3(y^*) \leqslant T$, $y' = \hat{t}_3(y^*)$, otherwise y' solves $\hat{t}_1(y', y^*) = T$. If $\hat{t}_2(0) > y^*$, $y'' = 0$, otherwise y'' solves $\hat{t}_2(y'') = y^*$.

Proof: see Appendix 3.A.2.

$$\Pi(y^*, y^*) \geqslant \Pi(y, y^*), \forall y \in [0, T] \xLeftrightarrow{\text{if not } \mathbf{I1}} \Pi_4(y^*, y^*) \geqslant \Pi_4(y, y^*), \forall y \in [0, T]$$

$$\Big\Updownarrow \text{if } \mathbf{I1}$$

$$\Pi_1(y^*, y^*) \geqslant \Pi(y, y^*), \forall y \in [0, T]$$

$$\Big\Updownarrow \text{Lemma 4}$$

$$\begin{cases} \Pi_1(y^*, y^*) \geqslant \Pi_1(y, y^*), \forall y \in (y'', y') & \xLeftrightarrow{\text{Lemma 5}} \mathbf{I2} \\ \Pi_1(y^*, y^*) \geqslant \Pi_2(0, y^*) & \xLeftrightarrow{\text{Lemma 6}} \mathbf{I3} \\ \Pi_1(y^*, y^*) \geqslant \Pi_4(T, y^*) & \xLeftrightarrow{\text{Lemma 7}} \mathbf{I4} \end{cases}$$

Figure 3.5: Structure of the proof for Proposition 5.

Lemma 5. $\Pi_1(y^*, y^*) \geqslant \Pi_1(y, y^*), \quad \forall y \in (y'', y')$, *if and only if condition* ***I2*** *is satisfied:*

$$\frac{\log(\frac{n-m-H/L}{I^{-1}(n-1)})}{\log(\gamma)} - T' > 0, \qquad (\boldsymbol{I2})$$

Proof: see Appendix 3.A.2.

Lemma 5 specifies conditions under which there are no profitable deviations to strategies in the interval $[y'', y']$. In addition, equilibrium payoffs must be higher than the payoffs from choosing $y = 0$ and $y = T$. Conditions under which there are no incentives to deviate to such strategies are specified in Lemma 6 and Lemma 7.

Lemma 6. $\Pi_1(y^*, y^*) \geqslant \Pi_2(0, y^*)$ *if and only if condition* $I3$ *is satisfied:*

$$\hat{t}_1(y^*, y^*) - y^* L/H \leqslant \hat{t}_2(0), \qquad (\boldsymbol{I3})$$

Proof: see Appendix 3.A.2.

Lemma 7. $\Pi_1(y^*, y^*) \geqslant \Pi_4(T, y^*)$ *if and only if condition* ***I4*** *is satisfied:*

$$\hat{t}_1(y^*, y^*) - y^* L/H \leqslant T(1 - L/H), \qquad (\boldsymbol{I4})$$

Proof: see Appendix 3.A.2.

Taken together, Lemmas 5, 6 and 7 prove Proposition 5. Conditions **I1**, **I2**, **I3** and **I4** are jointly sufficient because if all of them are satisfied there are no incentives to deviate to any strategy in $[0, T]$. If one of these conditions is violated, there will be a strategy in some region that exceeds the equilibrium payoff.

■

3.3.6.2 Corner Solution $y^* = 0$

In a second type of a symmetric sophisticated player equilibrium all sophisticated players switch to A at the start of the game, so that equilibrium strategies are $y^* = y^* = 0$.

Proposition 6. *A combination of strategies* $(0, 0)$ *is a sophisticated player equilibrium ("teaching equilibrium") if and only if conditions* ***T1*** *and* ***T2*** *are satisfied:*

$$\frac{n - m - H/L}{n - 1} \leqslant \gamma^{T'} I^{-1}, \quad \textit{(T1)}$$
$$\hat{t}_1(0, 0) \leqslant T(1 - L/H), \quad \textit{(T2)}$$

Proof.

$$\begin{array}{ccc}
\Pi(0,0) \geqslant \Pi(y,0), \forall y \in [0,T] & \xleftrightarrow{\text{if } \hat{t}_1(0,0)>T} & \Pi_4(0,y^*) \geqslant \Pi(y,y^*), \forall y \in [0,T] \\
\Big\Updownarrow \text{if } \hat{t}_1(0,0) \leqslant T & & \\
\Pi_1(0,0) \geqslant \Pi(y,0), \forall y \in [0,T] & & \\
\Big\Updownarrow & & \\
\left\{ \begin{array}{l} \Pi_1(0,0) \geqslant \Pi_1(y,0), \forall y \in [0,y') \\ \Pi_1(0,0) \geqslant \Pi_4(T,0) \end{array} \right. & \begin{array}{c} \xleftrightarrow{\text{Lemma 8}} \\ \xleftrightarrow{\text{Lemma 9}} \end{array} & \begin{array}{c} \textbf{T1} \\ \textbf{T2} \end{array}
\end{array}$$

Figure 3.6: Structure of the proof for Proposition 6.

Structure of the proof is shown in figure 3.6 and is similar to the proof of the interior equilibrium. The teaching equilibrium exists if (3.23) is satisfied for $y^* = 0$:

$$\Pi(0, 0) \geqslant \Pi(y, 0), \quad \forall y \in [0, T]$$

If $\hat{t}_1(0, 0) > T$, condition (3.22d) is satisfied and equilibrium payoffs are determined by $\Pi(0, 0) = \Pi_4(0, 0) = 0$, while deviation payoffs are determined by $\Pi(y, 0) = yL$. Then a teaching equilibrium would not exist because there is a profitable deviation to strategy $y = T$ that provides a payoff of TL. If $\hat{t}_1(0, 0) \leqslant T$, equilibrium payoffs are calculated by $\Pi_1(0, 0)$. Condition $\hat{t}_1(0, 0) \leqslant T$ is therefore necessary for the existence of an interior equilibrium. We do not list this condition separately because it is implied by **T2**.

Deviation payoffs are determined in a similar way to the deviation payoffs for an interior equilibrium. Payoffs for a small $y \in [0, t']$ are calculated by $\Pi_1(y, 0)$,

where y' solves $\hat{t}_1(y', 0) = T$. If the deviation is larger, that is $y \in [t', T]$, myopic player would never switch to A and deviation profits would be calculated by $\Pi_4(y, 0) = yL$. All strategies in this interval would be dominated by strategy $y = T$ that provides a payoff of TL. Overall, there are two requirements that need to be satisfied for a teaching equilibrium to exist. First, equilibrium payoffs should be higher than the payoffs from any other $y \in [0, y')$, calculated by $\Pi_1(y, 0)$. We will derive the conditions under which this requirement is satisfied in Lemma 8. Second, equilibrium payoffs should be higher than the payoff of strategy $y = T$; we will derive the conditions for this requirement in Lemma 9.

Lemma 8. $\Pi_1(0, 0) \geqslant \Pi_1(y, 0), \quad \forall y \in [0, y')$ *if and only if condition* ***T1*** *is satisfied:*

$$\frac{n - m - H/L}{n - 1} \leqslant \gamma^{T'} I^{-1} \quad \textbf{\textit{(T1)}}$$

where y' *solves* $\hat{t}_1(y', 0) = T$.

Proof: see Appendix 3.A.2.

Lemma 9. $\Pi_1(0, 0) \geqslant \Pi_4(T, 0)$ *if and only if condition* ***T2*** *is satisfied:*

$$\hat{t}_1(0, 0) \leqslant T(1 - L/H) \quad \textbf{\textit{(T2)}}$$

Proof.

Deviation payoffs are calculated from (3.22d): $\Pi_4(T, 0) = TL$. There are no incentives to deviate if

$$\Pi_1(0, 0) \geqslant \Pi_4(T, 0) \quad \Leftrightarrow \quad \hat{t}_1(0, 0) \leqslant T(1 - L/H)$$

■

If both **T1** and **T2** hold, equilibrium payoffs are calculated by $\Pi_1(0, 0)$ and there are no incentives to deviate neither to neighbouring strategies nor to strategy $y = T$. If one of these conditions is violated, there would be profitable deviation and a teaching equilibrium would not exist.

■

3.3.6.3 Corner Solution $y^* = T$

In the third type of a symmetric sophisticated player equilibrium all sophisticated players choose B for the entire duration of the game, that is $y^* = \bar{y}^* = T$.

Proposition 7. *A combination of strategies (T, T) is a sophisticated player equilibrium ("delay equilibrium") if and only if condition **D1** is satisfied:*

$$\hat{t}_2(0) \geqslant T(1 - L/H) \qquad \textbf{(D1)}$$

Proof.

$$\Pi(T,T) \geqslant \Pi(y,T), \forall y \in [0,T]$$
$$\Updownarrow$$
$$\Pi_4(T,T) \geqslant \Pi(y,T), \forall y \in [0,T]$$
$$\Updownarrow$$
$$\begin{cases} \Pi_4(T,T) \geqslant \Pi_2(0,T) & \Longleftrightarrow \quad \textbf{D1} \\ \Pi_4(T,T) \geqslant \Pi_4(y,T), \forall y \in [y',T] & \Longleftrightarrow \quad \text{Always satisfied} \end{cases}$$

Figure 3.7: Structure of the proof for Proposition 7.

If there is a symmetric equilibrium with $y^* = T$, it must hold that:

$$\Pi(T,T) \geqslant \Pi(y,T), \quad \forall y \in [0,T]$$

The structure of the proof is shown in figure 3.7. Condition (3.22d) is satisfied, therefore equilibrium payoffs are $\Pi(T,T) = \Pi_4(T,T) = TL$. Deviation payoffs $\Pi(y,T)$ are calculated either as $\Pi_4(y,T)$ if $y \in [y',T]$ or as $\Pi_2(y,T)$ if $y \in [0,y')$, where y' solves $\hat{t}_2(y') = T$. In the former case $\Pi_4(y,T) = yL$, which is less that the payoff of TL provided by strategy $y = T$, therefore the delay equilibrium would exist. In the latter case deviation payoffs are equal to:

$$\Pi(y,T) = \Pi_2(y,T) = yL + (T - \hat{t}_2(y))H$$

Lemma 3 implies that $\text{argmax}_y(\Pi_2(y,T)) = 0$, that is the most profitable deviation is to strategy $y = 0$. There will be no incentives to deviate to this strategy if the following holds:

$$\begin{aligned} \Pi_4(T,T) \geqslant \Pi_2(y,T) \quad & \Leftrightarrow \\ TL \geqslant (T - \hat{t}_2(0))H \quad & \Leftrightarrow \\ \hat{t}_2(0) \geqslant T(1 - L/H) & \end{aligned}$$

If this condition is satisfied, there will be no incentives to deviate to $y = 0$ and there would be no other profitable deviations, therefore a delay equilibrium would exist. If this condition is not satisfied, payoffs could be increased by choosing strategy $y = 0$.

■

Table 3.1: Summary of the types of symmetric Nash equilibria that may exist, speed of transition to the efficient state and the conditions that need to be satisfied for the particular type of equilibrium to exist.

Equilibrium	**Teaching**	**Interior**	**Delay**
Equilibrium strategy	$y^* = 0$	$y^* = \frac{\log(\frac{n-m-H/L}{I^{-1}(n-1)})}{\log(\gamma)} - T'$	$y^* = T$
Speed of transition	$\hat{t}(y^*, y^*) = \hat{t}_1(0,0)$	$\hat{t}(y^*, y^*) = \hat{t}_1(y^*, y^*)$	$\hat{t}(y^*, y^*) > T$
Equilibrium payoffs	$\Pi(0,0) = (T - \hat{t}_1(0,0))H$	$\Pi(y^*, y^*) = y^*L + (T - \hat{t}_1(y^*, y^*))H$	$\Pi(T,T) = TL$
Existence conditions			
No deviation to neighbouring strategies	**T1:** $\frac{n-m-H/L}{n-1} \leqslant \gamma^{T'} I^{-1}$	**I2:** $y^* > 0$	–
No deviation to $y = 0$	–	**I3:** $\hat{t}_1(y^*, y^*) - y^*L/H \leqslant \hat{t}_2(0)$	**D1:** $T(1-L/H) \leqslant \hat{t}_2(0)$
No deviation to $y = T$	**T2:** $\hat{t}_1(0,0) \leqslant T(1 - L/H)$	**I1:** $\hat{t}_1(y^*, y^*) < T$ **I4:** $\hat{t}_1(y^*, y^*) - y^*L/H \leqslant T(1 - L/H)$	–

3.4 Summary and Comparative Statics

Overall, three types of symmetric Nash equilibria can exist: in a "teaching" equilibrium all sophisticated players play A for the entire duration of the game and myopic players switch to A at some time $\hat{t}_1(0,0)$; in a "delay" equilibrium all sophisticated players choose B for the entire duration of the game, and myopic players never switch to A; in an interior equilibrium sophisticated players start by playing B and switch to A at time y^* while myopic players switch to A at time $\hat{t}_1(y^*, y^*)$. Table 3.1 summarizes all the conditions that need to be satisfied for each type of equilibrium to exist. Depending on the combination of parameters, it is possible that multiple equilibria will exist at the same time or that no symmetric equilibrium will exist.

We would like to make theoretical predictions about how the path of play depends on the game parameters, but precise predictions cannot be made due to the multiplicity of equilibria. Therefore we separately investigate how the factors of interest affect the existence conditions of each type of equilibria and the speed of transition to an efficient state. The factors that we consider are the length of planning horizon of sophisticated players (T), the number of myopic players (m) and the strength of initial lock-in (T').

3.4.1 Planning Horizon of the Sophisticated Players

The first parameter of interest is T, the length of the planning horizon for sophisticated players.

Proposition 8. *If sophisticated players have a longer planning horizon, then:*

1. *The speed of transition in any equilibrium is not affected.*
2. *Teaching equilibrium exists for a larger set of values of other parameters.*
3. *Interior equilibrium exists for a larger set of values of other parameters.*
4. *Delay equilibrium exists for a smaller set of values of other parameters.*

Proof.

Part 1 follows from the definition of the switching period, which depends only on myopic players who do not take future payoffs into account. For part 2, note that only condition **T2** depends on the planing horizon, and **T2** is satisfied for a larger set of parameters when T is higher. For part 3, note that conditions **I2** and **I4** depend on the length of the planning horizon, and both are satisfied for a larger set of parameters when T is larger. Part 4 holds because condition **D1** is satisfied for a smaller set of parameters when T is larger.

■

3.4.2 Player Composition

The second variable of interest is m, the number of myopic players, which reflects a different aspect of sophistication than the length of the planning horizon. Instead of making sophisticated players more sophisticated, we look at the effect of replacing some myopic players with sophisticated ones, while keeping the total number of players constant.

Proposition 9. *If there are more sophisticated players, then:*

1. *Transition is faster in the interior and in the teaching equilibria.*
2. *The effect on the existence of a teaching equilibrium or an interior equilibrium is ambiguous:*
   (a) *there are more incentives to deviate to neighboring action plans*
   (b) *there are less incentives to never choose A.*
3. *There is no change in the existence of the delay equilibrium.*

Proof.

This proof as well as other proofs on comparative statics rely on additional lemmas presented in Appendix 3.A.3. For part 1, see Lemmas 11 and 12. To see part 2 for the teaching equilibrium, note that both condition **T1** and condition **T3** depend on player composition. A smaller number of myopic players leads to **T1** being satisfied for a smaller set of values of other parameters. On the other hand, a smaller number of myopic players makes condition **T2** satisfied for a larger set of parameters because $\hat{t}_1(0,0)$ is increasing in m (see Lemma 11). For the interior equilibrium, all four conditions depend on the number of sophisticated players. Incentives to deviate to neighbouring strategies are determined by condition **I2**, which is satisfied for a smaller set of parameters when there are more sophisticated players. To see it, notice that $\frac{\partial y^*}{\partial m} > 0$ (Lemma 12), therefore as m decreases so does y^*, therefore **I2** is less likely to be satisfied. Incentives to deviate to corner solutions are determined by conditions **I1**, **I3** and **I4**, all of which are satisfied for a larger set of parameters when there are more sophisticated players. Conditions **I3** and **I4** are satisfied for a larger set of parameters because $\frac{\partial \hat{t}_1(y^*,y^*)}{\partial m} > \frac{\partial y^*}{\partial m} > \frac{\partial y^*}{\partial m} L/H$ (see Lemma 14). Condition **I1** is also satisfied for a larger set of parameters because $\frac{\partial \hat{t}_1(y^*,y^*)}{\partial m} > 0$, from Lemma 12. Part 3 holds because outcomes in the delay equilibrium are not affected by the number of myopic players. ■

The finding that an increase in the number of sophisticated players can reduce the incentives to use strategic teaching may sound counterintuitive, but it is a result of decreased delay costs as the number of teaching players grows. When the number of sophisticated players is large and all of them are choosing A in the teaching equilibrium, the decision of a single sophisticated player to delay teaching has only a small negative effect on the transition period, making free-riding an attractive alternative that could lead to a break-down of a teaching equilibrium. But if a teaching equilibrium does exist, a larger number of sophisticated players would make transition faster.

3.4.3 Length of the History of Inefficient Coordination

The third factor that we look at is the strength of the initial lock-in to an inefficient state, measured by the length of history of inefficient coordination, T'.

Proposition 10. *If the history of inefficient coordination is longer, then:*

1. *Transition is slower in the teaching equilibrium but faster in an interior equilibrium.*
2. *Teaching equilibrium exists for a smaller set of parameter values*
3. *The effect on the existence of an interior equilibrium is ambiguous*
4. *Delay equilibrium exists for a smaller set of parameter values.*

Proof.

Part 1 holds because the derivative of $\hat{t}_1(0,0)$ with respect to T' is positive while the derivative of $\hat{t}_1(y^*, y^*)$ is negative, as shown in Lemma 11 and Lemma 12. For part 2, parameter T' affects conditions **T1** and **T2**. An increase in T' leads to **T1** being satisfied for a smaller set of parameter values, because $\gamma^{T'}$ goes down. Condition **T2** is also less likely to be satisfied because of an increase in $\hat{t}_1(0,0)$. For part 3, notice that an increase in T' satisfies conditions **I1**, **I3** and **I4** for a larger set of parameter values, but satisfies condition **I2** for a smaller set of parameter values. Condition **I1** is satisfied for a larger set of parameter values because $\frac{\partial \hat{t}_1(y^*,y^*)}{\partial T'} < 0$. Conditions **I3** and **I4** are also satisfied for a larger set of parameter values because $\frac{\partial \hat{t}_1(y^*,y^*)}{\partial T'} = \frac{\partial y^*}{\partial T'} < \frac{\partial y^*}{\partial T'} L/H$ and $\frac{\hat{t}_2(0)}{\partial T'} > 0$, from Lemma 11, 12 and 13. Condition **I2** is satisfied for a smaller set of parameters because $\frac{\partial y^*}{\partial T'} < 0$, from Lemma 14.

■

Part 1 of Proposition 10 states that the history of inefficient coordination affects the transition speed in opposite ways in the teaching and in the interior equilibrium. The opposite sign of this effect is a result of changes of the equilibrium strategy in the interior equilibrium. If the equilibrium strategy in an interior equilibrium was held constant, a longer history of inefficient coordination would lead to a slower transition. However, to offset an increased history of inefficient coordination, in an equilibrium sophisticated players have to start teaching earlier. Lemma 12 shows that the latter effect is even stronger than the former.

3.5 Conclusion

In this chapter we present a model that combines the standard game-theoretic view of players who are strategic and farsighted with the idea that players best respond to beliefs that are formed using the observed history of play. We assume two types of players: myopic players make choices based on observed history of play while sophisticated players have correct beliefs about the actions of all other players, plan ahead and choose actions that maximize the sum of payoff flows. To make predictions in this modified game we propose a new solution concept based on a Nash equilibrium between sophisticated players who take the learning process of the myopic players into account.

This solution concept is applied to a critical mass coordination game in which play has converged to an inefficient equilibrium. The construction of a sophisticated player equilibrium involves several steps. Proposition 1 shows that myopic players will choose the efficient action if their beliefs exceed a certain threshold. Furthermore, in the sophisticated player equilibrium myopic players will switch from an inefficient to the efficient action at most once. The single switch and the assumption that there are sufficiently many myopic players means that the efficient state is absorbing, therefore the switching time is the only information needed for sophisticated players to calculate their payoffs. Proposition 3 shows exactly how the switching time of myopic players can be calculated if beliefs were formed using weighted fictitious play. The switching time depends on the strategies taken by sophisticated players, which could prescribe many switches from one action to the other. The task of finding the switching time of the myopic players is therefore greatly simplified by Proposition 2, which shows that only the sophisticated player strategies prescribing at most one switch from the inefficient to the efficient action survive the elimination of strictly dominated strategies, allowing a strategy to be identified by the switching time.

The ability to anticipate the speed of a transition allows sophisticated players to calculate how their payoffs depend on their own strategies and on the strategies chosen by other sophisticated players. The mapping from strategies to payoffs specified in Proposition 4 is used to identify strategy profiles in which all sophisticated players are best responding to each other. Three types of symmetric equilibria are possible: sophisticated players may play the efficient action right away, they may switch to the efficient action later or they may never switch. In the first two cases myopic players eventually start playing the efficient action, while in the third case all players choose the inefficient action. Which types of equilibria exist and how long the transition to an efficient state takes depends on the game parameters, as specified in Propositions 5, 6 and 7. Finally, Propositions 8, 9 and 10 show how these existence conditions depend on the history of inefficient coordination, length of the planning horizon of sophisticated players and the player composition. As the planing horizon of sophisticated players increases, teaching and interior equilibria exist for a larger set of parameters, while the delay equilibrium exists for a smaller set of parameters. A larger number of sophisticated players leads to faster transition to the efficient state in an interior or in a delay equilibrium, but the effect on the existence conditions is ambiguous: there are more incentives to deviate to neighboring strategies, but less incentives to deviate to corner solutions. Finally, we show that a longer history of observed inefficient coordination leads to a slower transition in a teaching equilibrium and to a smaller set of parameters under which a teaching equilibrium exists, while the set of parameters under which the delay equilibrium exists is larger. On the other hand, the transition to an efficient state in an interior equilibrium is faster, because a longer history of inefficient coordination forces sophisticated players to start teaching earlier.

The problem that motivated this chapter was the lack of a suitable theoretical model that could be used to make predictions in a game in which inefficient conventions have already been established. A small change in the assumptions – instead of assuming all players to be farsighted we assume that some players are learning from history – leads to a large differences in theoretical predictions. Not only can the new model be used to model inefficient conventions through the beliefs of myopic players, but it also reduces the set of predictions to only three types of equilibria, in contrast to almost limitless predictions made by the standard Nash equilibrium.

Appendix 3

3.A Proofs

3.A.1 Case 2 and Case 3 of Proposition 3

Case 2: $y < \hat{t} < \bar{y}$

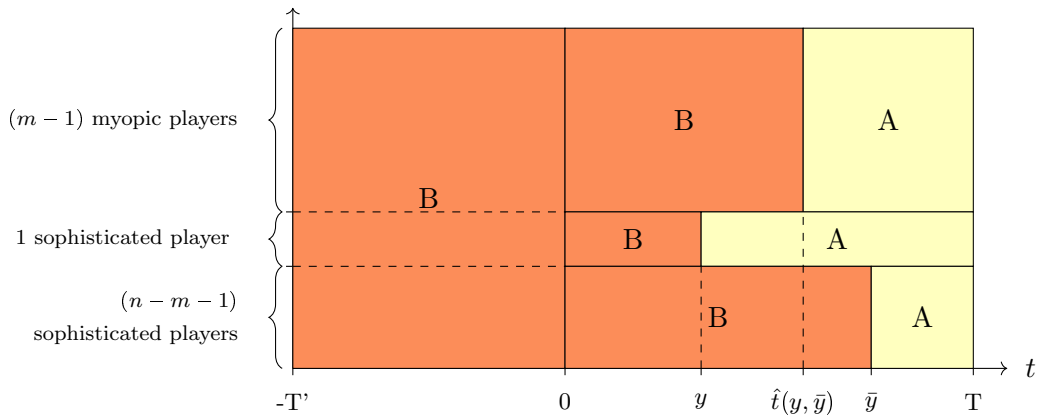


Figure 3.8: Illustration of the second case, where $y < \hat{t}(y, \bar{y}) \leqslant \bar{y}$. The height of the figure shows the fraction of players choosing action A or action B, the width shows the passage of time. The first sophisticated player switches from B to A in period y, other (n-m-1) sophisticated players switch in period $\bar{y}$ and the myopic players switch in period $\hat{t}$.

The second possibility is that $\hat{t}(y, \bar{y}) \leqslant \bar{y}$, that is myopic players switch to A earlier than $(n-m-1)$ sophisticated players. In this case the actual value of $\bar{y}$ will have no influence on the switching period of myopic players, as they will never observe any of the $(n-m-1)$ sophisticated players choosing A. Therefore the switching period will be a function only of the strategy chosen by a single sophisticated player. At time $t \in (y, \hat{t}]$ beliefs of a myopic player i are $x_i(t)$:

$$x_i(t) = \frac{\int_{k=0}^{t-y} \gamma^k (\frac{1}{n-1}) \mathrm{d}k}{\int_{k=0}^{t+T'} \gamma^k \mathrm{d}k} = \\ = \frac{(\gamma^{t-y} - 1)(\frac{1}{n-1})}{\gamma^{t+T'} - 1}$$

Player i will choose A in t if:

$$x_i(t) \geqslant I^{-1} \quad \Leftrightarrow$$

$$\gamma^{t+T'}(\gamma^{-y-T'} \frac{1}{n-1} - I^{-1}) \leqslant \frac{1}{n-1} - I^{-1} \tag{3.24}$$

If $\frac{1}{n-1} - I^{-1} \leqslant 0$, equation (3.24) is never satisfied. To see this, notice the following relationship that contradicts (3.24):

$$\gamma^{t+T'}(\gamma^{-y-T'}\frac{1}{n-1} - I^{-1}) > \gamma^{t+T'}(\frac{1}{n-1} - I^{-1}) \geqslant \frac{1}{n-1} - I^{-1}$$

The latter equation holds because $\gamma^{-y-T'} > 1$, $\gamma^{t+T'} < 1$ and $\frac{1}{n-1} - I^{-1} \leqslant 0$.

Alternatively, if $\frac{1}{n-1} - I^{-1} > 0$, (3.24) will be satisfied with equality at time $\hat{t}_2(y) \in (0, \infty)$ that satisfies:

$$\gamma^{-\hat{t}_2(y)} = \frac{\gamma^{-y}\frac{1}{n-1} - \gamma^{T'}I^{-1}}{\frac{1}{n-1} - I^{-1}} \quad \Leftrightarrow$$

$$\hat{t}_2(y) = \frac{\log(\frac{1}{n-1} - I^{-1}) - \log(\gamma^{-y}\frac{1}{n-1} - \gamma^{T'}I^{-1})}{\log(\gamma)} \tag{3.25}$$

$\hat{t}(y, \bar{y})$ can be calculated using (3.25) only if $\frac{1}{n-1} - I^{-1} > 0$, otherwise myopic players would never switch from A to B. The switching period if case 2 applies can be expressed as follows:

$$\hat{t}(y, \bar{y}) = \begin{cases} \hat{t}_2(y) & \text{if } \frac{1}{n-1} - I^{-1} > 0 \\ \infty & \text{otherwise} \end{cases} \tag{3.26}$$

Case 3: $\bar{y} < \hat{t} < y$

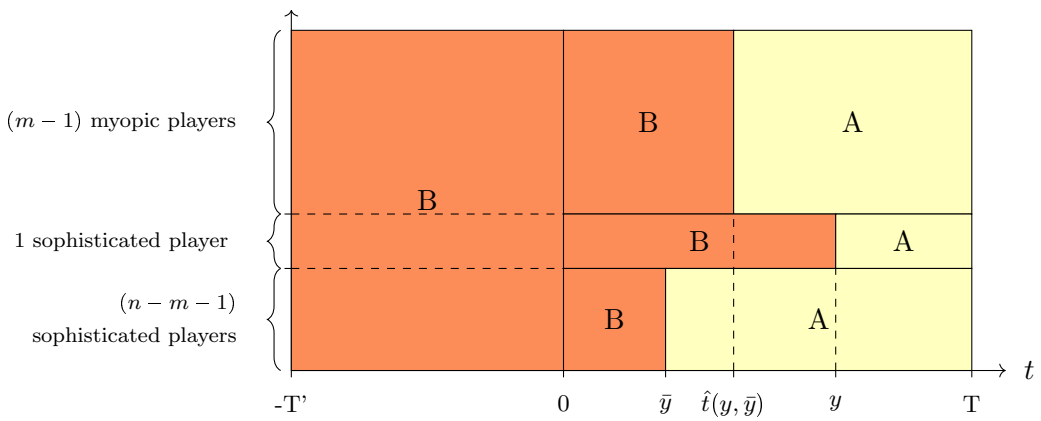


Figure 3.9: Illustration of the third case, where $\bar{y} < \hat{t}(y, \bar{y}) \leqslant y$. Height of the figure shows a fraction of players choosing action A or action B, the width shows the passage of time. The first sophisticated player switches from B to A in period y, other $(n-m-1)$ sophisticated players switch in period $\bar{y}$ and myopic players switch at time $\hat{t}(y, \bar{y})$.

The third possibility is that $\bar{y} < \hat{t}(y, \bar{y}) \leqslant y$, that is at first $(n-m-1)$ sophisticated players switch to A, then m myopic players switch and the last sophisticated player may switch some time after the myopic ones. In this case the switching time is a function only of $\bar{y}$. At time $t \in (\bar{y}, \hat{t}(y, \bar{y})]$ beliefs of a

myopic player i are $x_i(t)$:

$$x_i(t) = \frac{\int_{k=0}^{t-\bar{y}} \gamma^k (\frac{n-m-1}{n-1}) \mathrm{d}k}{\int_{k=0}^{t+T'} \gamma^k \mathrm{d}k} = \\ = \frac{(\gamma^{t-\bar{y}} - 1)(\frac{n-m-1}{n-1})}{\gamma^{t+T'} - 1}$$

Player i will choose A in t if:

$$x_i(t) \geqslant I^{-1} \Leftrightarrow$$

$$\gamma^{t+T'}(\gamma^{-\bar{y}-T'} \frac{n-m-1}{n-1} - I^{-1}) \leqslant \frac{n-m-1}{n-1} - I^{-1} \tag{3.27}$$

If $\frac{n-m-1}{n-1} - I^{-1} \leqslant 0$, condition (3.27) is never satisfied. To see this, notice the following relationship that contradicts (3.27):

$$\gamma^{t+T'}(\gamma^{-\bar{y}-T'} \frac{n-m-1}{n-1} - I^{-1}) > \gamma^{t+T'}(\frac{n-m-1}{n-1} - I^{-1}) \geqslant \frac{n-m-1}{n-1} - I^{-1}$$

The latter conditions holds because $\gamma^{-\bar{y}-T'} > 1$, $\gamma^{t+T'} < 1$ and $\frac{n-m-1}{n-1} - I^{-1} \leqslant 0$. Therefore if $\frac{n-m-1}{n-1} - I^{-1} \leqslant 0$, equation (3.27) is never satisfied and myopic players would choose B at any time t.

Alternatively, if $\frac{n-m-1}{n-1} - I^{-1} > 0$, (3.27) will be satisfied with equality at time $\hat{t}_3(y) \in (0, \infty)$ that satisfies:

$$\gamma^{-\hat{t}_3(\bar{y})} = \frac{\gamma^{-\bar{y}} \frac{n-m-1}{n-1} - \gamma^{T'} I^{-1}}{\frac{n-m-1}{n-1} - I^{-1}} \Leftrightarrow \tag{3.28}$$

$$\hat{t}_3(\bar{y}) = \frac{\log(\frac{n-m-1}{n-1} - I^{-1}) - \log(\gamma^{-\bar{y}} \frac{n-m-1}{n-1} - \gamma^{T'} I^{-1})}{\log(\gamma)} \tag{3.29}$$

$\hat{t}(y, \bar{y})$ can be calculated using (3.29) only if $\frac{n-m-1}{n-1} - I^{-1} > 0$. Therefore, the switching period if case 3 applies can be expressed as follows:

$$\hat{t}(y, \bar{y}) = \begin{cases} \hat{t}_3(\bar{y}) & \text{if } \frac{n-m-1}{n-1} - I^{-1} > 0 \\ \infty & \text{otherwise} \end{cases} \tag{3.30}$$

3.A.2 Proof of Lemmas

Lemma 1: *If two action plans of the sophisticated player prescribe the same action, the payoff flow is higher for the action plan with which myopic player beliefs are higher:*

$$\pi[a_s'(t), a_{-s}(t) \times a_i(t, a_s' \times a_{-s})] \geqslant \pi[a_s(t), a_{-s}(t) \times a_i(t, a_s \times a_{-s})]$$
$$\text{if} \quad x(t)' \geqslant x(t) \quad \text{and} \quad a_s'(t) = a_s(t)$$

where $x(t)'$ is the belief held by myopic players if sophisticated player uses action plan a_s' and $x(t)$ is the belief if sophisticated player uses action plan a_s.

Proof:

Consider two action plans a_s and a_s' that prescribe the same action at time t, but prescribe different actions prior to time t so that myopic players would hold higher beliefs following the history generated by a_s'.

From equation (3.7), $a_i(t, a_s \times a_{-s})$ is weakly increasing in beliefs $x_i(t)$, therefore:

$$a_i(t, a_s' \times a_{-s}) \geqslant a_i(t, a_s \times a_{-s})$$

Since we hold the action plans of other strategic players constant, a higher tendency to choose A by myopic players increases the total number of other players who choose A at time t. Because H > 0 and M $\geqslant 0$, equation (3.6) implies that payoffs are weakly increasing in the number of other players choosing A, therefore the payoff generated by a_s' must be at least as high as the payoff generated by a_s:

$$\pi[a_s'(t), a_{-s}(t) \times a_i(t, a_s' \times a_{-s})] \geqslant \pi[a_s(t), a_{-s}(t) \times a_i(t, a_s \times a_{-s})]$$

■

Lemma 2: *All action plan profiles for sophisticated players with which myopic players switch from A to B are strictly dominated:*

$$AB_M \cap U_s = \varnothing$$

Proof.

Suppose that $\times_{s \in S} a_s \in AB_M$, then there are two points in time t_1 and t_2 with $t_1 < t_2$ such that myopic players choose A at time t_1 and B at time t_2 . Find the first switching period $t_s \in (t_1, t_2]$ such that A is chosen in the interval $[t_1, t_s)$, but B is chosen at time t_s. Since all myopic players share the same history, the value of t_s will be the same for each myopic player so no myopic player will choose A in the interval $[t_1, t_s)$. If a myopic player observed all other sophisticated players choosing A in the interval $[t_1, t_s)$, the fictitious play rule would imply that $x_i(t_s) \geqslant x_i(t_1)$ therefore if A was optimal at time t_1 it will also be optimal at time t_s, contradicting the definition of t_s. Therefore if a myopic player chooses B at time t_s, at least one sophisticated player must be

choosing B in the interval $[t_1, t_s)$, that is $a_s(t) = 0$ for some $s \in S$ and $t \in [t_1, t_s)$. Denote the action plan of this sophisticated player by $\widetilde{a}_s$. We will show that $\widetilde{a}_s$ is dominated by an action plan a'_s that prescribes A in the entire interval $[t_1, t_s)$ and is otherwise the same as $\widetilde{a}_s$. First, the sum of payoff flows generated by a'_s in the interval $[t_1, t_s)$ is strictly higher than that generated by $\widetilde{a}_s$ because all myopic players are choosing A in this interval, and therefore Assumption 2 implies that the threshold will be exceeded. Second, payoffs generated in the interval $(t_s, T]$ will be equal or higher than those of $\widetilde{a}_s$ because myopic players will hold higher beliefs if a'_s is chosen (due to more A choices being observed) and consequently Lemma 1 implies that higher beliefs will lead to weakly higher payoffs for the sophisticated player at any time $t > t_s$.

■

Lemma 3: $\frac{\partial \hat{t}_2(y)}{\partial y} > 1$.

Proof.

Use the definition of $\hat{t}_2(y)$ from equation (3.25):

$$\hat{t}_2(y) = \frac{\log(\frac{1}{n-1} - I^{-1}) - \log(\gamma^{-y}\frac{1}{n-1} - \gamma^{T'}I^{-1})}{\log(\gamma)}$$

The partial derivative is calculated as follows:

$$\frac{\partial \hat{t}_2(y)}{\partial y} = \frac{1}{-\log(\gamma)} \times \frac{1}{\gamma^{-y}\frac{1}{n-1} - \gamma^{T'}I^{-1}} \times \gamma^{-y}\frac{-1}{n-1}\log(\gamma) =$$
$$= \frac{\gamma^{-y}\frac{1}{n-1}}{\gamma^{-y}\frac{1}{n-1} - \gamma^{T'}I^{-1}} > 1$$

■

Lemma 4:

$$\Pi_1(y^*, y^*) \geqslant \Pi(y, y^*), \forall y \in [0, T] \quad \Leftrightarrow \quad \begin{cases} \Pi_1(y^*, y^*) \geqslant \Pi_2(0, y^*) \\ \Pi_1(y^*, y^*) \geqslant \Pi_1(y, y^*), \forall y \in [y'', y'] \\ \Pi_1(y^*, y^*) \geqslant \Pi_4(T, y^*) \end{cases}$$

If $\hat{t}_3(y^*) \leqslant T$, $y' = \hat{t}_3(y^*)$, otherwise y' solves $\hat{t}_1(y', y^*) = T$. If $\hat{t}_2(0) > y^*$, $y'' = 0$, otherwise y'' solves $\hat{t}_2(y'') = y^*$.

Proof. To specify the deviation payoff, $\Pi(y, y^*)$, we will first look at deviations upwards ($y > y^*$) and then at deviations downwards ($y < y^*$). First, consider a deviation upwards to a strategy $y = y_D > y^*$. The calculation of payoff $\Pi(y_D, y^*)$ depends on the size of the deviation: if y_D is sufficiently small, the payoff is determined by $\Pi(y_D, y^*) = \Pi_1(y_D, y^*)$, but if y is large, myopic players may switch to A prior to y (see an illustration in figure 3.10, panel a), or myopic players may never switch to A (figure 3.10, panel b). The first

option is possible only if the myopic players switch to A without ever observing player s choose A, that is if $\hat{t}_3(y^*) < T$. Then the deviation payoffs for an action plan $y_D \in (\hat{t}_3(y^*), T]$ are calculated by $\Pi_3(y_D, y^*)$. But $\Pi_3(y, y^*)$ is decreasing in y, thus any strategy in this interval would be strictly dominated by strategy $y = \hat{t}_3(y^*)$. In figure 3.10 we indicate dominance with an arrow pointing towards the dominant strategy. Checking for profitable deviations upwards therefore only requires checking for potential deviations in the interval $(y^*, \hat{t}_3(y^*)]$. Also note that $y_D \leqslant \hat{t}_3(\bar{y})$ together with condition **I1** imply that deviation payoffs for strategies $y_D \in (0, \hat{t}_3(\bar{y}))$ are equal to $\Pi_1(y_D, y^*)$.

The second possibility is that $\hat{t}_3(y^*) \geqslant T$, so that myopic players do not switch prior to T if they observe only $n-m-1$ sophisticated players switching at y^* (see figure 3.10, panel b). Then because $\hat{t}_1(T, y^*) = \hat{t}_3(y^*) > T$, $\hat{t}_1(y^*, y^*) < T$ (from condition **I1**) and $\hat{t}_1(\cdot, y^*)$ is continuous, there must be a number $y' \in (y^*, T)$ such that $\hat{t}_1(y', y^*) = T$. If $y_D \in (y^*, y']$, (3.22a) is satisfied and $\Pi(y_D, y^*) = \Pi_1(y_D, y^*)$, because $\hat{t}_1(y, y^*) \leqslant T$, $\hat{t}_2(y_D) > y_D > y^*$ and $\hat{t}_3(y^*) > T > y^*$. The payoff from any $y_D > y'$ is determined by $\Pi_4(y, y^*) = yL$, and thus all strategies $y_D \in (y', T]$ are dominated by $y_D = T$. Overall, to check for the existence of an interior equilibrium it is sufficient to compare equilibrium payoffs to the payoffs from $y_D \in (y^*, y') \cup T$.

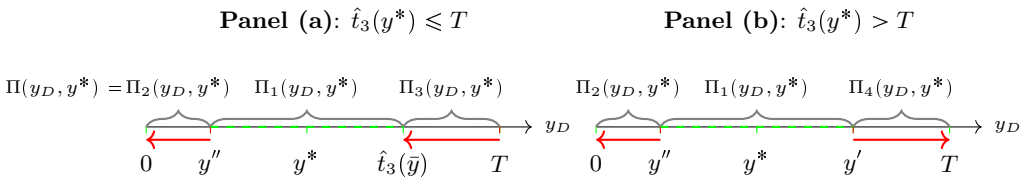


Figure 3.10: Calculation of deviation payoffs, $\Pi(y_D, y^*)$ for every possible value of y_D. Green dashed line and green ticks mark undominated strategies. Red arrows mark dominated strategies and the arrow points to the dominant strategy.

Now consider a possible deviation downwards to $y_D < y^*$. If y_D is only slightly below y^*, the switching period is $\hat{t}_1(y_D, y^*)$ and the deviation payoffs are $\Pi_1(y_D, y^*)$. But if y_D is low enough, myopic players may switch to A prior to y^*, at time $\hat{t}_2(y_D)$. If this does not happen, that is if $\hat{t}_2(0) > y^*$, payoffs from all deviations downwards are calculated by $\Pi_1(y_D, y^*)$. Otherwise, if $\hat{t}_2(0) \leqslant y^*$, there will be some value y'' that satisfies $\hat{t}_2(y'') = y^*$. For any y below this value, payoffs will be determined by $\Pi_2(y, y^*)$. From Lemma 3, any $y \in (0, y'')$ is dominated by $y = 0$, therefore to check if there are any profitable deviations downwards it is necessary to compare equilibrium payoffs to payoffs from strategies $y_D \in (y'', y^*) \cup 0$.

■

Lemma 5: $\Pi_1(y^*, y^*) \geqslant \Pi_1(y, y^*), \quad \forall y \in (y'', y')$, if and only if condition **I2** is satisfied:

$$\frac{\log(\frac{n-m-H/L}{I^{-1}(n-1)})}{\log(\gamma)} - T' > 0, \qquad (\mathbf{I2})$$

Proof.

We will calculate the first derivative of the profit function and determine under what conditions the derivative at the equilibrium point is equal to 0 and the second derivative is is non-positive, which ensures that the equilibrium is a maximum point and there are no incentives to deviate to strategies in the nearest neighbourhood. Instead of taking the derivative of the profit function, we will first transform it by applying a strictly increasing function $-\gamma^{(\cdot/H)}$, which preserves the sign of the derivative when $\gamma \in (0, 1)$. The transformed payoff function is calculated as follows:

$$\begin{aligned}
-\gamma^{\Pi(y,y^*)/H} &= -\gamma^{yL/H+T}\gamma^{-\hat{t}_1(y,y^*)} = \\
&= \frac{1}{I^{-1} - \frac{n-m}{n-1}}(\gamma^{y(L/H-1)+T}\frac{1}{n-1} + \gamma^{yL/H+T-y^*}\frac{n-m-1}{n-1} - \gamma^{yL/H+T+T'}I^{-1}) = \\
&= \frac{\gamma^{yL/H+T}}{I^{-1} - \frac{n-m}{n-1}}(\gamma^{-y}\frac{1}{n-1} + \gamma^{-y^*}\frac{n-m-1}{n-1} - \gamma^{T'}I^{-1}) \qquad (3.31)
\end{aligned}$$

where $\gamma^{\hat{t}_1(y,y_{-i})}$ has been substituted from (3.19). Differentiate the transformed profit function in (3.31) with respect to y to get

$$\begin{aligned}
\frac{\partial - \gamma^{\Pi(y,y^*)/H}}{\partial y} &= \frac{\log(\gamma)}{I^{-1} - \frac{n-m}{n-1}} \times (\gamma^{yL/H+T-y}\frac{1}{n-1}(L/H-1) + \\
&+ \gamma^{yL/H+T-y^*}\frac{n-m-1}{n-1}L/H - \gamma^{yL/H+T+T'}I^{-1}L/H) = \\
&= \frac{\log(\gamma)\gamma^{yL/H+T}}{I^{-1} - \frac{n-m}{n-1}}\left(\gamma^{-y}\frac{1}{n-1}(L/H-1) + \gamma^{-y^*}\frac{n-m-1}{n-1}L/H - \gamma^{T'}I^{-1}L/H\right) \qquad (3.32)
\end{aligned}$$

The first derivative is non-negative if:

$$\begin{aligned}
\frac{\partial - \gamma^{\Pi(y,y^*)/H}}{\partial y} \geqslant 0 \quad &\Leftrightarrow \\
\gamma^{-y}\frac{L/H-1}{n-1} + \gamma^{-y^*}\frac{n-m-1}{n-1}L/H - \gamma^{T'}I^{-1}L/H \geqslant 0 \quad &\Leftrightarrow \\
\gamma^{-y} \leqslant \frac{\gamma^{-y^*}(\frac{n-m-1}{n-1}) - \gamma^{T'}I^{-1}}{\frac{H/L-1}{n-1}} \qquad &(3.33)
\end{aligned}$$

The first derivative at point $y = y^*$ is non-negative if:

$$\left.\frac{\partial - \gamma^{\Pi(y,y^*)/H}}{\partial y}\right|_{y=y^*} \geqslant 0 \quad \Leftrightarrow$$

$$\gamma^{-y^*}\frac{H/L-1}{n-1} \leqslant \gamma^{-y^*}(\frac{n-m-1}{n-1}) - \gamma^{T'}I^{-1} \quad \Leftrightarrow$$

$$\gamma^{T'+y^*} \leqslant \frac{n-m-H/L}{I^{-1}(n-1)} \quad \Leftrightarrow$$

$$y^* \geqslant \frac{\log(\frac{n-m-H/L}{I^{-1}(n-1)})}{\log(\gamma)} - T' \tag{3.34}$$

The derivative is equal to 0 only if y^* satisfies (3.34) with equality:

$$y^* = \frac{\log(\frac{n-m-H/L}{I^{-1}(n-1)})}{\log(\gamma)} - T' \tag{3.35}$$

There will be at most one y^* that satisfies (3.35) for any given set of parameters, therefore there can be at most one interior equilibrium in a given game, and the equilibrium strategy is determined by equation (3.35). A necessary condition for the existence of an interior equilibrium is $0 < y^* < T$. But note that condition **I1** from Proposition 5 implies that $y^* < T$ because $y^* < \hat{t}_1(y^*, y^*)$, therefore the only separate condition is that $y^* > 0$.

Condition I2:

$$\frac{\log(\frac{n-m-H/L}{I^{-1}(n-1)})}{\log(\gamma)} - T' > 0$$

The second derivative is obtained by differentiating (3.32) with respect to y:

$$\frac{\partial^2 - \gamma^{\Pi(y,y^*)/H}}{\partial y^2} = \frac{\log(\gamma)^2\gamma^{yL/H+T}}{I^{-1} - \frac{n-m}{n-1}} \times (\gamma^{-y}\frac{1}{n-1}(L/H-1)^2 + \\ + \gamma^{-y^*}\frac{n-m-1}{n-1}(L/H)^2 - \gamma^{T'}I^{-1}(L/H)^2) \tag{3.36}$$

The second order derivative is negative if:

$$\frac{\partial^2 - \gamma^{\Pi(y,y^*)/H}}{\partial y^2} < 0 \quad \Leftrightarrow$$

$$\gamma^{-y}\frac{1}{n-1}(L/H-1)^2 + \gamma^{-y^*}\frac{n-m-1}{n-1}(L/H)^2 - \gamma^{T'}I^{-1}(L/H)^2 > 0 \tag{3.37}$$

If condition **I2** is satisfied, the expression of y^* in (3.35) can be used to rewrite (3.37) as follows:

$$\gamma^{y^*-y} > \frac{L}{L-H} \tag{3.38}$$

Because $L < H$ and $\gamma \in (0,1)$, condition (3.38) is satisfied for all y. The first order condition is therefore both necessary and sufficient for $y = y^*$ to be a local maximum point. Moreover, equation (3.38) states that the second derivative is negative not only at $y = y^*$, but also for any other value of y. Since the first derivative is equal to 0 at point $y = y^*$, and it is decreasing at all y, the payoff function must be increasing at any point $y < y^*$ and decreasing at any point $y > y^*$. Continuity of the profit function therefore implies that $y = y^*$ is not only a local, but also a global maximum in the interval (y'', y) as long as condition **I2** is satisfied.

■

Lemma 6: $\Pi_1(y^*, y^*) \geqslant \Pi_2(0, y^*)$ if and only if condition $I3$ is satisfied:

$$\hat{t}_1(y^*, y^*) - y^* L/H \leqslant \hat{t}_2(0), \qquad (\mathbf{I3})$$

Proof.

Use the profit specification in (3.13) to get the following expressions for the two profit functions:

$$\Pi_1(y^*, y^*) = yL + (T - \hat{t}_1(y^*, y^*))H$$

$$\Pi_2(0, y^*) = (T - \hat{t}_2(0))H$$

There are no incentives to deviate to $y = T$ if the former expression exceeds the latter:

$$\Pi_1(y^*, y^*) \geqslant \Pi_2(0, y^*) \quad \Leftrightarrow \quad \hat{t}_1(y^*, y^*) - y^* L/H \leqslant \hat{t}_2(0)$$

■

Lemma 7: $\Pi_1(y^*, y^*) \geqslant \Pi_4(T, y^*)$ if and only if condition **I4** is satisfied:

$$\hat{t}_1(y^*, y^*) - y^* L/H \leqslant T(1 - L/H), \qquad (\mathbf{I4})$$

Proof.

From (3.13), deviation payoffs are as follows:

$$\Pi_4(T, y^*) = TL$$

There are no incentives to deviate to $y = T$ if

$$\Pi_1(y^*, y^*) \geqslant \Pi_4(T, y^*) \quad \Leftrightarrow \quad \hat{t}_1(y^*, y^*) - y^* L/H \leqslant T(1 - L/H)$$

■

Lemma 8: $\Pi_1(0, 0) \geqslant \Pi_1(y, 0), \quad \forall y \in [0, y')$ if and only if condition **T1** is satisfied:

$$\frac{n-m-H/L}{n-1} \leqslant \gamma^{T'} I^{-1} \quad \textbf{(T1)}$$

where y' solves $\hat{t}_1(y', 0) = T$.

Proof.

Payoffs for any $y \in [0, y')$ are calculated the following way, from equation (3.22a):

$$\Pi_1(y, 0) = yL + (T - \hat{t}_1(y, 0))H \tag{3.39}$$

A necessary condition for the payoff to be maximized at $y = 0$ is the non-positive sign of the first derivative of (3.39) with respect to y at $y = 0$. We first apply a strictly increasing function $-\gamma^{(\cdot/H)}$ to the payoff function an then differentiate the transformed function with respect to y to obtain the following condition:

$$\begin{aligned}
\frac{\partial - \gamma^{\Pi_1(y,0)/H}}{\partial y} &\leqslant 0 \quad \Leftrightarrow \\
\frac{\log(\gamma)\gamma^T}{I^{-1} - \frac{n-m}{n-1}} \times \gamma^{-y}\left(\frac{1}{n-1}(L/H - 1) + \frac{n-m-1}{n-1}L/H - \gamma^{T'} I^{-1} L/H\right) &\leqslant 0 \quad \Leftrightarrow \\
\gamma^{-y}\frac{1}{n-1}(L/H - 1) + \frac{n-m-1}{n-1}L/H - \gamma^{T'} I^{-1} L/H &\leqslant 0
\end{aligned} \tag{3.40}$$

Inequality (3.40) must hold for $y = 0$:

$$\begin{aligned}
\left.\frac{\partial - \gamma^{\Pi_1(y,0)/H}}{\partial y}\right|_{y=0} &\leqslant 0 \quad \Leftrightarrow \\
\frac{1}{n-1}(L/H - 1) + \frac{n-m-1}{n-1}L/H - \gamma^{T'} I^{-1} L/H &\leqslant 0 \quad \Leftrightarrow \\
L/H\frac{n-m}{n-1} - \frac{1}{n-1} &\leqslant \gamma^{T'} I^{-1} L/H \quad \Leftrightarrow \\
\frac{n-m-H/L}{n-1} &\leqslant \gamma^{T'} I^{-1}
\end{aligned} \tag{3.41}$$

To obtain the second derivative, differentiate the the left-hand side of (3.40) with respect to y and simplify to get:

$$\frac{\partial^2 - \gamma^{\Pi_1(y,0)/H}}{\partial y^2} = \frac{1}{n-1}(L/H - 1)(-1)\log\gamma$$

Note that the second derivative is always negative because $\gamma \leqslant 1$ and $H > L$. If condition **T1** is satisfied, the first derivative will be non-positive at point $y = 0$, and it will non-positive for any $y \in (0, t')$. Payoffs would therefore

be maximized by choosing $y = 0$. If **T1** does not hold, the first derivative is positive at point $y = 0$ and profits could be increased by choosing $y > 0$.

■

Lemma 10. $\hat{t}_1(y, \bar{y}) \leqslant \hat{t}_2(y)$ *and* $\hat{t}_1(y, \bar{y}) \leqslant \hat{t}_3(\bar{y})$

Proof.

Note that $\hat{t}_1(y, \bar{y})$ is increasing both in y and in $\bar{y}$, from equation (3.20). If y is held constant, at any given time t the maximum value of $\hat{t}_1(y, \bar{y})$ will be reached at $\bar{y} = t$. Substituting $\bar{y} = t$ into equation (3.17) reduces it to equation (3.24), thus $\max_{\bar{y}} \hat{t}_1(y, \bar{y}) = \hat{t}_2(y)$. Likewise, setting $y = t$ in equation (3.17) reduces it to equation (3.27), thus $\max_y \hat{t}_1(y, \bar{y}) = \hat{t}_3(\bar{y})$. Therefore $\hat{t}_1(y, \bar{y})$ can never exceed $\hat{t}_2(y)$ or $\hat{t}_3(\bar{y})$.

■

3.A.3 Comparative Statics

3.A.3.1 Speed of Transition in the Teaching Equilibrium

We will prove the effect of the parameter changes on the general function $\hat{t}_1(y, y)$, and all results will of course hold for the special case $y = 0$.

Lemma 11. *Speed of transition to the efficient state in a teaching equilibrium depends on the parameter values the following way:*

1. $\frac{\partial \hat{t}_1(y,y)}{\partial y} > 0$
2. $\frac{\partial \hat{t}_1(y,y)}{\partial m} > 0$
3. $\frac{\partial \hat{t}_1(y,y)}{\partial T'} > 0$

Proof.

Assume that a teaching equilibrium exists, so that $\hat{t}_1(0, 0) < T$ and $\frac{n-m}{n-1} > I^{-1}$. We will show how the speed of transition in this type of equilibrium respond to changes in parameter values. The switching period $\hat{t}_1(y, y)$ is calculated using equation (3.20):

$$\hat{t}_1(y, y) = \frac{1}{-\log(\gamma)} \left[\log(\gamma^{-y} \frac{n-m}{n-1} - \gamma^{T'} I^{-1}) - \log(\frac{n-m}{n-1} - I^{-1}) \right] \tag{3.42}$$

1. *Derivative with respect to y:*

$$\frac{\partial \hat{t}_1(y, y)}{\partial y} = \frac{\frac{n-m}{n-1}}{\frac{n-m}{n-1} - \gamma^{y+T'} I^{-1}} \tag{3.43}$$

$\frac{\partial \hat{t}_1(y,y)}{\partial y} > 0$ because $\frac{n-m}{n-1} > I^{-1}$ and $\gamma \in (0, 1)$.

2. *Derivative with respect to m:*

$$\frac{\partial \hat{t}_1(y,y)}{\partial m} = -\frac{1}{\log(\gamma)}\frac{1}{n-1}\left(\frac{1}{\frac{n-m}{n-1} - I^{-1}} - \frac{1}{\frac{n-m}{n-1} - \gamma^{T'+y}I^{-1}}\right) =$$
$$= -\frac{1}{\log(\gamma)}\frac{1}{n-1}\frac{I^{-1}(1-\gamma^{T'+y})}{(\frac{n-m}{n-1} - I^{-1})(\frac{n-m}{n-1} - \gamma^{T'+y}I^{-1})} \tag{3.44}$$

$\frac{\partial \hat{t}_1(y,y)}{\partial m} > 0$ because $\frac{n-m}{n-1} > I^{-1}$ and $\gamma \in (0,1)$.

3. *Derivative with respect to T':*

$$\frac{\partial \hat{t}_1(y,y)}{\partial T'} = -\frac{1}{\log(\gamma)}\frac{1}{\gamma^{-y}\frac{n-m}{n-1} - \gamma^{T'}I^{-1}} \times -I^{-1}\gamma^{T'}\log(\gamma) \tag{3.45}$$

$\frac{\partial \hat{t}_1(y,y)}{\partial T'} > 0$ because $\frac{n-m}{n-1} > I^{-1}$ and $\gamma \in (0,1)$.

■

3.A.3.2 Speed of Transition in the Interior Equilibrium

Lemma 12. *Speed of transition to the efficient state in an interior equilibrium depends on the parameter values the following way:*

1. $\dfrac{\partial \hat{t}_1(y^*,y^*)}{\partial m} > 0$
2. $\dfrac{\partial \hat{t}_1(y^*,y^*)}{\partial T'} = -1$

Proof.

Assume that an interior equilibrium exists, so that $\hat{t}_1(y^*,y^*) < T$ and $\frac{n-m}{n-1} > I^{-1}$. In an interior equilibrium changes in parameter values affect both the equilibrium strategies of sophisticated players and the switching period of myopic players, holding the strategies of sophisticated players constant. To measure the total effect we substitute the expression of y^* from equation (3.35) into (3.42) to obtain the following result:

$$\hat{t}_1(y^*,y^*) = \frac{1}{-\log(\gamma)}\left[\log\left(\frac{I^{-1}(n-1)\gamma^{T'}}{n-m-H/L}\frac{n-m}{n-1} - \gamma^{T'}I^{-1}\right) - \log\left(\frac{n-m}{n-1} - I^{-1}\right)\right] =$$
$$= \frac{1}{-\log(\gamma)}\left[\log(H/L) + \log(\gamma^{T'}) + \log(I^{-1}) - \log(n-m-H/L) - \log\left(\frac{n-m}{n-1} - I^{-1}\right)\right]$$

1. *Derivative with respect to m:*

$$\frac{\partial \hat{t}_1(y^*,y^*)}{\partial m} = -\frac{1}{\log(\gamma)}\left(\frac{1}{n-m-H/L} + \frac{n-1}{\frac{n-m}{n-1} - I^{-1}}\right) \tag{3.46}$$

$\frac{\partial \hat{t}_1(y^*,y^*)}{\partial m} > 0$ because $\frac{n-m}{n-1} > I^{-1}$, $\gamma \in (0,1)$ and $n - m - H/L > 0$ (if an interior equilibrium exists).

2. *Derivative with respect to T':*

$$\frac{\partial \hat{t}_1(y^*,y^*)}{\partial T'} = -1 \tag{3.47}$$

■

3.A.3.3 Speed of Transition if One Player is Teaching

Here we will calculate how the parameters of interest affect $\hat{t}_2(0)$, which measures the transition speed if a single sophisticated player always plays A while all others play B. This derivative is necessary for Proposition 9 and Proposition 10 because the existence of a delay equilibrium depends on $\hat{t}_2(0)$

Lemma 13. *Speed of transition to the efficient state if only one sophisticated player is choosing A depends on the parameter values the following way:*

1. $\dfrac{\partial \hat{t}_2(0)}{\partial m} = 0$
2. $\dfrac{\partial \hat{t}_2(0)}{\partial T'} > 0$

Proof.

Suppose that $\hat{t}_2(0) < T$, which holds only if $\frac{1}{n-1} > I^{-1}$. Then $\hat{t}_2(0)$ is calculated the following way, from expression 3.25:

$$\hat{t}_2(0) = \frac{\log(\frac{1}{n-1} - I^{-1}) - \log(\frac{1}{n-1} - \gamma^{T'} I^{-1})}{\log(\gamma)} \tag{3.48}$$

1. *Derivative with respect to m:*

$$\frac{\partial \hat{t}_2(0)}{\partial m} = 0 \tag{3.49}$$

2. *Derivative with respect to T':*

$$\frac{\partial \hat{t}_2(0)}{\partial T'} = \frac{\gamma^{T' I^{-1}}}{\frac{1}{n-1} - \gamma^{T'} I^{-1}} \tag{3.50}$$

$\frac{\partial \hat{t}_2(0)}{\partial T'} > 0$ because $\frac{1}{n-1} > I^{-1}$.

■

3.A.3.4 Equilibrium Strategies in the Interior Equilibrium

Another variable if interest is the strategy used by sophisticated players in an interior equilibrium, y^*, which has an effect on the existence conditions of the interior equilibrium.

Lemma 14. *The strategies used by sophisticated players in an interior equilibrium depend on parameter values the following way:*

1. $\dfrac{\partial y^*}{\partial m} > 0$
2. $\dfrac{\partial y^*}{\partial T'} = -1$

   In addition:

3. $\dfrac{\partial y^*}{\partial m} < \dfrac{\partial \hat{t}_1(y^*, y^*)}{\partial m}$

Proof.

Equilibrium strategy is determined by equation (3.35):

$$y^* = \frac{\log(\frac{n-m-H/L}{I^{-1}(n-1)})}{\log(\gamma)} - T'$$

1. *Derivative with respect to m:*

$$\frac{\partial y^*}{\partial m} = -\frac{1}{\log(\gamma)}(\frac{1}{n-m-H/L}) \tag{3.51}$$

   $\frac{\partial y^*}{\partial m} > 0$ because $n - m - H/L > 0$ (because an interior equilibrium exists).

2. *Derivative with respect to T':*

$$\frac{\partial y^*}{\partial T'} = -1 \tag{3.52}$$

3. *Comparison to the derivative of $\hat{t}_1(y^*, y^*)$:*

   Recall the derivative of $\hat{t}_1(y^*, y^*)$ from equation (3.46):

$$\frac{\partial \hat{t}_1(y^*, y^*)}{\partial m} = -\frac{1}{\log(\gamma)}(\frac{1}{n-m-H/L} + \frac{n-1}{\frac{n-m}{n-1} - I^{-1}})$$

   The derivative of y^* calculated in (3.51) is strictly lower than the derivative of $\hat{t}_1(y^*, y^*)$ because $\frac{n-1}{\frac{n-m}{n-1} - I^{-1}} > 0$.

■

Chapter 4

Overcoming Coordination Failure in a Critical Mass Game: Strategic Motives and Action Disclosure

4.1 Introduction

This chapter addresses the same problem as chapter 3—convergence to an inefficient equilibrium in critical mass games—but uses experiments instead of theory. Chapter 3 has shown that strategic players could deviate from an inefficient convention to teach others and increase their own future earnings. This chapter presents results of an experiment that tests whether players are actually motivated by such strategic considerations. The strategic teaching hypothesis is tested in several ways. First, an independent task was run to classify participants according to their ability to anticipate future actions and to test the prediction that sophisticated players deviate from an inefficient equilibrium more often than myopic players. Of course, strategic teaching is not the only possible explanation: we show that heterogeneity in risk preferences, social preferences and beliefs could also in theory explain heterogeneity in choices, therefore additional tests were run to elicit these preferences and beliefs. The second approach to determine whether an action is taken for strategic reasons relies on an observation that strategic players value their actions not only for the immediate utility that they generate, but also for the feedback these actions provide to other group members. Sophisticated players may therefore be willing to pay to make their actions observable, while players who care only

about immediate utility would never do so. Disclosures should also follow a particular pattern: only deviations from an inefficient equilibrium should be disclosed, while a decision to remain in an inefficient equilibrium should not be.

The second aim of this chapter is to test whether the possibility to disclose actions at a lower cost facilitates transitions to an efficient equilibrium. Knowing whether disclosure costs facilitate transitions may help design policies to reduce inefficiencies in real life situations. Such knowledge could also help explain what factors contributed to transitions in the past. For example, lower disclosure costs could be one mechanism through which information technology has contributed to the democratization process, a situation that could be modeled as a game with two alternative states following coordination failure. Citizens in oppressed countries might take no political action because the history of no revolts establishes an inefficient convention, as defined by Lewis (1969) and Young (1996): taking no action is customary (everyone has been taking no action), expected (everyone expects others to take no action) and self-enforcing (taking no action is optimal as long as others do so). However, recent events in Tunisia, Egypt and other countries have shown that protest movements can be successful, and this success has often been attributed to information and communication technologies (ICT) that spread at the onset of the protest movements and were used by many protesters.[1] On the other hand, there is evidence that ICT may also have adverse effects and its role in mobilizing participants may have been exaggerated.[2] Unfortunately, scientific evidence about the role played by ICT in the democratization process is lacking and the few studies that do address this question use empirical country-level data, making the causal relationship hard to establish (see Kedzie, 1997, Best and Wade, 2009, and a review in Meier, 2011).

In the experiment we seek to implement some features that inhibit revolts in real world situations. We allow participants to choose between a safe action (call it "stay out") with a fixed payoff, and a risky action (call it "revolt") that pays a large payoff if sufficiently many group members choose this action. Such critical mass games are often used to represent situations faced by citizens in oppressed countries (Angeletos et al., 2007; Edmond, 2013). Three

[1]For example, almost half of the protesters in Egypt said that they have taken and shared pictures or videos of the protests (Tufekci and Wilson, 2012) and social networking services in both Egypt and Tunisia were primarily used to raise awareness inside the country about the civil movements (Mourtada and Salem, 2011)

[2]It has been stated that social media may be a result rather than a cause of discontent with the regime (Shirky, 2011), online activism may crowd out activism in the streets (Morozov, 2012) and protests movements are not hampered by attempts to shut down communication channels (Hassanpour, 2014). For a broader overview of the debate see Kalathil and Boas (2003), Morozov (2012) or Lynch (2011).

outcomes are possible in this game: efficient coordination, when everyone is revolting, coordination failure, when everyone is staying out, and miscoordination, when both actions are played by some participants.[3] Even though the situation in which everyone is revolting is preferred by all citizens, it may not be implemented due to the history of no revolts and imperfect observability of the actions taken by others. We established an inefficient convention by starting the game with a high participation threshold, which in similar games leads to coordination on the inefficient equilibrium (e.g. Brandts and Cooper, 2006a). Efficient coordination may also be inhibited by imperfect observability of revolt levels because participants are unable to observe deviations from an inefficient convention. However, observability is improved if players can use information technology, which provides an easier way to disclose information about one's actions. In the experiment we gave players an option to disclose their action by paying a certain cost, and only the disclosed actions were observed in the feedback stage. Disclosure costs were varied across treatments: it was cheap to do it in the *LOW* treatment but expensive in the *HIGH* treatment.

We find that low disclosure costs do not prevent coordination failure, but they help overcome it when the participation threshold is lowered. In the first block all groups in both treatments converge to an inefficient equilibrium and in *HIGH* this coordination failure persists throughout the second block in all groups. In *LOW*, on the other hand, half of the groups manage to overcome coordination failure. In these groups efficient coordination is also sustained in the third block, in which the participation threshold is increased back to its original level. Revolt levels are similar in the two treatments right after the decrease in the threshold, but more revolting players disclose their actions in *LOW*, leading to higher observed revolt levels. We present evidence suggesting that it is this difference in observed revolt levels that allows half of the groups in *LOW* to move to the efficient equilibrium while no group in *HIGH* does so.

We also find that strategic considerations are an important factor motivating deviations from an inefficient convention. Many players are willing to pay to disclose their actions and disclosures are systematic: revolts are disclosed much more often than stay outs and disclosures are most frequent at the start of each block. Finally, when disclosure costs are low farsighted players are much more likely to initiate revolts than their less farsighted counterparts.

[3]The term "coordination failure" may be used to denote failure to coordinate on an equilibrium, but we use terminology that is consistent with the literature, such as Van Huyck et al. (1991) and Devetag and Ortmann (2007).

4.1.1 Related Literature

Experiments on overcoming coordination failure in coordination games have been conducted by Brandts and Cooper (2006a) and Hamman et al. (2007), although their focus on coordination failures in organizations has lead to different design choices.[4] Both studies find that whether a group overcomes coordination failure can be predicted by the outcomes in the first round, and some players (called "leaders" by Brandts and Cooper, 2006a) are more likely to initiate the transition than others ("laggards"). Our study extends these findings by showing that whether a person is a leader or a laggard partly depends on their farsightedness level.

Other studies have focused on the role of feedback in coordination games. Brandts and Cooper (2006b) show that groups receiving full feedback about the choices of other participants overcome coordination failure more often than groups receiving information only about the minimum effort level. Devetag (2003) finds that full feedback improves efficient coordination in a game that is similar to ours, but without the initial period of inefficient coordination. On the other hand, Van Huyck et al. (1990) note that informing players about the distribution of actions does not improve coordination in a minimum effort game. In our study the level of feedback is determined endogenously, allowing us to investigate not only the effects, but also the causes of action disclosure. Also, heterogeneity in received feedback allows us to investigate the relationship between observed actions and choices more precisely and to determine how efficient coordination is affected by small increases in observed actions.

Another factor that improves efficient coordination is pre-play communication (Cooper et al., 1992; Blume and Ortmann, 2007), but in many situations, both in the lab and in the field, explicit communication is not possible. However, even without explicit communication players are often able to send signals to others using their actions, as long as these actions are observed. Action disclosure decisions allow us to determine whether players are willing to send such signals even when doing so is costly and whether these signals increase the rates of efficient coordination.

The hypothesis that some players are acting strategically to influence future

[4]The game used in this study does not have the weak-link structure, thus it is not necessary for all participants to cooperate for the efficient equilibrium to be reached (although we do use a weak-link structure in the first block of the game, as it has been shown that in such games most groups converge to the inefficient equilibrium); after the initial block in which groups coordinate on the inefficient equilibrium, literature on coordination failure in organizations changes the bonus that is paid to group members if the efficient equilibrium is achieved, while we model a decrease in the strength of the regime by lowering the participation threshold; participants in our experiment choose one of two actions, leading to two Nash equilibria, while the strategy space and the number of equilibria in the minimum effort game are larger.

actions of other participants has been investigated in two player games with two Pareto-ranked Nash equilibria (Hyndman et al., 2009), three non Pareto-ranked equilibria (Terracol and Vaksmann, 2009) and a unique equilibrium (Hyndman et al., 2012). All three studies find evidence for strategic teaching by showing that (efficient) coordination is less likely when: (i) the cost of strategic teaching is increased (Hyndman et al., 2009), (ii) players are rematched in each round or (iii) information about the payoffs of other players is not displayed (Hyndman et al., 2012). We contribute to the strategic teaching literature in several ways. First, we investigate whether strategic teaching is possible in games with more than two players. For strategic teaching to be profitable in two player coordination games the teacher must be able to affect the beliefs of the other party about one's future choices. In N-person games strategic teaching by one player may be not enough, even if all other players are receptive learners. To find strategic teaching profitable, players must expect that others are using strategic teaching too, and each player would prefer others to perform strategic teaching rather than performing it themselves, leading to a collective action problem and potentially lowering the willingness to use strategic teaching. Second, we test the strategic teaching hypothesis in a different way and using a smaller treatment manipulation, making it unlikely that the treatment difference could be explained by a theory that is not based on players acting strategically.[5] Costly action disclosure also allows us to identify teaching attempts more precisely, and we show that some personal characteristics, in particular farsightedness, can partly explain why some players use strategic teaching while others do not.

More generally, the critical mass game resembles other repeated coordination games with strategic complementarities, complete information and more than two players, such as entry games (Heinemann et al., 2004; Duffy and Ochs, 2012), order-statistic games (Van Huyck et al., 1990, 1991, 2007; Kogan et al., 2011), games with network externalities (Ruffle et al., 2010; Mak and Zwick, 2010) and step-level public goods games (Rapoport and Eshed-Levy, 1989; Offerman et al., 1996; Sonnemans et al., 1998). For similar one-shot games, see Heinemann et al. (2009) or Keser et al. (2012). Global games of regime change use a similar payoff function, but assume incomplete information about the participation threshold (Angeletos et al., 2007, tested in an experiment by Shurchkov, 2013).

[5]Findings made in two player games, while consistent with strategic teaching, might also be explained by other factors: larger costs of strategic teaching would lower the likelihood that a player will choose R even for players who are myopic but choose stochastically; rematching players every round could increase strategic uncertainty and make learning more difficult; social preferences might be switched off when players are not informed about payoffs of other participants, leading to different choices.

Several theoretical studies in political science proposed a hypothesis that political action could be taken strategically, to increase future turnout levels. Lohmann (1994, 1993, 2000) develops a signaling theory of mass action in which citizens with private information about regime strength take part in mass political action to induce others to join; De Mesquita (2010) models a coordination game with uncertainty about the level of anti-government sentiment, and vanguards undertake some costly action (e.g. acts of terror) to make the anti-government support seem larger and increase future revolt levels; Edmond (2013) uses a model in which citizens are not perfectly informed about the strength of the regime and thus the regime has incentives to provide biased information. Although the setting which we analyze is different from any of these theoretical models[6], our experiment reinforces the suggestion that participation in political action could be used strategically.

4.2 Experimental Design

At the start of the experiment participants were matched into groups of six and played a variant of a critical mass game (Heinemann et al., 2009; Devetag, 2003) for 33 rounds. In each round players went through four steps: they (i) chose an action, (ii) chose whether to disclose this action, (iii) reported beliefs about the actions chosen by other group members and (iv) received feedback about the actions of the group members who disclosed their actions.

4.2.1 Action Choice

In each round participants chose an action, R (for *"risky"* or *"revolt"*) or S (for *"safe"* or *"stay out"*).[7] The payoff of S was fixed while the payoff of R depended on whether the total number of players who chose R (denoted by $\#R$) exceeded the participation threshold (θ):

$$\pi(R) = \begin{cases} 100 \text{ ECU} & \text{if } \#R \geqslant \theta \\ 5 \text{ ECU} & \text{if } \#R < \theta \end{cases}$$
$$\pi(S) = \quad 60 \text{ ECU}$$

[6]De Mesquita (2010) and Edmond (2013) assume uncertainty about the strength of the regime (corresponding to the participation threshold in this study) or the preferences of the population. A strategic player thus affects the actions of other players by altering their beliefs about the unknown variable. The experiment in this study makes the strength of the regime public knowledge, and the preferences of all participants are aligned, but there is strategic uncertainty that can be affected by a strategic choice.

[7]In the experiment we used neutral language, labeling actions as "A" and "B". Screenshots of the decision screen are reproduced in Appendix 4.H

This payoff function with $2 \leqslant \theta \leqslant 6$ leads to two Pareto-ranked stage game Nash equilibria in pure strategies:[8] in the efficient equilibrium everyone chooses R and in the inefficient one everyone chooses S.

Our main objective is to study how an efficiency-improving transition could take place following coordination failure. To establish coordination failure we divided the game into three blocks and set the participation threshold equal to 6 in the first one, so that 100 ECU is received only if all six players in a group choose R. With this parameter choice, a critical mass game reduces to a binary minimum effort game, in which large groups typically converge to the inefficient equilibrium (Van Huyck et al., 1990). To test whether coordination failure can be overcome we made efficient coordination easier in the second block by reducing the participation threshold to 5. In the third block the threshold was increased back to 6 to test if the history of efficient coordination improves coordination in a more difficult environment.

The transfer of precedent from one block to the other was facilitated by keeping all elements of the game constant, except for the participation threshold, and by framing the game in terms of several blocks in a game rather than as separate games. At the start of the block players were informed about the duration of the current block (13 rounds in block 1, 15 rounds in block 2, 5 rounds in block 3) and the value of the participation threshold, but parameter values in future blocks were not known.

4.2.2 Action Disclosure, Belief Elicitation and Feedback

After choosing an action, R or S, players had an option to disclose their action to other group members. The cost of not disclosing an action was 1 ECU, while the cost of disclosing depended on the treatment to which a player was assigned:

- In *LOW* the disclosure cost was 2 ECU.
- In *HIGH* the cost was 80 ECU.

After the action disclosure stage we elicited beliefs using a procedure adapted from Heinemann et al. (2009): players were asked to report the probability that a randomly selected other group member will choose R in the current round. The belief elicitation task was incentivised and we used a binarized scoring rule that elicits truthful beliefs even if subjects are not risk neutral (Hossain and

[8] There also are stage game equilibria in mixed strategies. In a symmetric stage game mixed strategy equilibrium all players choose R with probability $I^{-1}_{\frac{11}{19}}(\theta - 1, n - \theta + 1)$, where I^{-1} is an inverse of the regularized incomplete beta function. R is chosen with probability 0.72 if $\theta = 5$, and with probability 0.90 if $\theta = 6$.

Okui, 2013). In coordination games stated beliefs may not be truthful because of hedging (see Armantier and Treich, 2013), and we reduced the incentives to hedge by paying either for the main part of the experiment or by playing a lottery in which the probability to receive the higher prize is determined by performance in a randomly selected belief elicitation task. Details about the belief elicitation task are presented in Appendix 4.D.

At the end of each round players were informed only about the actions taken by group members who disclosed their actions. In particular, players were informed about:

- The total number of group members who chose R and disclosed ("observed R")
- The total number of group members who chose S and disclosed ("observed S")
- The total number of group members who did not disclose their actions ("unobserved")

In addition, players were informed about their round income and whether at least θ players chose R. A history box displaying own actions, stated beliefs and observed feedback in previous rounds was always visible on the computer screen. Feedback about the accuracy of reported beliefs was provided only at the end of the experiment to suppress information about the choices of group members who did not disclose their actions.

4.2.3 Procedure

Experiments were conducted in the BEElab at Maastricht University in February, 2014, using z-Tree (Fischbacher, 2007) and ORSEE (Greiner, 2015). 72 participants took part in the experiment, 36 in each treatment. The average duration of the experiment was 100 minutes and average earnings were 18.50 euros. Earnings in all parts of the experiment were denoted in ECU and exchanged into euros using the conversion rate of 250 ECU = 1 EUR. Negative round income was possible in *HIGH*, therefore in both treatments players started with an initial balance of 400 ECU.

After reading the instructions but before starting the experiment subjects answered several questions about the payment scheme and calculated earnings in hypothetical situations to ensure that instructions were well understood (these questions are reproduced in figures A.5 and A.6 in Appendix 4.H). The questionnaire was computerised and subjects could not continue until all the questions have been correctly answered.

4.2.4 Theoretical Predictions

In each stage game players chose an action (R or S) and whether to disclose it (D) or not (ND), thus each player had four strategies: (R, D), (R, ND), (S, D) and (S, ND). Disclosure is costly, therefore (R, ND) dominates (R, D) and (S, ND) dominates (S, D) so no action should be disclosed if a stage game was played once or if all players were myopic. However, actions might be disclosed on the path of a subgame perfect Nash equilibrium (SPNE) in the repeated game. For example, consider a strategy that prescribes R if all players disclosed their actions in all previous rounds and prescribes S otherwise. There is a SPNE in which all players use such strategies, and on the equilibrium path all players would be choosing R and disclosing their actions in all but the last round as long as disclosure costs are sufficiently small. In a similar way, there are SPNE in which players do not coordinate on a common action if, for example, all players use strategies that prescribe R only following such history of miscoordination.

Nash equilibrium is useful in making predictions about stage-game outcomes that may be reached in the long run, but the multiplicity of equilibria leads to vacuous predictions about the dynamic process in a repeated game. An alternative approach is to make predictions about the path of play using a learning model, such as weighted fictitious play (Cheung and Friedman, 1997). In this model players form beliefs based on observed history and choose an action that maximises immediate utility conditional on those beliefs. Fictitious play predicts that players who observed others choosing a particular action will expect this action to be chosen more often, and will therefore be more likely to choose this action themselves. We formulate this first prediction the following way:

Prediction M1. *Players who observe action R chosen more often are more likely to choose R and to believe that others will choose R.*

However, players in learning models assume that others are using stationary strategies, ignoring the possibility that others could be learning as well. These models therefore overlook strategic teaching, which may explain deviations from an inefficient convention. To test whether behaviour is driven by such strategic motives we extend the learning model to include sophistication, and compare the predictions of this model to the predictions of a standard learning model. The experiment was designed in a way such that the two learning models would make different predictions, allowing us to determine: (i) whether choice dynamics can be accurately approximated by belief learning and (ii) whether the explanatory power of a learning model is increased by adding

sophistication.

We add sophistication to the standard learning model by assuming that some players are sophisticated and others are myopic. Myopic players act as predicted by the standard belief learning model (see prediction M1). Sophisticated players anticipate the learning process of myopic players and take into account not only the differences in immediate expected payoffs, but also the differences in future payoff flows generated by each action. A solution concept for N-person critical mass games with sophisticated and myopic players is presented in Chapter 3; here we will only use the features of the model that are useful in making predictions about the treatment difference.[9]

Sophisticated players may deviate from an inefficient convention to increase future payoffs from efficient coordination. Furthermore, sophisticated players would be willing to pay to disclose their action because they value the deviation for its informational content rather than for the immediate payoff it generates. A decision to not deviate should never be disclosed because it makes myopic players less likely to choose R in the future, which can only decrease the payoffs of sophisticated players.

Prediction S1. *Action S is never disclosed, R is disclosed more often than S.*

The model with sophistication also predicts differences in disclosure patterns between the two treatments. Action R should be disclosed more often in *LOW* than in *HIGH* because lower disclosure costs reduce the cost of strategic teaching. Since action S should never be disclosed, its disclosure rates should not depend on disclosure costs.

Prediction S2. *R is more often disclosed in LOW than in HIGH.*

In addition to the immediate effect of action disclosure, lower disclosure costs are predicted to help overcome coordination failure. This prediction follows from predictions M1, S1 and S2. Notice that S1 and S2 predict that only R will be disclosed, and more so in *LOW*, therefore M1 predicts that over time the learning process should lead to higher R levels in *LOW* than in *HIGH*. Higher R levels result in an increased frequency of instances when θ or more players choose R, in which case coordination failure would be overcome.

Prediction S3. *R is more frequently chosen and coordination failure is more often overcome in LOW than in HIGH.*

[9]For alterantive ways to model sophistication see Milgrom and Roberts (1991), Kalai and Lehrer (1993) or Camerer et al. (2002a).

Finally, players with a longer planning horizon would be more likely to deviate from an inefficient convention because they put more weight on future benefits of efficient coordination.

Prediction S4. *The tendency to choose R and disclose it is higher among farsighted subjects.*

It is important to note that predictions S1, S2, S3 and S4 would not hold if all players were myopic, because no myopic player would be willing to pay to disclose an action.[10] Predictions S1, S2, S3 and S4 are therefore not only interesting on their own, but they also help to test whether players are motivated by strategic considerations.

4.3 Experimental Results and Analysis

We address two main questions in this section. First, we want to know if lower action disclosure costs help overcome coordination failure. Second, we want to see if the transition process can be explained by a learning model, either with sophistication or without it.

4.3.1 Is Coordination Failure More Often Overcome with Low Action Disclosure Costs?

To compare the treatments, we look at two measures: *R levels*, defined as the total number of players in a group who choose R, and *transition frequency*, defined as the fraction of groups that coordinate on S in the end of block 1 and go on to coordinate on R by the end of block 2. The former metric captures the willingness to initiate a transition while the latter measures the success of these efforts.

4.3.1.1 Differences in R Levels

In block 1 there is no significant difference in terms of R levels (figure 4.1). In both treatments about half of the players choose R in round 1, but the tendency to do so decreases until almost nobody is choosing R at the end of block 1. In the last five rounds of block 1 the fraction of players who choose R is 6% in *LOW* and 3% in *HIGH* while average beliefs about R choice are respectively

[10] Action disclosure could be myopically optimal only if players were willing to throw away money to reduce advantageous inequality ($\beta > 1$ in the model of Fehr and Schmidt, 1999). We tested the existence of such preferences in an additional task, where players could reduce their earnings without affecting the earnings of the other person. We found that such preferences were rare: 94% of the sample chose not to decrease their own earings.

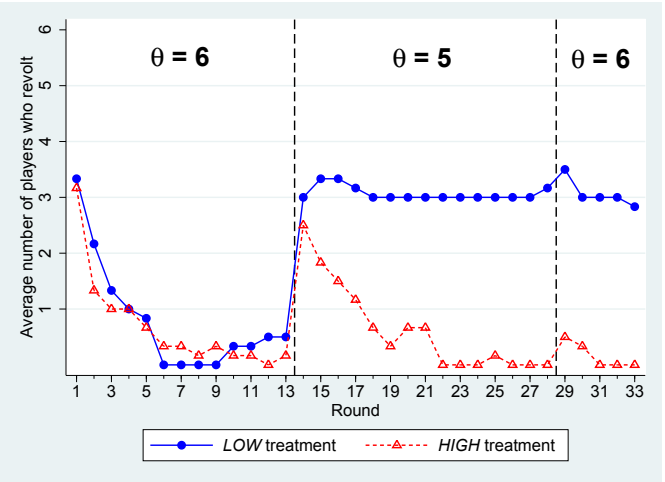


Figure 4.1: Comparison of R levels, by treatment. Vertical dashed lines separate the three blocks with different values of θ.

14% and 13%. A high participation threshold therefore leads to coordination failure regardless of disclosure costs.

The effect of disclosure costs becomes apparent when the participation threshold is lowered in block 2. The treatment difference in round 14 is small and insignificant (Mann-Whitney U test $p > 0.4810$) as a decreased threshold leads to a similar jump in both treatments. However, the gap between two treatments soon widens and eventually almost all players in *HIGH* stop choosing R while average R levels hardly change in *LOW*.

Result 1. *R levels in LOW are significantly higher than in HIGH in blocks 2 and 3, but there is no significant difference in block 1. In block 2 the treatment difference increases over time.*

Result 1 can be supported by regressing R levels on the treatment variable, number of rounds remaining until the end of the block and the interaction of these two variables. Table 4.1 shows that the main effect of the treatment variable is significantly different from zero in blocks 2 and 3, but not in block 1.[11] We use the estimated coefficients to predict the treatment difference in each round (table A.1 in the Appendix) and find no significant treatment difference in block 1 and at the start of block 2, but a significant difference at the end of block 2 and in block 3. The observation that the gap between two treatments grows over time in block 2 is also evident from the negative coefficient of the interaction term in the second column of table 4.1.

[11] If we do not include the interaction term, the treatment coefficient is significant in block 2 and 3 at a 5% level and not significant in block 1.

Table 4.1: Random effects GLS regression using group level data (12 groups). Dependent variable: R levels. Standard errors clustered on the group level.

	Block 1	Block 2	Block 3
LOW treatment	0.060	3.437**	2.900**
	(0.24)	(2.44)	(2.20)
Rounds remaining	0.172***	0.155***	0.133***
	(3.81)	(2.90)	(2.76)
Rounds remaining * *LOW* treatment	0.009	-0.143**	0.000
	(0.15)	(-2.28)	(0.00)
Constant	-0.353***	-0.450***	-0.100**
	(-2.79)	(-2.94)	(-2.28)
Observations	156	180	60

t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

4.3.1.2 Differences in Transition Frequency

A treatment difference is also observed if the comparison is performed using transition frequency instead of R levels. Convergence patters on a group level are illustrated in figure 4.2. All groups converge to an inefficient equilibrium in the first block, and no group is in equilibrium after the change in participation threshold in round 14. However, the equilibrium to which groups converge in block 2 depends on disclosure costs: in *HIGH* all groups return to the inefficient equilibrium, while in *LOW* half of the groups move to the efficient one. The flat line of average R levels in *LOW* shown in figure 4.1 therefore does not tell the entire story: while no group is in equilibrium in round 14, all converge to one by round 18. All three groups in *LOW* that do overcome coordination failure in block 2 stay in the efficient equilibrium in block 3, when the participation threshold is increased to its original value. A temporary decrease in the participation threshold can therefore have a lasting positive effect on the levels of efficient coordination.

Result 2. *No groups overcome coordination failure in HIGH. Half of the groups overcome coordination failure in LOW.*

Overall, we find that lower action disclosure costs increase both the R levels and the frequency of transitions to the efficient equilibrium. Since action disclosure costs should have no effect if all players were myopic, these findings seem to indicate that some players act strategically, a hypothesis that will be investigated in the next subsection.

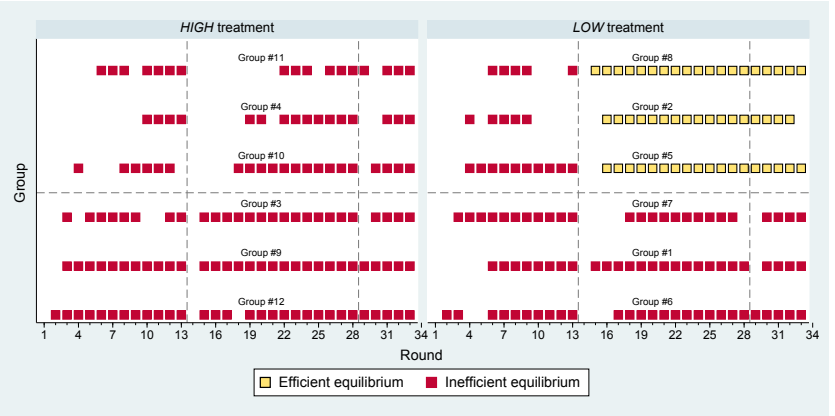


Figure 4.2: Convergence to an equilibrium by group. A group is said to be in an efficient equilibrium if all members choose R and in an inefficient equilibrium if all group members choose S. Empty space denotes rounds in which a group is not in equilibrium. Vertical dashed lines separate three blocks with different threshold value. Groups are ordered by the total number of R choices in the entire experiment. Horizontal dashed lines show the median split, groups with higher R levels being at the top.

4.3.2 Is Adaptation Explained by Belief Learning Models?

Treatment differences in R levels and transition frequency are mainly driven by the adaptation process: figure 4.3 shows that groups which overcome coordination failure in *LOW* (top three groups in the right panel of figure 4.2) and groups that are most likely to choose R in *HIGH* (top three groups in the left panel of figure 4.2) experience similar histories in block 1 and round 14, but all the groups in *LOW* overcome coordination failure while no group does so in *HIGH*. This subsection tests whether the adaptation process can be explained by a learning model with or without sophistication: we first test the predictions made by each model and then fit a learning model to the data and compare the choices simulated by each model to experimental data.

4.3.2.1 Action Disclosure

A learning model with sophistication predicts that R will be disclosed more often than S and more often in *LOW* than in *HIGH* (predictions S1 and S2). These predictions are mostly confirmed by the data (figure 4.4). In the first round of block 2 all players who choose R disclose their actions in *LOW*, but

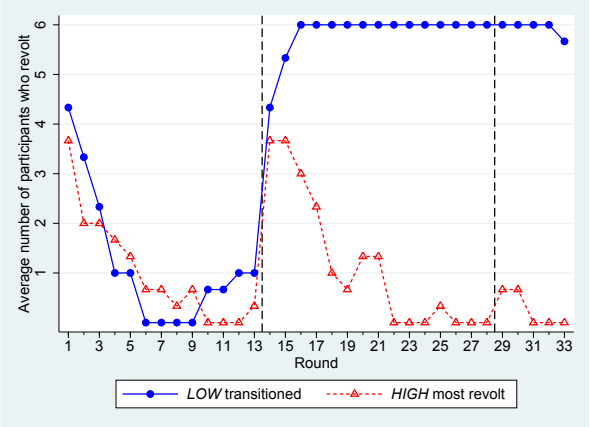
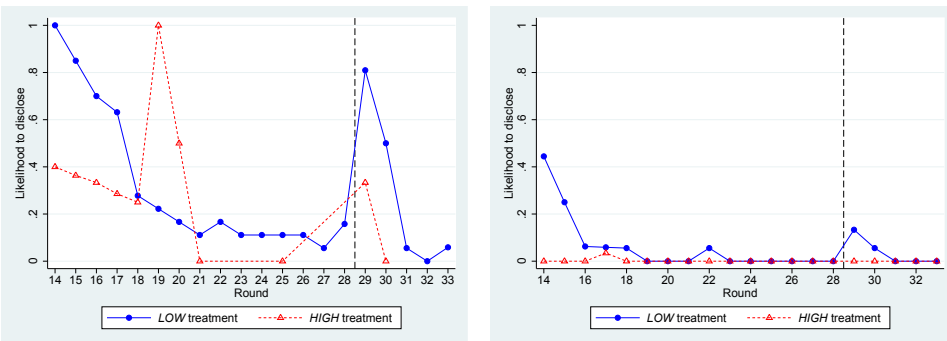


Figure 4.3: R levels for groups in *LOW* that transition to an efficient equilibrium and groups in *HIGH* that choose R most (three uppermost groups in figure 4.2 for each treatment).

only 40% do so in *HIGH*, a difference that persists until all groups converge to an equilibrium in round 18. Actions are more often disclosed at the start of a block than at the end; although this decrease is not the main interest of this chapter, we suspect that disclosures may go down because the lower number of remaining rounds decreases incentives for strategic teaching or because strategic uncertainty decreases as players converge to an equilibrium, reducing the need for disclosure.[12]



(a) Disclosure conditional on choosing R (b) Disclosure conditional on choosing S

Figure 4.4: Frequency of disclosure in *LOW* and *HIGH*, blocks 2 and 3.

The difference between panel (a) and panel (b) of figure 4.4 indicates that S is disclosed much less often than R in both treatments. Still, some players do disclose S in *LOW*, especially at the start of each block, suggesting that such disclosures may also be used for strategic reasons.

[12]Large fluctuations and missing data in *HIGH* after round 18 are caused by a small sample of disclosed actions.

Result 3. *R is disclosed more frequently than S. R is disclosed more frequently in LOW than in HIGH.*

Result 3 is supported by a probit regression of a disclosure decision on action choice and treatment dummies (see table A.3 in Appendix 4.B). The treatment effect is highly significant if all data is taken into account or if we look only at disclosure conditional on choosing R. Action R is disclosed significantly more often than S, and the rate of disclosure decreases over time.

4.3.2.2 Belief Formation

The prediction that observing R increases both the beliefs about R choice and the likelihood to choose R (prediction M1) can be tested on a group level and on an individual level. On a group level, groups with higher initial observed R levels should overcome coordination failure more often. Table 4.2 provides evidence for this prediction: all groups in which at least 4 players choose and disclose R in round 14 overcome coordination failure, while all other groups do not. This difference cannot be explained by path dependence alone, as groups in which many players choose R but do not disclose their actions do not overcome coordination failure (e.g. groups 4 and 11). Overall, the likelihood of a transition depends on the level of *observed* R choices, and this feedback is lacking when disclosure costs are high.

Table 4.2: The number of players in each group who chose R ("R levels")and the number of players who chose R and disclosed it ("Disclosed R") in round 14. A transition is said to occur if all group members were choosing R in the last round of block 2.

Treatment	Group ID	R levels	Disclosed R	Transition?
HIGH	3	2	0	No
HIGH	4	4	3	No
HIGH	9	1	0	No
HIGH	10	3	1	No
HIGH	11	4	2	No
HIGH	12	1	0	No
LOW	1	1	1	No
LOW	6	1	1	No
LOW	7	3	3	No
LOW	2	4	4	Yes
LOW	5	4	4	Yes
LOW	8	5	5	Yes

Result 4. *Transitions to an efficient equilibrium are correlated with R levels a group observed in round 14.*

Table 4.3: Three-level nested random effects model and two-level random effects GLS model (12 clusters). Dependent variable: stated probability that a randomly chosen group member chose R (from 0 to 100). Standard errors are robust to correlation within the observations of each group.

	Random effects GLS		Nested random effects	
	Two-period lags	One-period lags	Two-period lags	One-period lags
Observed R in (t-1)	11.20***	11.58***	10.47***	10.83***
	(12.90)	(9.19)	(9.36)	(8.47)
Observed R in (t-2)	2.440**		1.829	
	(2.33)		(1.57)	
Observed S in (t-1)	-2.940**	-2.041	-3.639***	-3.191**
	(-2.50)	(-1.37)	(-2.98)	(-2.00)
Observed S in (t-2)	0.916		-0.0777	
	(1.01)		(-0.07)	
Actual R level in (t-1)	0.362	3.682***	1.128	3.940***
	(0.31)	(3.84)	(0.89)	(4.50)
Actual R level in (t-2)	1.680**		1.836**	
	(1.98)		(2.10)	
Treatment: 1 = *LOW*	-1.029	0.0160	0.189	1.116
0 = *HIGH*	(-0.49)	(0.01)	(0.09)	(0.54)
Constant	13.91***	13.91***	13.82***	13.96***
	(6.14)	(6.77)	(6.30)	(6.86)
Observations	1944	2160	1944	2160

t statistics in parentheses
* p<0.10, ** p<0.05, *** p<0.01

On an individual level, we can test if players who observe more R choices report higher beliefs. To test this prediction we regress stated beliefs on the level of observed R and observed S in the previous two rounds. A correct specification must take into account the fact that our units of observation are not independent: stated beliefs might be correlated with the person's stated beliefs in other rounds as well as the stated beliefs of other group members in the current round. Thus the most appropriate specification should include an individual random effect nested within a group random effect (three-level nested random effects model). Additionally, we estimate a standard two-level model with only a group-specific random component. As an additional robustness check we include a model with only one round lags.

Table 4.3 shows that observed R levels have a significant positive effect on beliefs about R choice in the subsequent round. The actual R level is also significant, but the significance level and the economic size are much lower than that of the observed R level. Observing more players choosing S is not significant, suggesting that players who do not disclose their actions are perceived in a similar way to those who choose S.

Models in table 4.3 treat observations from all rounds equally, therefore results are influenced by the correlation between choices and beliefs after play has converged to an equilibrium, potentially overstating the effect of belief

learning. We therefore replicate the estimation using data only from rounds in which play has not yet converged, where convergence is said to occur if in the previous round all group members chose R or if all chose S. In this subset the effect of observed feedback on beliefs is lower, but still highly significant (see table A.5 in Appendix 4.B).

4.3.2.3 Estimation of a Weighted Fictitious Play Model

After establishing that beliefs are related to observed R levels, we can specify the adaptation process by fitting a belief learning model. We used weighted fictitious play (Cheung and Friedman, 1997), modified to accommodate three features of the game. First, we have to extend the model to N-person games because Cheung and Friedman (1997) specify it only for two player games. We could have done it by allowing separate beliefs about the action of each particular group member, but players could not update such beliefs because in the experiment individual choices could not be tracked. Instead, we make a homogeneity assumption (Rapoport, 1985; Rapoport and Eshed-Levy, 1989), which states that players form a single belief $b_i(t) \in [0, 1]$ about the probability that any other group member will choose R.[13] Second, the option to not disclose one's action results in incomplete feedback, thus additional assumptions need to be made about how unobserved actions are perceived. If beliefs were formed using Bayes rule, the interpretation of undisclosed actions would depend on beliefs about the probability that R is disclosed, which is unknown to the players. Instead of imposing Bayesian updating and correct beliefs we use a different approach and add a new parameter (ρ) which specifies how an unobserved action is perceived. In particular, if the number of other group members who did not disclose in round t is denoted by $\#U(t)$, the perceived R level would be calculated by $\#R(t) + \rho * \#U(t)$, with $\rho \in [0, 1]$.[14] If $\rho = 1$, unobserved actions are perceived as R while if $\rho = 0$, unobserved actions are perceived as S. The third feature is the change in the participation threshold from one block to the next, which might mean that experience in one block is not directly applicable to the next block. We therefore allow players to discount experience from a previous block by multiplying the observations from the first block by β_1 and the observations from the second block by β_2. Putting everything together, we

[13] This specification of fictitious play assumes that players expect each other member to use a stationary mixed strategy. In a game with two strategies prior beliefs about the probability to choose R follow a beta distribution, and are updated using Bayes rule, assuming that the aggregate number of choices were independently drawn from a binomial distribution.

[14] Equivalently, we could allow players to form three separate beliefs $b_i^R(t)$, $b_i^S(t)$ and $b_i^U(t)$ about the observed number of R, S and U in a standard way, and calculate beliefs about the perceived number of R as a convex combination of $b_i^R(t)$ and $b_i^U(t)$.

use the following formula to determine beliefs of player i about the likelihood that that any other group member will choose R in round t, for all $t > 1$:

$$b_i(t) = \frac{\sum_{u=1}^{t-1} \gamma^{u-1} \frac{\#R(t-u)+\rho\#U(t-u)}{n-1} D(t-u)}{\sum_{u=1}^{t-1} \gamma^{u-1} D(t-u)} \tag{4.1}$$

$$\text{where } D(\tau) = \begin{cases} \beta_1 \times \beta_2 & \text{if } \tau \in [1, 13] \\ \beta_2 & \text{if } \tau \in [14, 28] \\ 0 & \text{otherwise} \end{cases}$$

The model was fit by finding parameters γ, ρ, β_1 and β_2 that minimise the mean squared deviation between $b_i(t)$ and the stated belief. The estimated coefficient values are as follows: $\hat{\gamma} = 0.43$, $\hat{\rho} = 0.15$, $\hat{\beta}_1 = 0.28$ and $\hat{\beta}_2 = 0.99$. The low estimated value of ρ shows that most players react to an unobserved action as if it was S rather than R. Adding the β_1 parameter improves the fit substantially, and the low estimate shows that history in the first block has only a weak effect on choices in block 2. The estimated value of β_2, however, is close to 1, indicating that history in block 2 has a strong effect on beliefs in round 3. Differences between the estimated values of β_1 and β_2 reflect an asymmetric effect of changes in θ: a decrease allows some groups to move from the inefficient equilibrium to the efficient one, but a subsequent increase does not lead back to the inefficient equilibrium.

We evaluate the fit of a belief learning model by comparing stated beliefs with fictitious play beliefs, calculated using the estimated population level parameters and the actual history of observed feedback. Figures A.1 and A.2 (in Appendix 4.A) show that predicted beliefs track stated beliefs rather well even on an individual level. Largest discrepancies arise for players whose stated beliefs are not sensitive to observed history. Overall, both the regressions and the estimation of a belief learning model indicate that stated beliefs are affected by the observed history, as is predicted by a belief learning model.

Result 5. *Players who observed larger R levels in previous rounds expect that others will be more likely to choose R.*

4.3.2.4 Response to Stated Beliefs

The previous subsection has shown that beliefs can be predicted using observed feedback and a weighted fictitious play model. Next, we want to determine if stated beliefs are informative of choices. Rational players should base their decisions on the expected utilities, calculated by

$$EU(S) = u(60)$$

$$EU(R) = u(100) \times Pr(\#R \geqslant \theta) + u(5) \times (1 - Pr(\#R \geqslant \theta))$$

where $EU(S)$ and $EU(R)$ are the expected utilities of actions S and R and $Pr(\#R \geqslant \theta)$ is the subjective probability assigned to an event that the threshold will be exceeded. To calculate this probability we use elicited probabilistic beliefs $b_i(t)$ to construct a probability distribution over the number of group members who choose R. In these calculations we assume that players expect the actions of other group members to be independently drawn from a binomial distribution with the stated probability and that beliefs are homogeneous (Rapoport, 1985; Rapoport and Eshed-Levy, 1989), so that the same probability is assigned to each group member.[15] The probability that the threshold will be exceeded is calculated as follows:

$$Pr(\#R \geqslant \theta) = \sum_{k=\theta-1}^{n-1} (b_i(t))^k (1 - b_i(t))^{n-1-k} \binom{n-1}{k}$$

How expected utilities are mapped into choices is determined by a choice rule. We compare the predictions of the following six choice rules:

1. **Deterministic choice** (DET). The action with a higher expected utility is always chosen, utility is equal to the monetary payoff.

2. **Logistic choice** (LOG). Utility is equal to the monetary payoff, but each action can be chosen with positive probability, determined by a logistic choice rule:

$$Pr(R) = \frac{e^{\lambda EU(R)}}{e^{\lambda EU(R)} + e^{\lambda EU(S)}}, \text{ with } \lambda \in [0, \infty)$$

3. **Risk preferences** ($RISK$). In addition to a logistic choice rule, this specification includes a utility function with constant relative risk aversion:

$$u(\pi_i) = \frac{\pi_i^{1-r}}{1-r}$$

4. **Pro-social preferences** (SOC). In addition to their own payoff, subjects also take into account the average earnings of other group members, and the strength of social preferences is captured by parameter $\alpha \in (0, 1)$:

$$u(\pi_i, \pi_{-i}) = (1 - \alpha)\pi_i + \alpha \sum_{j \in N \setminus \{i\}} \frac{\pi_j}{n-1}$$

[15] If separate beliefs were formed, the scoring rule is designed to elicit the expected value of all probabilistic beliefs.

Table 4.4: Estimated parameter values and mean squared deviation (MSD) for each model.

	Model					
Parameter	*DET*	*LOG*	*RISK*	*SOC*	*FAR*	*ALL*
$\hat{\lambda}$	–	11.41	0.02	34.60	12.93	2.82
$\hat{r}$	–	–	-1.85	–	–	-0.43
$\hat{\alpha}$	–	–	–	0.65	–	0
$\hat{h}$	–	–	–	–	6.75	3.60
MSD	242	197.88	181.31	183.83	181.31	180.25
Performance (0 = worst, 100 = best)	0.00	71.45	98.28	94.20	98.28	100.00

5. **Farsightedness** (*FAR*). This choice rule assumes that players take into account the payoffs from h_i rounds, where $h_i \in [1, \infty)$ is the length of the planning horizon.[16] We assume that farsighted players expect all other group members to use the *DET* choice rule, with beliefs formed using weighted fictitious play. The ability to anticipate the reaction of others allows players to choose a best response path to the actions of other players in every future round. The expected utility of an action is calculated by adding the expected utilities from all rounds along this best response path.

6. **All preferences and farsightedness** (*ALL*). This specification combines logistic choice with farsightedness, risk and pro-social preferences.

We estimate the parameters of all six choice rules by minimizing the mean squared deviation (MSD) between the predictions and experimental data. Table 4.4 shows the estimated parameter values and the goodness of fit for each choice rule. Unsurprisingly, the rule that combines all preferences and farsightedness fits data best, although specifications with only risk preferences or farsightedness fit almost equally well with a lower number of parameters. The estimated parameters of the *ALL* model indicate slight risk aversion, farsightedness and no pro-social preferences. Figure A.3 in Appendix 4.A compares predictions with experimental data across rounds. All choice rules fit data well at the end of each block, but models *DET* and *LOG* under-predict R choice following coordination failure in *LOW* at the start of block 2 and all models under-predict R choice in *LOW* at the start of block 3. Overall, deviations from a theoretical best response could be explained either by risk aversion, farsightedness or pro-social preferences.

[16] Players know the length of the game and cannot take into account the earnings from rounds that will not be played. Therefore the actual number of rounds taken into account is $\min\{h_i, T - t + 1\}$, where T is the number of rounds in a block and t is the current round.

Table 4.5: Frequency of transitions to the efficient equilibrium in simulations (measured by the fraction of iterations in which R is the deterministic best response to beliefs in round 28). Left panel shows the number of group members who choose R and disclose in in round 14 and whether a group overcomes coordination failure in the experiment.

	Experiment		Simulations				
Group ID	(R $\wedge$ D) in 14	Transition?	*DET*, *LOG*	*RISK*	*SOC*	*FAR*	*ALL*
8	5	Yes	0%	100%	99%	89%	100%
2	4	Yes	0%	76%	53%	86%	84%
5	4	Yes	0%	84%	57%	86%	88%
4	3	No	0%	29%	10%	47%	31%
7	3	No	0%	21%	2%	48%	22%
11	2	No	0%	2%	0%	4%	2%
1	1	No	0%	0%	0%	0%	0%
6	1	No	0%	0%	0%	0%	0%
10	1	No	0%	3%	0%	2%	2%
3	0	No	0%	0%	0%	0%	0%
9	0	No	0%	0%	0%	0%	0%
12	0	No	0%	0%	0%	0%	0%

Next we want to determine how well each choice rule can predict whether a group will overcome coordination failure. We calculate beliefs in round 14 using estimated population-level fictitious play parameters and observed feedback in rounds 1-13, averaged across all players. Then we use each choice rule with weighted fictitious play to simulate the path of choices in rounds 15-28 for every 6-person group. We run 100 simulations for each model and count the frequency of simulations in which coordination failure is overcome, which is said to occur if in round 28 all group members hold beliefs under which R is a deterministic best response. Results of these simulations and experimental data are provided in table 4.5. No group overcomes coordination failure when choices are simulated with *DET* or *LOG*, while other rules correctly predict higher success rates for groups that actually overcome coordination failure in the experiment (#8, #2 and #5).

Simulations show that belief learning models with risk preferences, social preferences or farsightedness are able to explain how the likelihood to overcome coordination failure depends on R levels observed in round 14. This is, of course, not the only possible explanation, and we review other possibilities in Appendix 4.C. In short, we look at possibilities that players in the two treatments have different personal characteristics, experience different histories in block 1, or that the adaptation process is driven by reinforcement learning or reciprocity. Each of these explanations could potentially explain the treatment difference, but we show that none of them can fully explain experimental data.

4.3.3 Individual Heterogeneity in *LOW*

Whether a group in *LOW* overcomes coordination failure depends on the number of group members who choose R and disclose it in round 14. In this section we will explore whether R is more often chosen by players who are more risk seeking, pro-social or farsighted, as predicted by the theory (see Appendix 4.E). Before showing the results, we will explain the tasks used to elicit these characteristics. Each task was incentivized and these earnings were added to the earnings from the main part of the experiment.

4.3.3.1 Elicitation of Risk and Social Preferences

Risk preferences were elicited using a multiple price list (Holt and Laury, 2002), and the number of safe choices was used as a measure of risk aversion.[17] Note that objective risk was not present in the experiment and uncertainty about payoffs originated only from strategic uncertainty. But if players form probabilistic beliefs, a utility function with a lower Arrow-Pratt measure of risk aversion would make R myopically optimal for a wider range of beliefs (see Appendix 4.E.3 for a proof).

Social preferences were elicited using a Social Value Orientation (SVO) Slider Measure (Murphy et al., 2011), following a z-Tree implementation by Crosetto et al. (2012). In this measure subjects complete six allocation tasks[18] between themselves and some other participant, randomly selected from a different group to prevent meeting former group members. A graphical representation of all tasks is shown in figure A.4, Appendix 4.F. At the end of the experiment one task was chosen for payment and subjects received either the amount that they sent or the amount that was sent to them by the other participant. Answers to each task were used to construct a continuous, uni-dimensional scale of SVO, known as the SVO angle, using the following formula:

$$\text{SVO angle} = \arctan\left(\frac{\sum_{t\in\{1,..,6\}}(P_t^O - 50)}{\sum_{t\in\{1,..,6\}}(P_t^S - 50)}\right)$$

where P_t^S is the amount allocated to oneself in task t and P_t^O is the amount allocated to the other person. SVO angle ranges from -16.26° for perfectly competitive individuals to 61.39° for perfectly altruistic individuals.

[17] 5 subjects switched more than once, therefore results are slightly different if risk aversion is measured using the first switching round. In such case the coefficient of risk aversion decreases and becomes less significant in models 5 and 6 in table 4.7. The coefficient of farsightedness turns significant at a 10% level in model 3 in table 4.7.

[18] We added a seventh task, completed after the six original ones. This task tested whether subjects were willing to give up their earnings to reduce the level of advantageous inequality (see footnote 10). This additional task was not used to calculate the SVO angle.

4.3.3.2 Elicitation of Farsightedness

Players who choose R for strategic reasons must be able to reason about future events and take future payoffs into consideration. If there is heterogeneity in the ability to plan ahead, prediction S4 states that farsighted players, i.e. those with a longer planning horizon, should be more likely to initiate a transition.

Farsightedness was measured using three individual choice tasks, each presented as a decision tree with two decision nodes and two chance nodes (see figures A.10 - A.12 in Appendix 4.H). These tasks were designed by Bone et al. (2009) to test whether subjects are able to plan ahead and anticipate their own future choices. In the second decision node one move dominates the other, making the decision trivial (97% of players chose the dominant move, both in Bone et al., 2009, and in our experiment). Which move is dominant in the first node depends on whether players anticipate their own choices in the second node: players who do so and ignore the payoffs that will not be reached would choose one move (we call it "correct move" as it would maximise payoffs for 97% of the players) while those who treat the second decision node as a chance node, expecting all final payoffs to be equally likely, would choose a different move.

Bone et al. (2009) found that the correct move is chosen by only a third of the subjects. To classify subjects more precisely we designed three tasks with different difficulty levels,[19] leading to an increase in the frequency of correct moves from 21% in the hardest task (task 1) to 42% in the task of intermediate difficulty (task 2) and 65% in the easiest task (task 3). Each subject completed all three tasks in the order of decreasing difficulty and one task was randomly chosen for payment. 88% of subjects chose the correct move in an easier task if they had chosen the correct move in a more difficult task, conforming with the Guttman scale (Guttman, 1950). To more precisely classify subjects who do not conform with the Guttman scale, the farsightedness score was calculated as follows:

- Score of 0 for players who failed in the easy task
- Score of 1 for players who solved the easy task but failed in the intermediate task

[19] We manipulated the task difficulty by changing payoff differences. Two types of payoff differences are present in each task: the actual difference in expected payoffs in favour of the correct move if the second node decision was taken into account ("farsighted payoff difference"), and the difference in the average of all payoffs in favour of the incorrect move if all payoffs were treated as being equally likely ("myopic payoff difference"). Task 1 was adapted from Bone et al. (2009), with a farsighted payoff difference of 125 ECU and the myopic payoff difference of 138 ECU. Task 2 made the correct choice more attractive by increasing the farsighted payoff difference to 175 ECU and at the same time reducing the myopic payoff difference to 125 ECU. In task 3 the distribution of payoffs was the same for both actions, reducing the myopic payoff difference to 0, but the payoffs were arranged in a way that increased the farsighted payoff difference to 300 ECU.

- Score of 2 for players who solved the easy and intermediate tasks but failed in the hard task
- Score of 3 for players who solved all tasks

Using the number of correct answers to measure farsightedness has two drawbacks: classification is coarse and farsightedness may be confounded with other traits such as intelligence. To overcome these downsides a second measure of farsightedness was constructed using a ratio of time spent in the first decision node relative to time spent in the second node. Time ratio reflects the willingness to think ahead (in node 1) instead of afterwards (in node 2), and is a continuous variable, capturing individual differences more precisely. Still, both measures are closely related (see table A.4 in Appendix 4.B). At the end of experiment we also administered a self-reported temporal orientation questionnaire that measured time perspective, anticipation of future consequences and the ability to plan ahead (Steinberg et al., 2009).

4.3.3.3 Results: Personal Characteristics and R Choice in *LOW*

We identify the types of players who are more likely to choose R by comparing the characteristics of subjects who choose R and then running a probit regression with four specifications. In the first two specifications the dependent variable is set to 1 for players who chose R and disclosed it in round 14 and is set to 0 otherwise. We favour this specification because observed R levels in round 14 accurately predict transition success, but it suffers from a problem of independence, as observations in round 14 are affected by the shared history in block 1. To account for the different histories we add the total number of observed R and S in block 1 as control variables. Furthermore, we replicate the analysis in models 3 and 4 using choices in round 1, which are independent. The difference between model 1 and 2 and between model 3 and 4 lies in the measurement of farsightedness: *farsighted score* is based on the number of correctly solved individual choice tasks, while *time ratio* measures time spent in the first decision node relative to time spent in the second decision node.

Both the average scores in table 4.6 and the probit regressions in table 4.7 show that players who are more farsighted and who report higher beliefs are more likely to choose R in rounds 1 and 14. Table 4.6 shows a large difference in the magnitude of farsightedness between players who choose R and S, although Mann-Whitney U test fails to reject the null hypothesis when the farsightedness score is used. Table 4.7 reports the marginal effects at the means of covariates for the four probit models. Farsightedness is positively related to the tendency

Table 4.6: Average scores of subjects who choose R and S in rounds 1 and 14, *LOW* treatment. Significance levels reflect p-values of a Mann-Whitney U test, where the null hypothesis is that characteristics of subjects who choose S are drawn from the same distribution as the characteristics of those who chose R.

	Round 1		Round 14	
	S	R	S	R
Farsightedness: score	0.75	1.25	0.78	1.28
Farsightedness: time ratio	1.53	1.69	1.34*	1.89*
Risk aversion	6.38	5.60	6.06	5.83
Pro-sociality	20.75	24.52	23.34	22.35
Beliefs	30.75***	66.80***	35.44**	52.67**
Average income	63.54	67.99	62.68	69.35
N	16.00	20.00	18.00	18.00

* p<0.10, ** p<0.05, *** p<0.01

Table 4.7: Probit models, where the dependent variable is set to 1 if players choose R and disclose it (1-4), or choose R (5-6). Controls for gender and age are included. Standard errors are heteroskedasticity-robust.

	(1)	(2)	(3)	(4)	(5)	(6)
	R14	R14	R1, R $\wedge$ D	R1, R $\wedge$ D	R1, R	R1, R
Farsighted score	0.153*		0.147		0.119	
	(1.67)		(1.62)		(1.17)	
Time ratio		0.289***		0.249***		0.107
		(3.08)		(2.86)		(0.88)
Risk aversion	-0.00779	-0.0233	-0.0835	-0.0838	-0.274**	-0.308**
	(-0.12)	(-0.34)	(-1.24)	(-1.13)	(-2.24)	(-2.28)
Pro-sociality	0.00279	0.00334	0.0106	0.0104	0.0210**	0.0220**
	(0.46)	(0.55)	(1.28)	(1.29)	(2.16)	(2.23)
Beliefs	0.00673	0.0117**	0.0104***	0.0120***	0.0191***	0.0204***
	(1.46)	(2.36)	(3.38)	(3.85)	(3.54)	(3.28)
Observed R	0.0448	0.0437				
	(1.40)	(1.40)				
Observed S	-0.0373	-0.0567				
	(-0.90)	(-1.34)				
Constant						
N	36	36	36	36	36	36
pseudo R^2	0.174	0.264	0.300	0.355	0.508	0.495

Marginal effects; t statistics in parentheses

(d) for discrete change of dummy variable from 0 to 1

* p<0.10, ** p<0.05, *** p<0.01

to choose R and disclose, and the significance level depends on the dependent variable and on the method used to measure farsightedness (p-values range from 0.002 to 0.131). In addition to statistical significance, we find a large economic effect: the predicted likelihood to choose R and disclose in round 14 is 30% for subjects with a farsightedness score of 0 and increases to 81% for those with a score of 3, holding other factors constant at their mean (Wald test shows that the difference is statistically significant, p-value = 0.0324). The difference between subjects with a score of 0 and those with scores 1 or 2 is of expected sign, but not significant at the 5% level. The ratio of decision times predicts choices better than the number of correct answers, as is evident from higher significance levels and a higher pseudo-R^2. The coefficient of stated beliefs is statistically significant in all specifications except for model 1.

The estimated coefficients for risk aversion and pro-social attitudes are of expected sign, but not statistically significant. But a player who chooses R because of pro-social or risk-loving preferences might not be willing to pay even a trivial amount to disclose the action, thus it would be more appropriate to evaluate the effect of these motives using an unconditional decision to choose R. In round 14 all players who choose R also disclose this action, therefore the estimated coefficients would be identical to models 1 and 2 in table 4.7. In round 1, however, some players do not disclose their actions, and the choice of the dependent variable affects results. When disclosure is not taken into account, pro-social and risk-loving individuals are significantly more likely to choose R, while the coefficients for farsightedness decrease and become insignificant (columns 5 and 6 in table 4.7). Thus it seems that players driven by pro-social or risk-seeking preferences choose R in round 1, but they do not necessarily disclose their actions, while farsighted players are more likely to choose R and disclose both in round 1 and in round 14.

Result 6. *Participants in LOW who are classified as more farsighted are significantly more likely to choose R and disclose this action in rounds 1 and 14. Risk-loving and pro-social participants are more likely to choose R in round 1.*

4.4 Discussion

Overall, we find evidence that there are players who act strategically, and there also are players who update beliefs based on observed feedback. We find that many players are willing to pay to make their actions observable, and those who choose R are much more likely to do so. Strategic motives seem to be the most likely explanation for such behavior as action disclosure provides no

monetary benefits in the current round. If disclosures were driven only by the outcomes in the current round, it would be hard to explain why the frequency of disclosures is decreasing over time and why less than 20% of players disclose their actions at the end of block 2 when it costs 2 ECU while 40% disclose at the start of block 2 when it costs 80 ECU (see figure 4.4). Some disclosures might be a result of mistakes, but a large gap between disclosures of R and S and a concentration of disclosures at the start of each block indicate that most disclosures are not mistakes. We also find that R levels and disclosures of R are higher and transitions to the efficient equilibrium are more frequent in *LOW*, just as predicted by the learning model with sophistication. Evidence for belief learning provides further support that choosing R following coordination on S could be part of a payoff maximizing strategy if future earnings are taken into account. In fact, players who choose R in *LOW* and disclose it in round 14 on average earn 69 ECU per round, compared to 63 ECU earned by other players. In *HIGH*, however, strategic teaching is not profitable: those who choose R and disclose on average make only 42 ECU, compared to 54 ECU made by others. Finally, we find that farsighted players are more likely to initiate revolts and disclose their actions, while risk and social preferences have a significant effect only on the R choices in the first round.

However, even as players seem to understand the benefits of having their actions observed, there are indications for a limited ability to make intertemporal trade-offs. First, higher action disclosure costs discourage action disclosure, but have only a small direct effect on R levels in round 14. A sophisticated player would be expected to compare the payoff of choosing R and disclosing with the payoff of choosing S and not disclosing, and choose R if the former is larger than the latter. Consequently, higher costs should reduce not only the tendency to disclose, but also the tendency to choose R. Yet we find that higher costs discourage only action disclosure, while average R levels are similar both in round 14 and in the rest of the experiment if differences in beliefs are taken into account.[20] Thus low disclosure costs help overcome coordination failure mostly through higher observed R levels that alter beliefs. The second issue is that 60% of players in *HIGH* who choose R do not disclose their actions. It may seem that sophisticated players should either choose R and disclose, or choose S and not disclose. One explanation for choosing R and not disclosing could be that such choices are driven by non-strategic motives, such as beliefs

[20] To test the treatment effect we used a probit regression with the chosen action as a dependent variable and a treatment dummy as the main independent variable while controlling for beliefs. We find that the estimated likelihood to choose R is about 8 percentage points higher in *LOW*, and marginally significant (p-value = 0.093), but the effect disappears if all rounds in which play has converged to the efficient equilibrium are removed from the analysis (p-value = 0.332).

or preferences. But sophisticated players could also choose R without disclosing even if R was not myopically optimal, as long as a sufficiently large probability is assigned to the event that exactly $\theta - 1$ other group members will choose R. In such an event one's decision would be pivotal, and a choice of R would raise R levels above the threshold, informing other players about the true R levels via their payoffs.

4.5 Concluding Remarks

This chapter addresses the social dilemma of inefficient conventions: a unilateral deviation from an inefficient convention is costly, even though a collective deviation would lead to a Pareto improvement. As an example of such inefficient conventions we use the failure to take political action, but our findings could be applied to other contexts too. An inefficient technology could become the standard if it exhibits steep returns to scale (Arthur, 1989); examples of such technological lock-in include QWERTY keyboard (David, 1985), inferior information technology (Shapiro and Varian, 1999) or light water nuclear reactors (Cowan, 1990). As older technologies are being superseded by new innovations, it would be desirable to design an environment that would allow the most efficient technology to become a standard. Our findings may also be applied to the problems of inefficient social customs (Akerlof, 1980) or inefficient economic and political institutions (North, 1990; Acemoglu, 2006).

To understand how the efficiency loss could be reduced, it is important to know what motivates individuals who deviate from an inefficient convention, and how such deviations could be facilitated. We address these questions by running a laboratory experiment to determine whether deviations from an inefficient convention are explained by strategic motives, and whether such deviations occur more often when players can disclose their actions at a lower cost. We find evidence supporting the hypothesis that deviations are motivated by strategic reasons and we also find that inefficient conventions are overcome only if the cost of action disclosure is sufficiently low. These findings may help design interventions to facilitate the diffusion of efficient conventions. For example, availability of information technology could make it easier to inform others about one's action both directly, through social media, and indirectly, through information shared by other parties. The physical location of the events might matter too, as a large central space should make it much easier to gauge the level of support, while also allowing information to be shared much more rapidly among the participants. However, before implementing any such interventions

more research has to be done to verify the external validity our findings and to understand the potential negative effects. For example, advances in information technology could crowd out physical protests or make it easier for the regime to use propaganda, censorship or surveillance (Morozov, 2012).

Appendix 4

4.A Figures

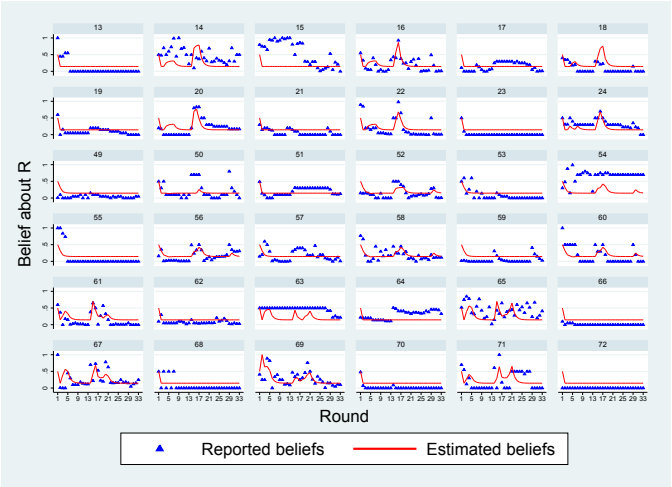


Figure A.1: Comparison of stated beliefs and the estimated beliefs using a weighted fictitious play model with parameters that minimise MSD. Graphs by subject in *HIGH*.

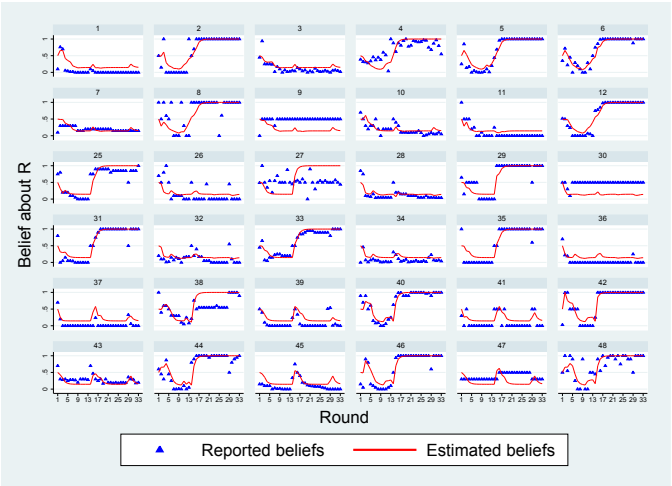


Figure A.2: Comparison of stated beliefs and the estimated beliefs using a weighted fictitious play model with parameters that minimise MSD. Graphs by subject in *LOW*.

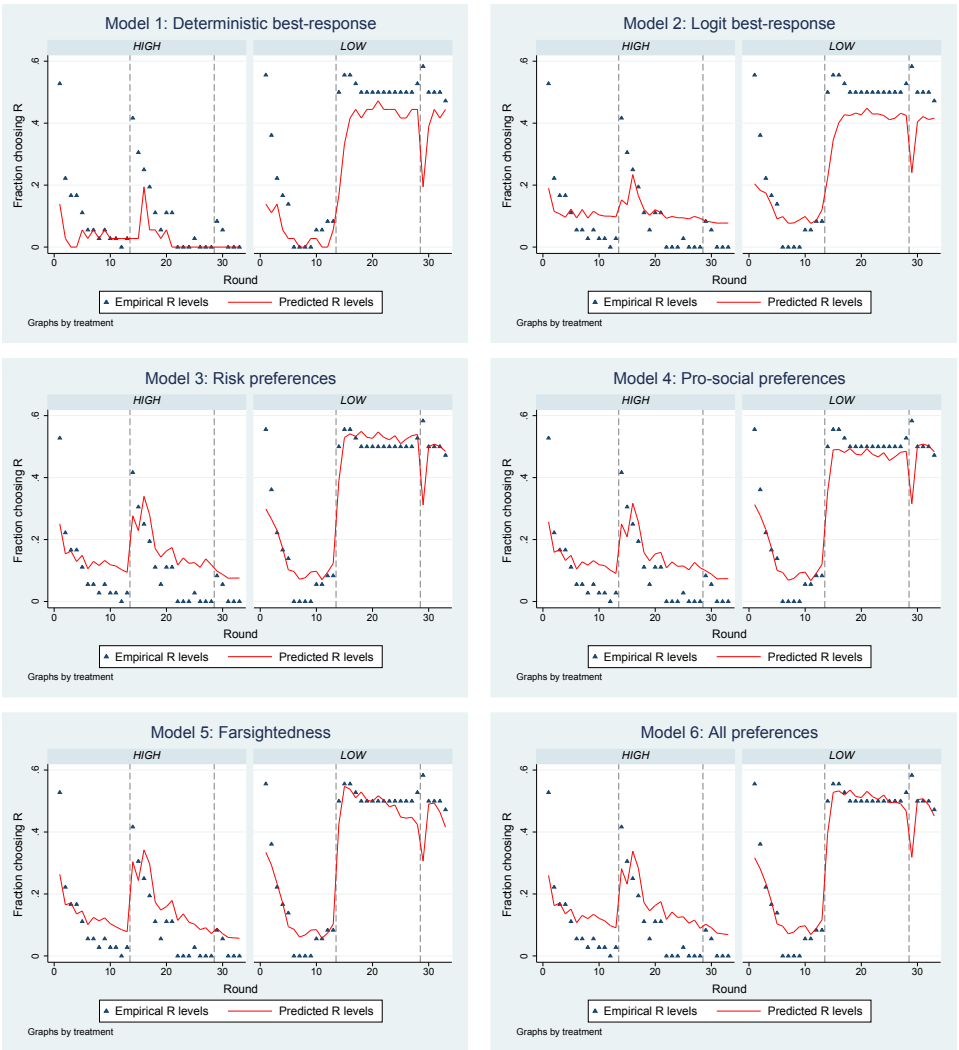


Figure A.3: Comparison of empirical R levels and R levels predicted by each of the six choice rules.

4.B Tables

Table A.1: Estimated treatment effect by round, calculated using the estimates of a multilevel GLS random effects regression and group level data. Standard errors clustered on group level.

Round	1	2	3	4	5	6	7	8	9	10	11
Coefficient	0.170	0.161	0.152	0.143	0.134	0.125	0.115	0.106	0.097	0.088	0.079
p-value	0.782	0.773	0.761	0.748	0.732	0.713	0.691	0.668	0.652	0.655	0.689
Round	12	13	14	15	16	17	18	19	20	21	22
Coefficient	0.070	0.060	1.429	1.573	1.716	1.860	2.003	2.146	2.290	2.433	2.577
p-value	0.747	0.807	0.228	0.183	0.146	0.116	0.092	0.073	0.058	0.047	0.038
Round	23	24	25	26	27	28	29	30	31	32	33
Coefficient	2.720	2.864	3.007	3.151	3.294	3.438	2.900	2.900	2.900	2.900	2.900
p-value	0.031	0.026	0.022	0.019	0.016	0.015	0.017	0.019	0.021	0.024	0.028

Table A.2: Time trends: estimated p-values for the coefficient of the number of rounds left.

Block	*HIGH* treatment	*LOW* treatment
1	0.0001	0.0000
2	0.0037	0.7351
3	0.0059	0.0410

Table A.3: Random-effects panel data probit regression. Standard errors clustered on the group level (12 clusters). Dependent variable is equal to 1 if the action is disclosed.

	All data	Decisions after choosing R
Treatment = *LOW*	1.128***	1.227***
	(5.25)	(3.33)
Action = R	1.742***	
	(15.48)	
Round	-0.0517***	-0.0554***
	(-9.55)	(-7.63)
Constant	-2.011***	-0.192
	(-10.65)	(-0.70)
Observations	2376	545

t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

Table A.4: Comparison of the two measures used to measure farsightedness: farsightedness score is based on the number of correct answers, time ratio is calculated as the ratio of time spent in node 1 relative to time spent in node 2.

Farsightedness score	Average time ratio	N
0	1.20	25
1	1.35	21
2	1.91	18
3	2.50	8
Total	1.57	72

Table A.5: Three-level nested random effects model and two-level random effects GLS model (12 clusters). Dependent variable: stated probability that a randomly chosen group member chose R (from 0 to 100). Standard errors are robust to correlation within the observations of each group. Observations from rounds in which groups are converged to an equilibrium are dropped (a group is converged if no player observed two different actions being chosen in the previous round)

	Random effects GLS		Nested random effects	
	Two-period lags	One-period lags	Two-period lags	One-period lags
Observed R in (t-1)	6.723***	6.815***	7.074***	7.051***
	(4.96)	(3.97)	(6.13)	(3.87)
Observed R in (t-2)	0.772		1.447	
	(0.36)		(0.66)	
Observed S in (t-1)	-3.950*	-3.019	-4.591**	-3.577*
	(-1.87)	(-1.36)	(-2.40)	(-1.91)
Observed S in (t-2)	1.110		1.083	
	(0.36)		(0.36)	
Actual R level in (t-1)	5.120**	6.680***	4.767***	6.508***
	(2.40)	(5.67)	(3.37)	(6.52)
Actual R level in (t-2)	1.857		1.231	
	(0.64)		(0.48)	
Treatment: 1 = *LOW*	-6.596*	-1.813	-6.148*	-1.633
0 = *HIGH*	(-1.66)	(-0.44)	(-1.89)	(-0.40)
Constant	17.51***	16.03***	17.51***	16.09***
	(3.03)	(3.22)	(3.44)	(3.37)
Observations	192	270	192	270

t statistics in parentheses

* p<0.10, ** p<0.05, *** p<0.01

4.C Alternative Explanations for the Treatment Difference

Treatment	Farsightedness score	Risk preference	SVO angle	Female	Age
High	1.22	5.92	16.85	0.58	21.06
Low	1.03	5.94	22.85	0.47	20.97
Range	0–3	2–10	-7.82 – 53.37	0–1	18–29

Table A.6: Descriptive statistics and their range by treatment

First, participants in the two treatments could differ in terms of their personal characteristics, which would lead to different responses to beliefs. To evaluate this possibility we compare the levels of farsightedness, risk and social preferences and demographic data, which were elicited at the end of the experiment (see section 4.3.3.1 for details). Table A.6 shows that most characteristics are similar, except for the SVO angle that measures pro-social preferences: subjects in *LOW* are more pro-social and the difference is marginally significant (Mann-Whitney U test p-value = 0.0615). If more pro-social participants were assigned to LOW, this difference could potentially explain the treatment effect because pro-social players should be more likely to choose R (see Appendix 4.E.4). However, differences in social preferences should increase R levels in all rounds, while we find that the treatment difference is very small at the start of block 2, but grows over time. Alternatively, choices in the SVO task could have been influenced by outcomes in the main part of the experiment. Social preference elicitation task was conducted at the end of the experiment and subjects from groups that coordinated on the efficient equilibrium might have acted in a more pro-social way, for motives such as reciprocity. Note that differences in social preferences cannot be explained by direct reciprocity, as subjects were never matched with their group members from the main part of the experiment, but they might be explained by generalised reciprocity (see Stanca, 2009, for evidence of generalised reciprocity).

Second, the process that drives the treatment difference could be reinforcement learning instead of belief learning. If this was the case, separate attractions could be formed for each of the four strategies in a stage game: 'choose R and disclose it' (R, D), 'choose R and not disclose' (R, ND), 'choose S and disclose' (S, D) and 'choose S and not disclose' (S, ND). If at the start of block 2 there are no treatment differences between the frequencies with which each strategy is chosen, payoffs to players who chose (R, D) will be lower in *HIGH* than in *LOW*. Under reinforcement learning the attraction of (R, D) would be reduced more strongly in *HIGH* than in *LOW*. If few players choose (S, D), this would

reduce R levels in *HIGH* more than in *LOW*. However, there is no evidence that large losses discourage subsequent R choice in *HIGH* at the start of block 2: players who choose (R, D) in round 14 make -75 ECU in that round, yet 83% of them continue choosing R in round 15; in comparison, those who choose (R, ND) earn 4 ECU, but only 22% of them choose R in round 15.

Another process that could explain the correlation between observed feedback and choices without the need to form beliefs is reciprocity: players who observe many group members choosing R could receive higher payoffs and they would be willing to reciprocate by choosing R more often in subsequent rounds. However, this explanation seems unlikely for several reasons. First, subjects are always informed about their payoffs, so observed R levels provide no additional information about one's earnings. Thus reciprocity should be induced by the actual R levels rather than the observed R levels. Second, R levels would affect earnings only if the participation threshold has been exceeded, and therefore there should be no difference in responses to higher observed R levels as long as the threshold has not been exceeded. Instead, we find that the effect of observed R levels is continuous: observing one additional group member choosing R has a positive effect, regardless of the R levels.[21] Third, reciprocity should be driven by subjects who chose R in the previous round because the payoffs of those who chose S are always fixed. In the probit regression we include a variable capturing the interaction between choosing R in the previous round and the observed R levels and find that the estimated coefficient is small and not significant ($p = 0.973$), suggesting that those who chose R and those who chose S are similarly affected by the observed history.

[21] p-values for the different numbers of observed R are: 0.008 (1 vs 0), 0.150 (2 vs 1), 0.068 (3 vs 2), 0.825 (4 vs 3), 0.000 (5 vs 4)

4.D Belief Elicitation Task

The belief elicitation task was designed to elicit precise and truthful probabilistic beliefs about the actions of other players. Images of the computer screen seen by the participants are reproduced in figures A.8 and A.9. Instructions are reproduced in Appendix 4.G.

To get a precise measure of beliefs we allowed subjects to choose any number between 0% to 100%, but elicitation was performed in two stages. The first stage displayed only the outcomes for all multiples of 5 between 0 and 100, and the second stage displayed the outcomes for the 11 strategies closest to the one chosen in the first stage (see figures A.8 and A.9 in Appendix 4.H).

To avoid hedging within one round, subjects were paid either their accumulated earnings from 33 rounds of the main part of the experiment, or they were paid for the belief elicitation tasks. In particular, at the end of the experiment a random draw was performed, so that with 80% probability a subject was paid for the main part of the experiment and with 20% probability a subject was paid for the belief elicitation task (Blanco et al., 2010, used a similar procedure to avoid hedging). The chances to be paid for the belief elicitation task were deliberately kept low to further reduce the temptation to hedge. We used a random-lottery incentive scheme by choosing one round for payment at the end of the experiment, a procedure that has been shown to avoid the income effect that appears if earnings are added up (Lee, 2008).

The second issue that may hinder the elicitation of truthful beliefs is risk preference: for example, risk-averse subjects may avoid reporting beliefs that are too extreme to lower the risk of receiving a low payment. It has been shown that deterministic payment schemes, such as the quadratic scoring rule (McKelvey and Page, 1990; Nyarko and Schotter, 2002), do not elicit truthful beliefs if subjects are not risk neutral (Schlag and van der Weele, 2013). Therefore we used a binarized scoring rule, which should elicit truthful beliefs irrespective of risk attitudes (Schlag and van der Weele, 2013). Experiments have also shown that a binarized scoring rule is less influenced by risk attitudes and is closer to the objective probabilities compared to the quadratic scoring rule (Hossain and Okui, 2013). The binarized scoring rule was implemented by paying subjects in lottery tickets instead of monetary earnings. If the belief elicitation task was chosen for payment, subjects participated in a lottery that yielded either 4000 ECU or 1000 ECU. Better performance in the belief elicitation task therefore did not affect the size of the earnings, but increased the chances to receive the higher reward.

4.E Risk and Social Preferences

4.E.1 Preliminaries

Assume that player $i \in N$ chooses action $a_i \in \{1, 0\}$ where $a_i = 1$ if R is chosen and $a_i = 0$ if S is chosen. Define the action profile of all players except for i by a vector $a_{-i} = \times_{j \in N \setminus \{i\}} a_j$. The total number of players other than i who chose action R can be calculated using a dot product: $\#R = a_{-i} \cdot a_{-i}$. The vector of profits for all players can be expressed as a function of the action taken by player i and of the total number of other players who choose R:

$$\Pi(a_i, \#R) = \begin{cases} \{100, 100 \cdot \mathbb{1}^{\#R}, 60 \cdot \mathbb{1}^{n-\#R-1}\} & \text{if } \#R + a_i \geqslant \theta \text{ and } a_i = 1 \\ \{60\ \ , 100 \cdot \mathbb{1}^{\#R}, 60 \cdot \mathbb{1}^{n-\#R-1}\} & \text{if } \#R + a_i \geqslant \theta \text{ and } a_i = 0 \\ \{5\ \ \ , 5\ \ \ \cdot \mathbb{1}^{\#R}, 60 \cdot \mathbb{1}^{n-\#R-1}\} & \text{if } \#R + a_i < \theta \text{ and } a_i = 1 \\ \{60\ \ , 5\ \ \ \cdot \mathbb{1}^{\#R}, 60 \cdot \mathbb{1}^{n-\#R-1}\} & \text{if } \#R + a_i < \theta \text{ and } a_i = 0 \end{cases}$$

where $\mathbb{1}^x$ is a vector of ones of length x. The first entry in the profit vector specifies the profit to player i, the second entry specifies the profits to $\#R$ players who chose R and the third entry specifies the profits for the $n - \#R - 1$ players who chose S. Define the payoff to player i as $\Pi_i(a_i, \#R)$ and the payoff vector of other players as $\Pi_{-i}(a_i, \#R)$.

Define subjective beliefs about the number of other group members who will choose R by a probability mass function $\sigma(\#R) : \{0, 1, \ldots, n-1\} \to [0, 1]$.

The expected utility of action a_i, conditional on subjective beliefs σ, is:

$$EU(a_i, \sigma) = \sum_{b \in \{0, \ldots, n-1\}} \sigma(b) u(\Pi(a_i, b))$$

Action R maximizes expected utility if a function that measures the difference between expected utilities is positive:

$$d(\sigma, u) = EU(R, \sigma) - EU(S, \sigma) = \sum_{b \in \{0, \ldots, n-1\}} \sigma(b)(u(R, b) - u(S, b)) \geqslant 0 \quad (4.2)$$

4.E.2 Risk Neutral, Self-interested Players

Self interested players would take into account only their own payoffs, specified by the first entry of a payoff vector $\Pi(a_i, \#R)$:

$$\Pi_i(a_i, \#R) = \begin{cases} 100 & \text{if } a_i = 1 \text{ and } \#R \geqslant \theta - 1 \\ 5 & \text{if } a_i = 1 \text{ and } \#R < \theta - 1 \\ 60 & \text{if } a_i = 0 \end{cases} \tag{4.3}$$

For risk neutral players utility is equal to payoffs: $u(\Pi_i(a_i, \#R)) = \Pi_i(a_i, \#R)$.

Substitute (4.3) into (4.2) and define $Pr(\#R \geqslant \theta - 1) = \sum_{b=\theta-1}^{n} \sigma(b)$ to get:

$$\begin{aligned} d(\sigma, u) \geqslant 0 \quad &\Leftrightarrow \\ Pr(\#R \geqslant \theta - 1) * 40 + (1 - Pr(\#R \geqslant \theta - 1)) * (-55) \geqslant 0 \quad &\Leftrightarrow \\ Pr(\#R \geqslant \theta - 1) \geqslant \frac{11}{19} \end{aligned} \tag{4.4}$$

A risk neutral player will choose R if and only if he assigns a probability of at least $\frac{11}{19}$ to the event that the participation threshold will be exceeded.

4.E.3 Risk Attitudes

Now suppose that players are not necessarily risk neutral. Define $v(x)$ satisfying $v'(x) > 0$ as the utility of receiving x, and use (4.3) to calculate utility:

$$u(a_i, \#R) = \begin{cases} v(100) & \text{if } a_i = 1 \text{ and } \#R \geqslant \theta - 1 \\ v(5) & \text{if } a_i = 1 \text{ and } \#R < \theta - 1 \\ v(60) & \text{if } a_i = 0 \end{cases} \tag{4.5}$$

Note that if $\sigma(\cdot)$ is degenerate, there is no strategic uncertainty and thus the strict monotonicity of $v(\cdot)$ ensures that the shape of the utility function will not affect the optimal choice. If $\sigma(\cdot)$ is not degenerate and $Pr(\#R > \theta - 1) \in (0, 1)$, R will be the optimal action if:

$$\begin{aligned} d(\sigma, u) \geqslant 0 \Leftrightarrow \\ Pr(\#R \geqslant \theta - 1) * (v(100) - v(60)) + (1 - Pr(\#R \geqslant \theta - 1)) * (v(5) - v(60)) \geqslant 0 \Leftrightarrow \end{aligned}$$

$$\frac{Pr(\#R \geqslant \theta - 1)}{1 - Pr(\#R \geqslant \theta - 1)} \frac{v(100) - v(60)}{v(60) - v(5)} \geqslant 0 \tag{4.6}$$

Consider two utility functions, $v_1(x)$ and $v_2(x)$ such that v_1 exhibits higher risk aversion $\left(\frac{-v_1''(x)}{v_1'(x)} > \frac{-v_2''(x)}{v_2'(x)}\right)$. Then $\frac{v_1(100)-v_1(60)}{v_1(60)-v_1(5)} < \frac{v_2(100)-v_2(60)}{v_2(60)-v_2(5)}$ (Pratt, 1964) and therefore $d(\sigma, v_1) < d(\sigma, v_2)$.

We conclude that players with lower risk aversion coefficients are weakly more likely to choose R, holding other factors constant.

4.E.4 Social Preferences

Suppose that utility depends not only on own earnings π_i, but also on the earnings of all other group members. In particular, assume that the utility function is separable and can be written the following way:

$$u_i(a_i, a_{-i}) = \Pi_i(a_i, \#R) + \alpha v(\Pi_{-i}(a_i, \#R))$$

where $v(\cdot)$ captures the effect of the payoffs received by other group members. We assume that $\nabla_{\mathbb{1}}(v(\pi_1, \pi_2, \ldots, \pi_{n-1})) \geqslant 0$, that is $v(\Pi_{-i}(a_i, \#R))$ does not decrease if there is no other player whose payoff decreased. A selfish player would be characterised by $\alpha = 0$, and would take into account only his own payoff. Players with $\alpha > 0$ would be pro-social and those with $\alpha < 0$ would be anti-social. If $\alpha > 0$, altruism could be modeled by setting $v(\cdot) = \sum(\cdot)$, maximin preferences could be modeled by setting $v(\cdot) = \min(\cdot)$. R is optimal if the following holds:

$$d(\sigma, u) =$$
$$\sum_{b \in \{0,\ldots,n-1\}} \sigma(b)(\Pi_i(R, b) + \alpha v[\Pi_{-i}(R, b)]) - \sum_{b \in \{0,\ldots,n-1\}} \sigma(b)(\Pi_i(S, b) + \alpha v[\Pi_{-i}(S, b)]) =$$
$$\sum_{b \in \{0,\ldots,n-1\}} \sigma(b)(\Pi_i(R, b) - \Pi_i(S, b)) + \alpha \sum_{b \in \{0,\ldots,n-1\}} \sigma(b)(v[\Pi_{-i}(R, b)] - v[\Pi_{-i}(S, b)])$$

The first term is not affected by social preferences because it includes only the earnings of player i. Notice that $\Pi_{-i}(R, b) \geqslant \Pi_{-i}(S, b)$, $\forall b \in \{1, \ldots, n-1\}$, that is payoffs received by other group members are at least as high when i chooses R as when i chooses S. Also, $\Pi_{-i}(R, \theta - 1) \geqslant \Pi_{-i}(S, \theta - 1)$, because if a person is pivotal, choosing R would lead to the threshold being exceeded and would therefore provide higher payoffs for those group members who chose R, compared to the case if S was chosen. Thus if a positive weight is given to the event that $\theta - 1$ others will choose R, as we have assumed, the second term would be increasing in α. The value of $d(\sigma, u)$ would be higher for players who are more pro-social, defined by higher values of α, thus for such players R would be optimal under a broader set of parameter values.

Note that the positive effect of pro-social preferences on R choice holds only if players make deterministic choices. If choices are stochastic and depend on the payoff difference, increased pro-sociality may decrease the payoff difference between R and S, making R less likely to be chosen.

4.F SVO Slider Measure Tasks

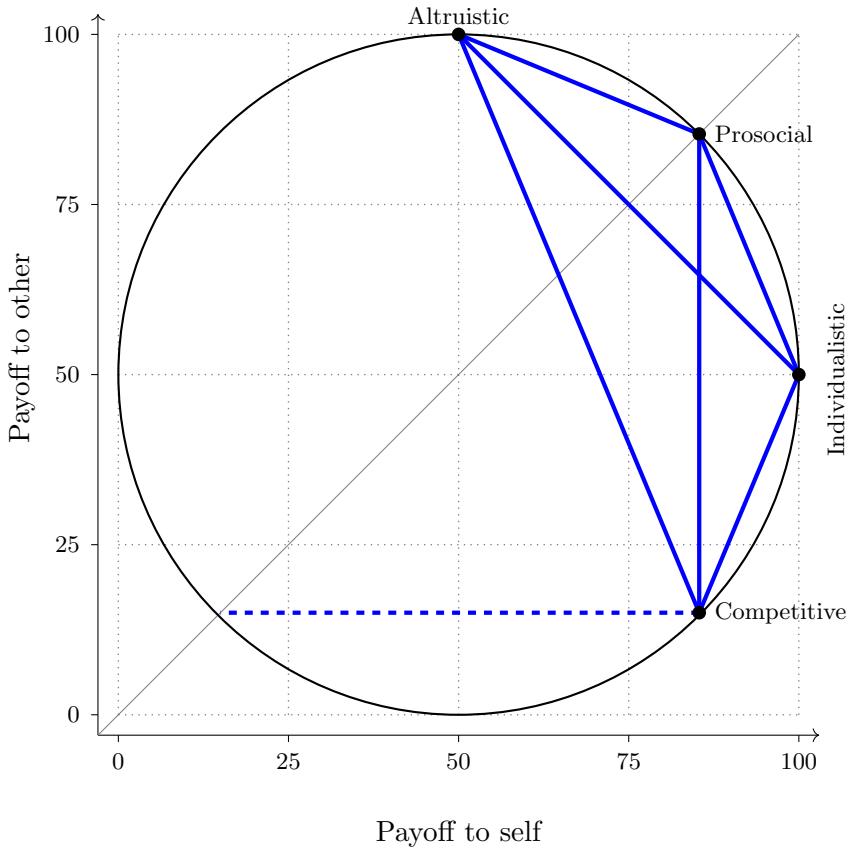


Figure A.4: Six allocation tasks of the Slider measure (solid lines) and the additional task (dashed line). Each line represents one allocation task, in which subjects choose one of the 9 allocations on the line.

4.G Instructions

Instructions were identical for treatments HIGH and LOW, except for the disclosure costs. We highlight this difference below.

INSTRUCTIONS

Welcome to the experiment. Please read the instructions carefully and take as much time as you need. The instructions are identical for all the participants with whom you will interact during this experiment.

If you have any questions please raise your hand. One of the experimenters will come to you and answer your questions. From now on communication with other participants is not allowed. If you do not conform to these rules you will be excluded from the experiment with no payment. Please also switch off your mobile phone at this moment.

In this experiment you can earn some money. How much you earn depends on your decisions and the decisions of the other participants. During the experiment we will refer to ECU (Experimental Currency Unit) instead of Euro. The total amount of ECU that you will have earned during the experiment will be converted into Euro at the end of the experiment and paid to you in cash confidentially. In this experiment the conversion rate that will be used to convert your ECU earnings into your Euro cash payment is: 250 ECU = 1 euro.

Overview

This experiment consists of multiple rounds. At the start of the experiment you will be randomly matched with 5 other participants. The group will remain constant throughout the experiment so you will interact with the same 5 other participants in all rounds. All other participants will face the same decision task as you, and you will not know the identity of each other. Every round consists of two tasks, a Decision Task and a Guess Task. You will do both tasks in every round, but you will be paid either for Decision Tasks or for the Guess Tasks. At the end of the experiment it will be randomly determined whether your payment is based on the Decision Tasks or on Guess Tasks. There is a chance of 4 out of 5 that your payment will be based on the Decision Tasks and a chance of 1 out of 5 that your payment will be based on the Guess Tasks.

Next we will describe what you have to do in each of the tasks.

Decision Task

- Your initial balance at the start of round 1 is 400 ECU.
- In each round you will have to choose either action A or action B.
- If you choose action B, your round income will be 60 ECU.
- If you choose action A, your round income depends on decisions of other participants in your group. If x or more others (not including yourself) choose A, your round income will be 100 ECU. If less than x others choose A, your round income will be 5 ECU. For example, if $x = 2$, your round income will be 5 ECU if 0 or 1 other participant chooses A and your round income will be 100 ECU if 2,3,4 or 5 other participants choose A.
- The rounds will be grouped into a certain number of blocks. At the start of a block you will be informed about the number of rounds that the block contains and the value of x in that block. The value of x will stay the same in every round of the block, and you will be reminded about the value of x in every round.
- Everyone will make decisions at the same time, so you will not know the decisions of other participants before making your own decision.

Table 1 summarizes your choices and possible round income in the Decision Task.

Your choice	**Your income if x or more other participants choose A**	**Your income if less than x other participants choose A**
A	100	5
B	60	60

Table 1: Round income for the Decision Task (in ECU)

Making your action public or hidden

- After choosing an action in the Decision Task you will have to choose if you want to make this action public or if you want to make it hidden.
- If you make your action public, all other participants in your group will be informed about the action (A or B) that you chose in the Decision Task, but only after the Guess Task has been completed. If you make

your action hidden, other participants will not know your action in the Decision Task.

- Making your action public will cost you 80 ECU [2 ECU]. Making your action hidden will cost you 1 ECU. This amount will be subtracted from your round income. If your round income is not sufficient to cover these costs, the difference will be subtracted from the initial balance or from your income in other rounds.

- Making your action public or hidden will affect only the information that others see, but it will not affect their income. Likewise, your income will not be affected by whether other participants in your group made their actions public or hidden.

Guess Task

- In the Guess Task you will be asked to guess what action, A or B, some other participant in your group chose in the Decision Task. We ask you to report your Guess on a scale from 0 to 100, where 0 means that you are sure the other participant chose B and 100 means that you are sure that the other participant chose A. The participant whose action you have to guess (call him other participant) will be randomly chosen by the computer. All 5 other participants in your group have equal chances to be chosen.

- You will not earn any ECU from the Guess Tasks, but you will earn some lottery tickets. If your payment is based on the Guess Tasks, at the end of the experiment you will play a lottery in which you will receive either 4000 ECU or 1000 ECU. The probability to receive 4000 ECU will depend on the number of tickets that you received in one of the Guess Tasks. Which Guess Task is used will be determined randomly by the computer and every Guess Task has an equal chance of being selected.

- The exact way how the number of lottery tickets depends on your Guess and on the action of the other participant in the Decision Task is shown in Table 2. If you guess 0, you will receive 4000 ECU for sure if the other participant chose B and will receive 1000 ECU for sure if the other participant chose A. If you guess 100, you will receive 4000 ECU for sure if the other participant chose A and will receive 1000 ECU for sure if the other participant chose B. All other Guesses give a positive probability to receive either sum. Notice that as you increase your Guess, you receive

more tickets if the other participant chose A but receive less tickets if the other participant chose B.

- Notice that Table 2 lists the payoffs only for the multiples of five, but you are free to choose any number between 0 and 100. To make it easier to choose, we will first ask you to choose a multiple of five, and then allow you pick a more precise Guess by showing earnings for 11 numbers closest to the one you picked. For example, if in the first table you choose number 50, you will be given numbers 45-55 to choose from in the second table.

Your Guess	Number of lottery tickets that you will receive if other participant chose A in the Decision Task	Number of lottery tickets that you will receive if other participant chose B in the Decision Task
0	0.00	100.00
5	9.75	99.75
10	19.00	99.00
15	27.75	97.75
20	36.00	96.00
25	43.75	93.75
30	51.00	91.00
35	57.75	87.75
40	64.00	84.00
45	69.75	79.75
50	75.00	75.00
55	79.75	69.75
60	84.00	64.00
65	87.75	57.75
70	91.00	51.00
75	93.75	43.75
80	96.00	36.00
85	97.75	27.75
90	99.00	19.00
95	99.75	9.75
100	100.00	0.00

Table 2: Number of lottery tickets received in the Guess Task

- Your choice in the Guess Task only affects your payoff. It does not affect the payoffs of other participants.

What you will see at the end of a round

After both the Decision Task and the Guess Task have been completed, you will see the following information on the screen:

- Your choice in the Decision Task
- The number of other participants in your group who chose action A and made it public
- The number of other participants in your group who chose action B and made it public

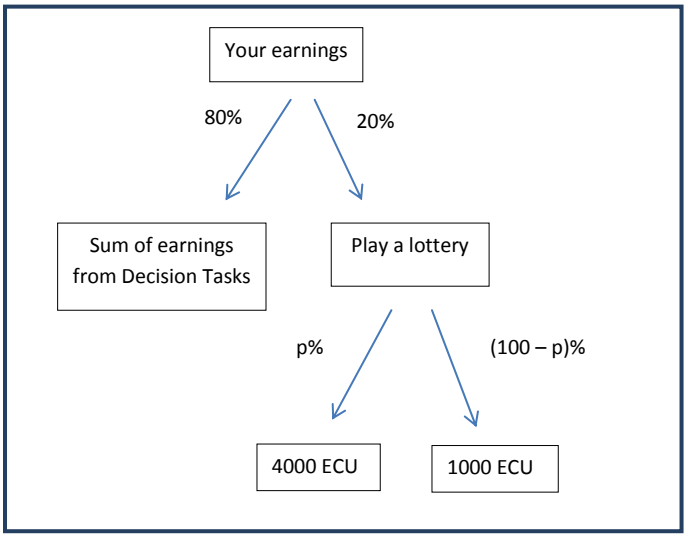


Figure 1. How your earnings from Part 1 will be determined. p is the number of lottery tickets that you got from one randomly chosen Guess Task.

- The number of other participants in your group who made their action hidden
- Your payoff from the Decision Task

You will not know the choices of participants who made their choices hidden.

Information about your choices and choices of other participants in previous rounds will always be available on the right side of your computer screen.

You will not receive any feedback about the Guess Task. The choice of a randomly selected participant and the number of lottery tickets that you earned in every round will be displayed at the end of the experiment.

How your cash earnings will be determined

At the end of the experiment it will be randomly determined whether you are paid for the Decision Tasks or for the Guess Tasks. There is a chance of 4 out of 5 that you will be paid for Decision Tasks and there is a chance of 1 out of 5 that you will be paid for Guess Tasks. Note that if you are paid for Decision Tasks, your choices in the Guess Task will not influence your final earnings. If you are paid for Guess tasks, you will not receive the earnings that you made in the Decision tasks.

If your payment is based on Decisions Tasks, your final earnings will be determined by adding your round income from all Decision Tasks to the initial balance of 400 ECU.

If your payment is based on Guess Tasks, your total earnings will be either 4000 ECU or 1000 ECU. Which sum you will receive will be determined by spinning a virtual Lottery Wheel. The Lottery Wheel will have two sectors, Red and Blue, and will randomly stop in one of them. If the Lottery Wheel stops in the Red sector, you will receive 4000 ECU; if it stops in the Blue sector, you will receive 1000 ECU. The size of the Red sectors will be determined by the number of lottery tickets that you received in one Guess Task. For example, if you earned p lottery tickets, $p\%$ of the lottery wheel will be colored Red and the remaining $(100 - p)\%$ will be colored blue. Therefore the probability that you will receive 4000 ECU will be $p\%$ and the probability that you will receive 1000 ECU will be $(100 - p)\%$. Which Guess Task is used to determine the size of the Red sector will be determined randomly by the computer and all Guess Tasks have equal probability to be selected. The spin of the virtual Lottery Wheel will be performed on your computer screen at the end of the experiment.

How it will be determined if you are paid for the Decision Tasks or for the Guess Tasks

At the end of the experiment we will ask you to choose a number between 1 and 5. Afterwards we will ask one participant in this room to randomly draw one card out of 5 cards, numbered from 1 to 5. The draw will be public, so everyone will be able to observe what number has been drawn. If the drawn number is the same as the number that you entered, your payment will be based on the Guess Tasks. If the drawn number is different from the one that you have entered, your earnings will be based on the Decision Tasks.

What you will see at the end of the experiment

At the end of the experiment you will be informed about the choices of all participants in all rounds and your earnings from all Decision Tasks and all Choice Tasks. You will see if you are paid for the Decision Tasks or for the Guess Tasks. If you are paid for the Guess Tasks, you will see the result of a virtual Lottery Wheel spin.

After you have made decisions in all rounds, but before seeing the final feedback screen you will have to do some additional tasks in which you will have a chance to earn more ECU. Instructions for these additional tasks will be shown on your computer screen. After these tasks you will see your total earnings from the experiment. We will also ask you to complete a short questionnaire. After completing the questionnaire, please stay seated until we ask you to come to receive your earnings from the experiment. Your earnings will be paid in cash

and in private.

If you have any further questions, please raise your hand now.

In order to ensure that everybody has understood the instructions, we will ask you to answer a few questions. Please click a button on your computer screen to start answering these questions.

4.H Screenshots

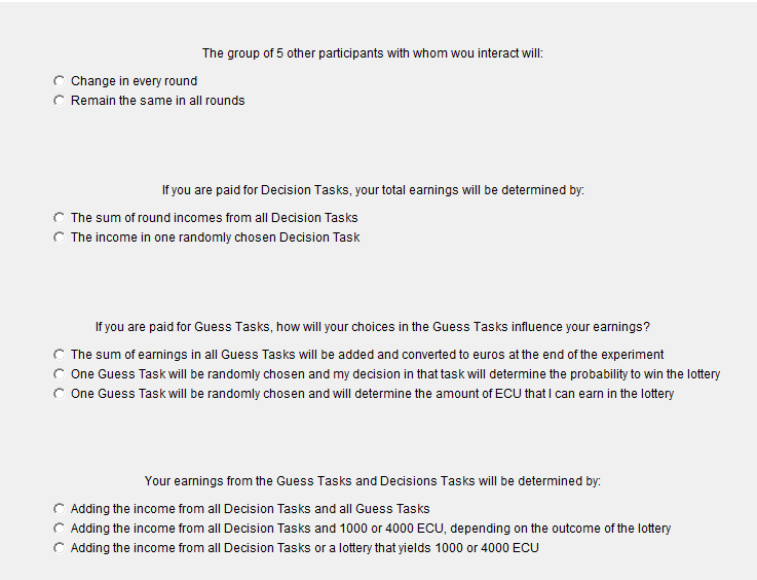


Figure A.5: First part of the questionnaire that had to be answered before subjects could start the experiment.

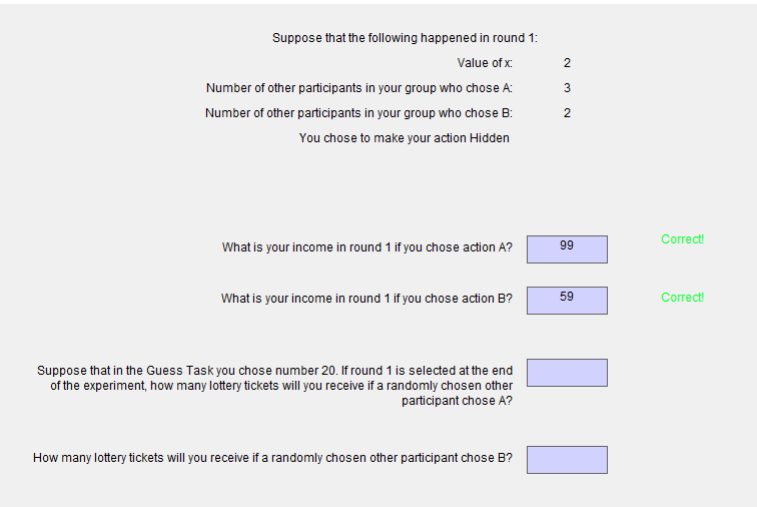


Figure A.6: Second part of the questionnaire that had to be answered before subjects could start the experiment. Values for the first question were generated randomly.

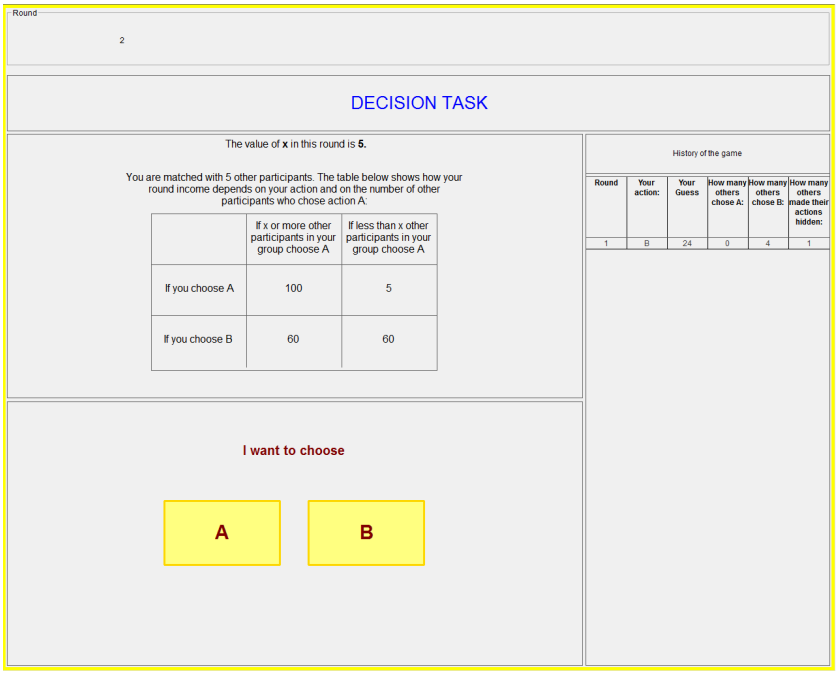


Figure A.7: Action choice stage

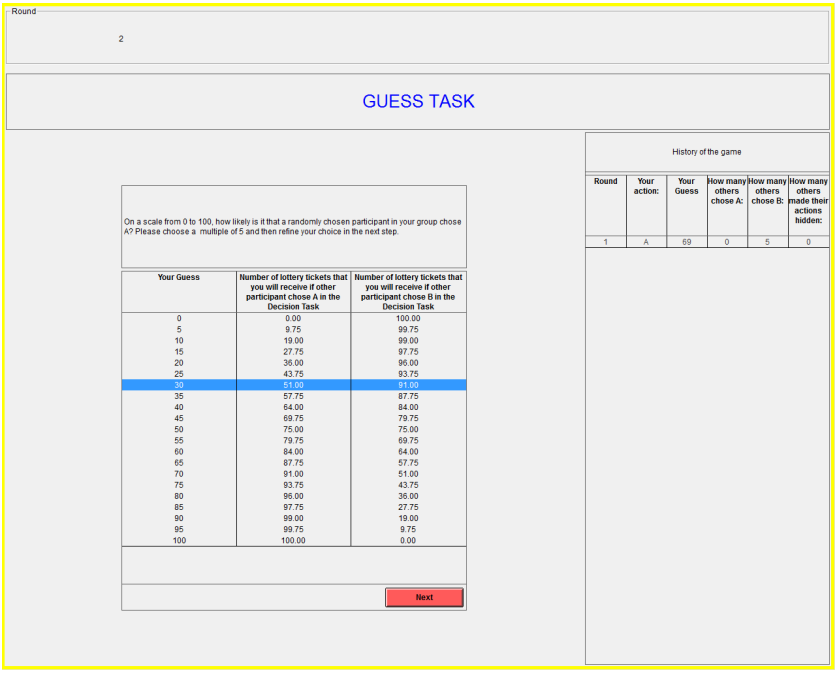


Figure A.8: First part of the belief elicitation task

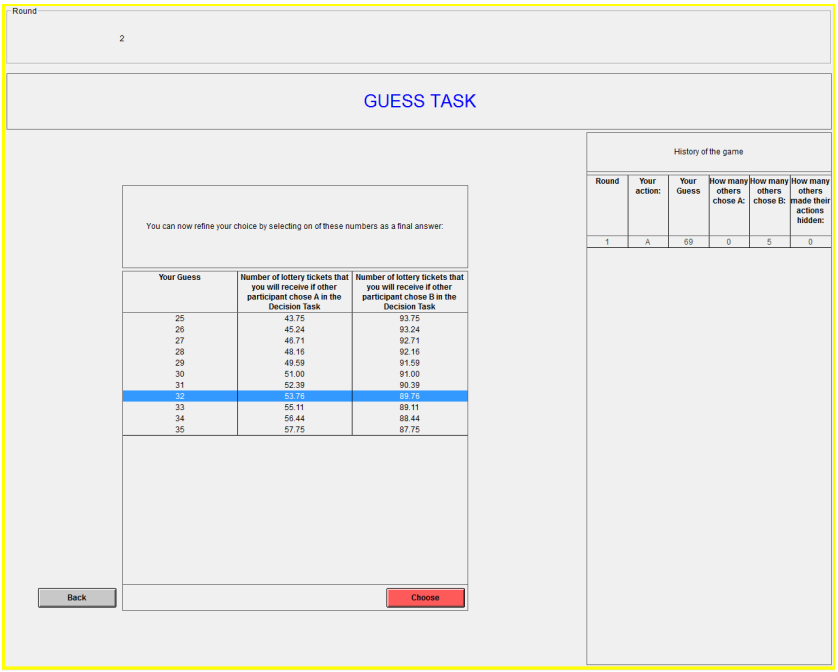


Figure A.9: Second part of the belief elicitation task

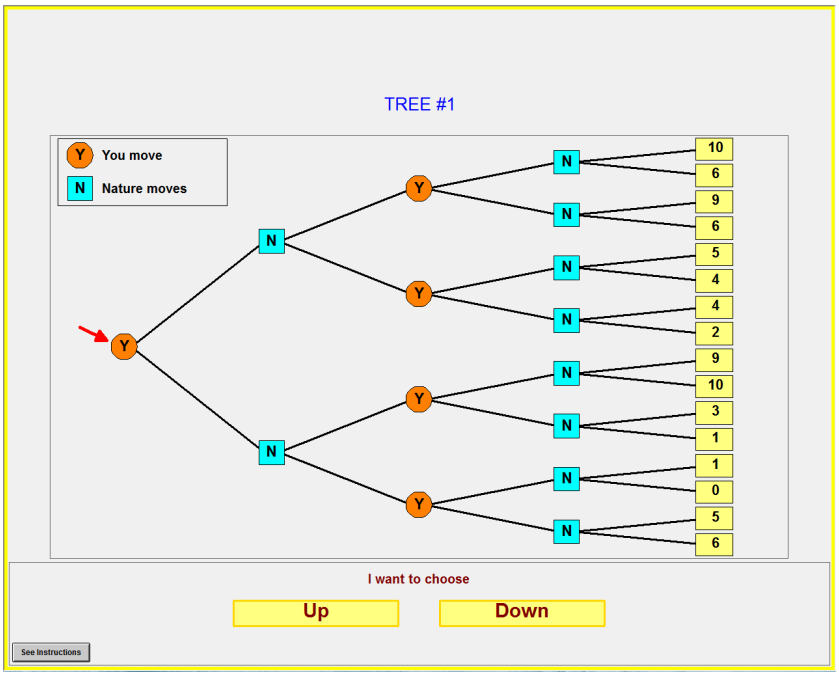


Figure A.10: Decision tree 1 from the task measuring farsightedness.

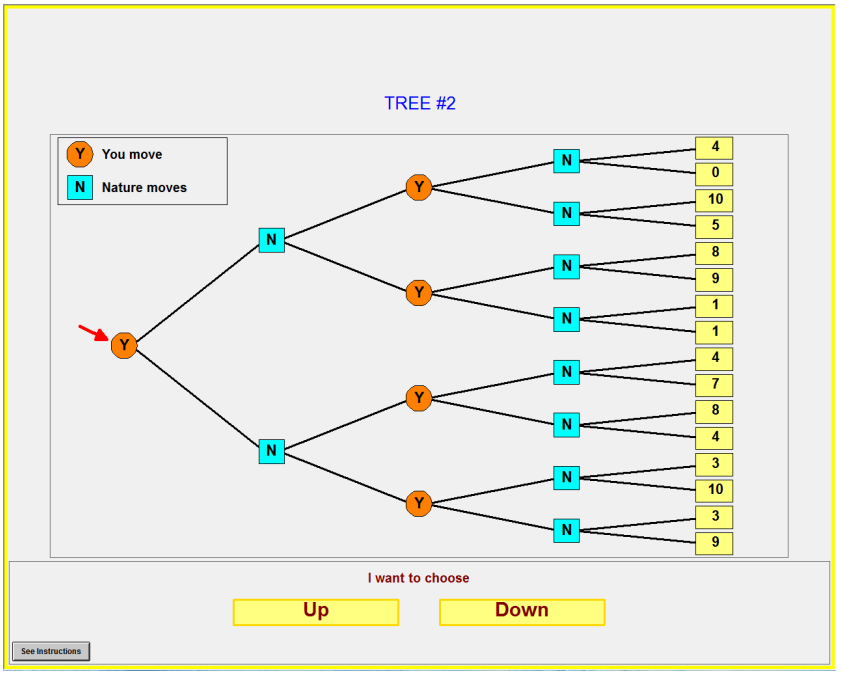


Figure A.11: Decision tree 2 from the task measuring farsightedness.

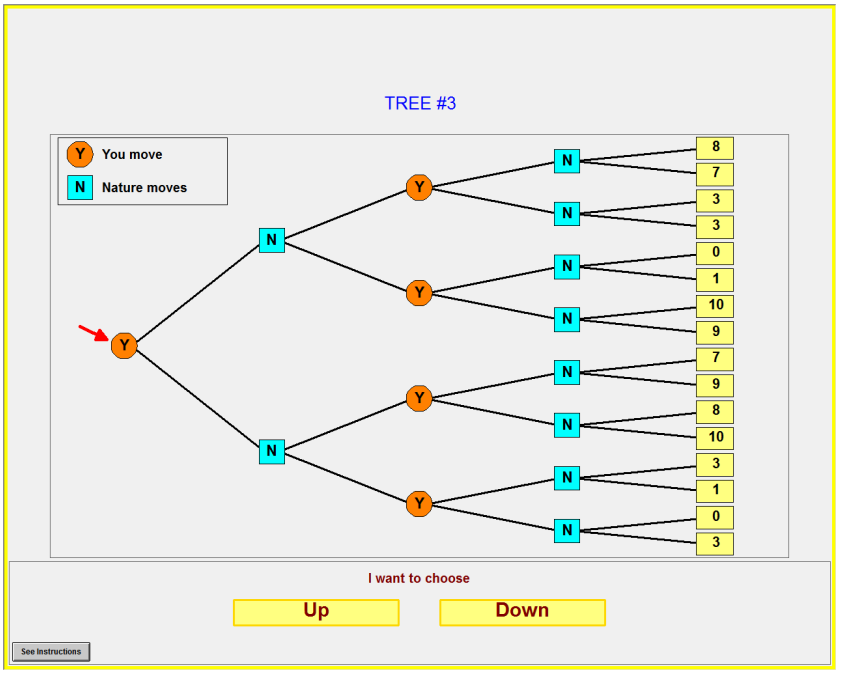


Figure A.12: Decision tree 3 from the task measuring farsightedness.

Chapter 5

Conclusion

The theme that unites the chapters in this thesis is feedback and its role in the learning process. The main question addressed here is whether better quality and quantity of feedback can help reduce the welfare loss that occurs when individuals are choosing dominated strategies in contests or when groups converge to an inefficient equilibrium in coordination games. If better feedback can help solve these problems, interventions that improve observed feedback could be easily implemented at a low cost, especially given the increasing use of the Internet, mobile devices and social media that make finding or sharing information easier than ever. Technological advances also make it easy for policy makers to shape the information received by the players, while other elements of the game, such as payoffs, order of moves or group size, are much harder to manipulate.

Overall, this thesis shows that better feedback can help reduce the efficiency loss. In chapter 2 we show that in contests dominated strategies are chosen less often and theoretical best response is chosen more often if the prize allocation is not probabilistic, if players know the action of the other participant and if there are opportunities to learn an optimal response. When one of these elements is missing, behavioral variation goes up and the explanatory power of Nash equilibrium goes down. The next challenge is to understand the exact mechanism that leads to the changes in explanatory power. One possibility is that matching players to computers and paying the expected value of the gamble lowers the risk and eliminates payoff consequences imposed on other players, which may lead to a difference in behavior if players have risk or social preferences. However, we show that such preferences cannot explain the large behavioral variation and low explanatory power typically found in contests and it is unlikely that such preferences are behind the treatment difference. An

alternative explanation is that lower complexity leads to more accurate feedback and makes it easier to find the best response. To further test this hypothesis one would need to design a contest in which feedback is improved without changing other elements of the game. One way to do so is by providing information about the payoffs that would have been generated by choosing a different action. This foregone payoff information should make it easier to directly compare different strategies and would also provide a larger quantity of information.

Chapter 3 shows that strategic players may deviate from an inefficient equilibrium solely for the information that such deviation conveys to other group members. A theoretical model presented in chapter 3 predicts that strategic players would be more likely to deviate from the inefficient equilibrium if they had a longer planning horizon or if the history of inefficient coordination was shorter, while player composition has an ambiguous effect. Overall, chapter 3 shows how game-theoretic approach can be combined with a learning model to narrow down the predictions of standard game-theoretic solution concepts. The chapter also demonstrates that a social dilemma that occurs due to convergence to an inefficient equilibrium can be solved even if all players are rational and self-interested. In other words, it is not necessary to assume social preferences because farsighted players who anticipate how others are learning could deviate from an inefficient convention for their own self-interest.

Players who deviate from inefficient conventions to improve future payoffs may be willing to pay to make their actions visible, and a policy maker may be able to facilitate the transition process by decreasing the cost of action disclosure. An experiment reported in chapter 4 finds evidence for both hypotheses: players are willing to disclose their actions even when doing so is costly, and the tendency to do so is much higher for those who deviate from an inefficient equilibrium. Players who have a longer planning horizon are also more likely to deviate from an inefficient convention and to disclose this action, while the effect of risk and social preferences is not significant. Finally, we find that half of the groups move from an inefficient to the efficient equilibrium when disclosure costs are low, but no groups do so when the costs are high, even though groups in both treatments experience a similar pattern of coordination failure. All groups that move to the efficient equilibrium remain there even when the participation threshold is increased back to its original level. These findings show that the efficiency loss which occurs from convergence to an inefficient equilibrium can be lowered by temporarily decreasing the participation threshold, as long as players are able to disclose their actions at a low cost.

Addendum: Valorization

This final chapter discusses the valorization opportunities of this thesis, in accordance with article 23.5 of the "Regulation governing the attainment of doctoral degrees" of Maastricht University. Knowledge valorization refers to the "process of creating value from knowledge, by making knowledge suitable and/or available for social (and/or economic) use and by making knowledge suitable for translation into competitive products, services, processes and new commercial activities" (adapted definition based on the National Valorization Committee). This section will discuss the broader implications and potential practical uses of the findings presented in this thesis. One should of course take these suggestions with caution and keep in mind that most findings come from experiments with small student samples in a laboratory setting. It is therefore not clear whether the interventions that work in experiment could be generalized. One should also carefully consider the potential negative side effects of the proposed suggestions.

The thesis uses theory and experiments to understand why an efficiency loss occurs and how it can be minimized in two types of situations. This chapter will briefly review the main findings and will show how these findings could be used by policy makers, managers or consumers.

Chapter 2 addresses the efficiency loss that occurs because of the choice of dominated strategies in a contest game. In contest experiments two parties compete for a fixed prize my making costly investments and the probability to receive the prize is proportional to the investments. There is robust experimental evidence that in such contests players often invest more than they should according to theoretical predictions. In some situations such overinvestments may be desirable while in others they may not be. If firms are competing for the the market share by investing in R&D, overinvestments may be beneficial to the society. Likewise, the chances that a worker will receive a bonus may be increasing in their effort and employers may attempt to design a system that maximizes the effort exerted by all workers. In other circumstances investments

may be wasteful, such as when politicians are competing in elections and when firms are investing in advertising or lobbying. Knowledge about the factors that affect investments could enable policy makers and managers to influence investments in a desirable way. For example, chapter 2 shows that when the actions of the other participant are know, subjects with higher cognitive abilities invest less than those with lower abilities.

However, chapter 2 shows that overinvestments are in large part a reflection of behavioral variability and noise that are present in contests. In fact, a Tullock contest is a good example of a game in which subjects make a clear mistake, which could not be explained by preferences or beliefs. More than half of the time players choose dominated strategies – these are strategies that generate lower earnings that some other strategy, regardless of what other players will choose – and there is high behavioral variability. Tools that could reduce the behavioral variability and increase the frequency of optimal strategies might also be used to improve decisions in other settings. We do find that in some treatments subjects with higher cognitive abilities perform better, but improving cognitive abilities is not an easy task and the computational complexity required to identify dominated strategies is so high that even those with the highest cognitive abilities are unable to perform it. A much easier way to avoid mistakes is by learning: subjects who track choices and payoffs, either their own or those made by others, could identify payoffs that lead to poor payoffs and avoid them. We find that when the complexity of the game is reduced, most people learn to play optimal strategies. For this to happen, three elements are of particular importance: subjects must be able to anticipate what others will choose, receive accurate feedback about their choices and they must be able to learn over time. If any of these elements is missing, the frequency of mistakes (i.e. choice of dominated strategies) and noise (i.e. behavioral variation) go up significantly.

These findings could be used to improve the quality of feedback that hinders learning. In fact, there already are programmes that aim to do that. Energy bills that compare current expenses to those in the past, as well to the performance of other households, have been shown to reduce expenditures (Allcott, 2011). Failure to understand probabilistic statements in medical decision making have lead to information being presented in formats that are easier to understand (Bornstein and Emler, 2001). Chapter 2 also shows that when conditions are suitable for learning, subjects are more likely to sample different actions when doing so is costless. Experimentation could therefore be another element that is necessary for learning. Decision makers who never

change their habits might never try better alternatives, and would therefore never learn about their superiority.

The second problem addressed in this thesis is the emergence of inefficient conventions, which may be implemented in situations with multiple alternative stable states. Multiplicity of stable states can occur because of strategic complementarity: for example, a person may cycle to work, recycle, pay taxes or avoid littering only if he expects that others will behave the same way. Incentives to coordinate on a common action drives choices towards one of the two stable states, and it is possible that an inefficient state will be implemented. Once this happens, an inefficient convention is hard to replace even if everyone involved would prefer a collective switch to the efficient convention. The understanding that the current situation is inefficient is not enough – participants must also expect sufficiently many others to deviate from an inefficient convention. For example, cycling might be dangerous if car drivers are not used to seeing cyclists on the road, making it hard to promote cycling even if everyone knew about its benefits and would willingly cycle to work as long as others do the same. To promote desirable actions a policy maker should therefore not only convince the public about its benefits, but also attempt to structure the environment in a way that increases the likelihood of efficiency-enhancing transitions. This thesis aims to increase our understanding about the ways how inefficient convention can be replaced, using both theory (chapter 3) and experiments (chapter 4).

Chapter 3 provides a framework that can be used to analyze situations in which an inefficient convention has been established. The model makes predictions about the types of symmetric equilibria that exist in a given game, and shows how these existence conditions can be calculated from the game parameters. It also shows how these existence conditions are affected by changes in parameter values. While some of the findings are intuitive – for example, a transition to the efficient state is more likely and faster if players have a longer planning horizon, other results are more surprising. If the number of sophisticated players is too large, chances of a transition may even decrease because each sophisticated player may find free-riding more attractive.

Chapter 4 studies inefficient conventions using an experiment. An important finding of this chapter is that farsighted players tend to deviate from inefficient conventions more often. This finding reinforces the results from chapter 3 and shows that a group can move from the inefficient to the efficient state if it has some sophisticated players. This could be potentially useful for managers, as coordination failure is not uncommon when output depends on the effort put in by the slowest member in the group. If a team gets stuck it in an

inefficient state, the presence of several farsighted players may help turn it around. Farsightedness could therefore be one of the criteria used to assign players into teams and allocate their roles.

One question that guided the design of the experiment presented in chapter 4 was the role of information technology in facilitating protest movements. Finding ways to help citizens in oppressed countries rise up against the regime could affect millions of people. If findings from chapter 4 could be taken at face value, uprisings could be facilitated by allowing citizens to make their actions observable. Information technology is certainly one possibility, as it enables the use of social media to communicate with each other at a low cost. But other tools that increase the observability could be effective too. Any accurate information about the current events might improve the participation levels, and such information is often lacking in the absence of free media. Protest movements might therefore be supported by providing better information or by improving the existing media channels. Interventions that seek to improve the quality and quantity of information that is available to the participants are rather easy to implement, which is one of the reasons why they were chosen to be studied in this thesis. It might not be possible to change the payoffs of the citizens, the order of moves or the group size, but interventions that provide better information or better tools for information sharing could be implemented easily and at a low cost.

Bibliography

Abbink, K., Brandts, J., Herrmann, B., and Orzen, H. (2010). Intergroup conflict and intra-group punishment in an experimental contest game. *The American Economic Review*, 100(1):420–447.

Acemoglu, D. (2006). A simple model of inefficient institutions. *The Scandinavian Journal of Economics*, 108(4):515–546.

Acemoglu, D. and Jackson, M. O. (2015). History, expectations, and leadership in the evolution of social norms. *The Review of Economic Studies*, 82(2):423–456.

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence. *The quarterly journal of economics*, pages 749–775.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9):1082–1095.

Andersson, M. and Iwasa, Y. (1996). Sexual selection. *Trends in Ecology and Evolution*, 11(2):53–58.

Angeletos, G.-M., Hellwig, C., and Pavan, A. (2007). Dynamic global games of regime change: Learning, multiplicity, and the timing of attacks. *Econometrica*, 75(3):711–756.

Armantier, O. and Treich, N. (2013). Eliciting beliefs: Proper scoring rules, incentives, stakes and hedging. *European Economic Review*, 62:17–40.

Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal*, pages 116–131.

Baharad, E. and Nitzan, S. (2008). Contest efforts in light of behavioural considerations. *The Economic Journal*, 118(533):2047–2059.

Bartling, B., Fehr, E., Marechal, M., and Schunk, D. (2009). Egalitarianism and competitiveness. *American Economic Review*, 99:93–98.

Baye, M., Kovenock, D., and Vries, C. D. (1993). Rigging the lobbying process: An application of the all-pay auction. *The American Economic Review*, 83(1):289–294.

Baye, M. R. and Hoppe, H. (2003). The strategic equivalence of rent-seeking, innovation and patent-race games. *Games and Economic Behavior*, 44(2):217–226.

Bednar, J., Chen, Y., Liu, T., and Page, S. (2011). Behavioral spillovers and cognitive load in multiple games: An experimental study. *Games and Economic Behavior*.

Best, M. L. and Wade, K. W. (2009). The internet and democracy global catalyst or democratic dud? *Bulletin of science, technology & society*, 29(4):255–271.

Blanco, M., Engelmann, D., Koch, A. K., and Normann, H.-T. (2010). Belief elicitation in experiments: is there a hedging problem? *Experimental Economics*, 13(4):412–438.

Blume, A. and Ortmann, A. (2007). The effects of costless pre-play communication: Experimental evidence from games with pareto-ranked equilibria. *Journal of Economic Theory*, 132(1):274–290.

Bone, J., Hey, J. D., and Suckling, J. (2009). Do people plan? *Experimental Economics*, 12(1):12–25.

Bornstein, B. H. and Emler, A. C. (2001). Rationality in medical decision making: a review of the literature on doctors? decision-making biases. *Journal of evaluation in clinical practice*, 7(2):97–107.

Brandts, J. and Cooper, D. J. (2006a). A change would do you good.... an experimental study on how to overcome coordination failure in organizations. *The American Economic Review*, pages 669–693.

Brandts, J. and Cooper, D. J. (2006b). Observability and overcoming coordination failure in organizations: An experimental study. *Experimental Economics*, 9(4):407–423.

Brosig, J. and Reiss, J. (2007). Entry decisions and bidding behavior in sequential first price procurement auctions: An experimental study. *Games and Economic Behavior*, 58(1).

Buchanan, J. and Tullock, G. (1962). *The Calculus of Consent: Logical Foundations of a Constitutional Democracy.* University of Michigan Press, Ann Arbor.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction.* Princeton University Press.

Camerer, C. F., Ho, T.-H., and Chong, J.-K. (2002a). Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *Journal of Economic Theory*, 104(1):137–188.

Camerer, C. F., Ho, T. H., Chong, J.-K., and Weigelt, K. (2002b). Strategic teaching and equilibrium models of repeated trust and entry games.

Cason, T., Masters, W., and R.M.Sheremeta (2010). Entry into winner-take-all and proportional-prize contests. *Journal of Public Economics*, 94:604–611.

Charness, G. and Rabin, M. (2002). Understanding social preferences with simple tests. *American Economic Review*, 117(3):817–869.

Chen, K. (2003). Sabotage in promotion tournaments. *Journal of Law, Economics and Organization*, 19(1):119–140.

Cheung, Y.-W. and Friedman, D. (1997). Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior*, 19(1):46–76.

Cooper, D. and Fang, H. (2008). Understanding overbidding in second price auctions: An experimental study. *Economic Journal*, 118:1572–1595.

Cooper, R., DeJong, D. V., Forsythe, R., and Ross, T. W. (1992). Communication in coordination games. *The Quarterly Journal of Economics*, 107(2):739–771.

Cowan, R. (1990). Nuclear power reactors: a study in technological lock-in. *The Journal of Economic History*, 50(03):541–567.

Crawford, V. P. (1995). Adaptive dynamics in coordination games. *Econometrica: Journal of the Econometric Society*, pages 103–143.

Crosetto, P., Weisel, O., and Winter, F. (2012). A flexible z-tree implementation of the social value orientation slider measure (murphy et al. 2011): Manual. Technical report, Jena Economic Research Papers.

David, P. A. (1985). Clio and the economics of QWERTY. *The American Economic Review*, pages 332–337.

De Mesquita, E. B. (2010). Regime change and revolutionary entrepreneurs. *American Political Science Review*, 104(03):446–466.

Devetag, G. (2003). Coordination and information in critical mass games: an experimental study. *Experimental economics*, 6(1):53–73.

Devetag, G. and Ortmann, A. (2007). When and why? a critical survey on coordination failure in the laboratory. *Experimental Economics*, 10(3):331–344.

Duffy, J. and Ochs, J. (2012). Equilibrium selection in static and dynamic entry games. *Games and Economic Behavior*, 76(1):97–116.

Edmond, C. (2013). Information manipulation, coordination, and regime change. *The Review of Economic Studies*, 80(4):1422–1458.

Ellison, G. (1997). Learning from personal experience: One rational guy and the justification of myopia. *Games and Economic Behavior*, 19(2):180–210.

Erev, I. and Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, pages 848–881.

Esteban, J. and Ray, D. (2011). A model of ethnic conflict. *The Journal of the European Economic Association*, 9:496–521.

Fallucchi, F., Renner, E., and Sefton, M. (2013). Information feedback and contest structure in rent-seeking games. *European Economic Review*, 64:223–240.

Fehr, E. and Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, pages 817–868.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178.

Fonseca, M. (2009). An experimental investigation of asymmetric contests. *International Journal of Industrial Organization*, 27(5):582–591.

Fudenberg, D. and Levine, D. K. (1998). *The theory of learning in games*, volume 2. MIT press.

Goodman, N. (1983). *Fact, fiction, and forecast.* Harvard University Press.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, pages 1–12.

Grimm, V. and Mengel, F. (2012). An experiment on learning in a multiple games environment. *Journal of Economic Theory*, 147(6):2220–2259.

Guttman, L. (1950). The basis for scalogram analysis. In Stouffer, S. A., Suchman, E. A., Lazarsfeld, P. F., Star, S. A., and Clausen, J. A., editors, *Measurement and prediction. [Studies in social psychology in World War II. Vol.4.]*, volume x. Princeton University Press, Princeton, NJ, US.

Hamman, J., Rick, S., and Weber, R. A. (2007). Solving coordination failure with "all-or-none" group-level incentives. *Experimental Economics*, 10(3):285–303.

Hassanpour, N. (2014). Media disruption and revolutionary unrest: Evidence from mubarak's quasi-experiment. *Political Communication*, 31(1):1–24.

Heinemann, F., Nagel, R., and Ockenfels, P. (2004). The theory of global games on test: experimental analysis of coordination games with public and private information. *Econometrica*, 72(5):1583–1599.

Heinemann, F., Nagel, R., and Ockenfels, P. (2009). Measuring strategic uncertainty in coordination games. *The Review of Economic Studies*, 76(1):181–221.

Hillman, A. L. and Katz, E. (1984). Risk-averse rent seekers and the social cost of monopoly power. *The Economic Journal*, 94(373):104–110.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655.

Hossain, T. and Okui, R. (2013). The binarized scoring rule. *The Review of Economic Studies*, 80(3):984–1001.

Huck, S., Jehiel, P., and Rutter, T. (2010). Feedback spillover and analogy based expectations: A multi game experiment. *Games and Economic Behavior*, 71(2):351365.

Huck, S. and Weizsäcker, G. (1999). Risk, complexity, and deviations from expected-value maximization: Results of a lottery choice experiment. *Journal of Economic Psychology*, 20(6):699–715.

Hyndman, K., Ozbay, E. Y., Schotter, A., and Ehrblatt, W. Z. (2012). Convergence: an experimental study of teaching and learning in repeated games. *Journal of the European Economic Association*, 10(3):573–604.

Hyndman, K., Terracol, A., and Vaksmann, J. (2009). Learning and sophistication in coordination games. *Experimental Economics*, 12(4):450–472.

Johnson, E., Camerer, C., Sen, S., and Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory*, 104(1):16–47.

Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgement of representativeness. *Cognitive Psychology*, 3(3):430–454.

Kalai, E. and Lehrer, E. (1993). Rational learning leads to nash equilibrium. *Econometrica*, pages 1019–1045.

Kalathil, S. and Boas, T. C. (2003). *Open networks, closed regimes: The impact of the Internet on authoritarian rule.* Carnegie Endowment.

Kedzie, C. (1997). Communication and democracy: Coincident revolutions and the emergent dictators. *RAND Dissertation.*

Keser, C., Suleymanova, I., and Wey, C. (2012). Technology adoption in markets with network effects: Theory and experimental evidence. *Information Economics and Policy*, 24(3-4):262–276.

Kogan, S., Kwasnica, A. M., and Weber, R. A. (2011). Coordination in the presence of asset markets. *American Economic Review*, 101(2):927–947.

Konrad, K. (2009). *Strategy and Dynamics in Contests.* Oxford University Press.

Lee, J. (2008). The effect of the background risk in a simple chance improving decision model. *Journal of Risk and Uncertainty*, 36(1):19–41.

Lewis, D. (1969). *Convention: A philosophical study.* John Wiley & Sons.

Lim, W., Matros, A., and Turocy, T. L. (2014). Bounded rationality and group size in tullock contests: Experimental evidence. *Journal of Economic Behavior & Organization*, 99:155–167.

Lohmann, S. (1993). A signaling model of informative and manipulative political action. *The American Political Science Review*, 87(2):319.

Lohmann, S. (1994). The dynamics of informational cascades: The monday demonstrations in leipzig, east germany, 1989-91. *World Politics*, 47(01):42–101.

Lohmann, S. (2000). Collective action cascades: An informational rationale for the power in numbers. *Journal of Economic Surveys*, 14(5):655–684.

Lynch, M. (2011). After egypt: The limits and promise of online challenges to the authoritarian arab state. *Perspectives on Politics*, 9(02):301–310.

Mailath, G. J. (1998). Do people play nash equilibrium? lessons from evolutionary game theory. *Journal of Economic Literature*, pages 1347–1374.

Mak, V. and Zwick, R. (2010). Investment decisions and coordination problems in a market with network externalities: An experimental study. *Journal of Economic Behavior & Organization*, 76(3):759–773.

McKelvey, R. D. and Page, T. (1990). Public and private information: An experimental study of information pooling. *Econometrica*, 58(6):13–21.

Meier, P. (2011). *Do liberation technologies change the balance of power between repressive states and civil society?* PhD thesis, The Fletcher School of Law and Diplomacy.

Mengel, F. (2014). Learning by (limited) forward looking players. *Journal of Economic Behavior & Organization*, 108:59–77.

Milgrom, P. and Roberts, J. (1991). Adaptive and sophisticated learning in normal form games. *Games and economic Behavior*, 3(1):82–100.

Millner, E. and Pratt, M. (1989). An experimental investigation of efficient rent-seeking. *Public Choice*, 62(2):139–151.

Monderer, D. and Shapley, L. S. (1996). Fictitious play property for games with identical interests. *journal of economic theory*, 68(1):258–265.

Morozov, E. (2012). *The net delusion: The dark side of Internet freedom*. PublicAffairs.

Mourtada, R. and Salem, F. (2011). Civil movements: The impact of facebook and twitter. *Arab Social Media Report*, 1(2):1–30.

Murphy, R. O., Ackermann, K. A., and Handgraaf, M. J. (2011). Measuring social value orientation. *Judgment and Decision Making*, 6(8):771–781.

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge university press.

Nyarko, Y. and Schotter, A. (2002). An experimental study of belief learning using elicited beliefs. *Econometrica*, 70(3):971–1005.

Offerman, T., Sonnemans, J., and Schram, A. (1996). Value orientations, expectations and voluntary contributions in public goods. *The Economic Journal*, 106(437):817.

Perea, A. (2007). A one-person doxastic characterization of nash strategies. *Synthese*, 158(2):251–271.

Peters, E., Dieckmann, N., Dixon, A., Hibbard, J., and Mertz, C. (2007). Less is more in presenting quality information to consumers. *Medical Care Research and Review*, 64(2):169–190.

Potters, J., De Vries, C., and Van Winden, F. (1998). An experimental examination of rational rent-seeking. *European Journal of Political Economy*, 14(4):783–800.

Rabin, M. and Weizsäcker, G. (2009). Narrow bracketing and dominated choices. *American Economic Review*, 99:1508–1543.

Rapoport, A. (1985). Provision of public goods and the MCS experimental paradigm. *The American Political Science Review*, 79(1):148.

Rapoport, A. and Eshed-Levy, D. (1989). Provision of step-level public goods: Effects of greed and fear of being gypped. *Organizational Behavior and Human Decision Processes*, 44(3):325–344.

Romero, J. (2011). The effect of hysteresis on equilibrium selection in coordination games. Technical report, Purdue University, Department of Economics.

Ruffle, B., Weiss, A., and Etziony, A. (2010). Coordination and critical mass in a network market: An experimental investigation. *Bar-Ilan University Department of Economics Research Paper*, (2010-03).

Schipper, B. C. (2011). Strategic control of myopic best reply in repeated games. *Available at SSRN 1804667*.

Schlag, K. H. and van der Weele, J. J. (2013). Eliciting probabilities, means, medians, variances and covariances without assuming risk neutrality. *Theoretical Economics Letters*, 3:38.

Schmitt, P., Shupp, R., Swope, K., and Cadigan, J. (2004). Multi period rent-seeking contests with carryover: Theory and experimental evidence. *Economics of Governance*, 5:187–211.

Shannon, C. (1948). The bell system technical journal 27. *A Mathematical Theory of Communication*, pages 379–423.

Shapiro, C. and Varian, H. R. (1999). *Information rules: a strategic guide to the network economy.* Harvard Business School Press, Boston, MA.

Sheremeta, R. (2010). Experimental comparison of multi-stage and one-stage contests. *Games and Economic Behavior*, 68(2):731–747.

Sheremeta, R. (2011). Contest design: an experimental investigation. *Economic Inquiry*, 49(2):573–590.

Sheremeta, R. (2013). Overbidding and heterogeneous behavior in contests. *Journal of Economic Surveys*, 27:491–514.

Shirky, C. (2011). The political power of social media-technology, the public sphere sphere, and political change. *Foreign Affairs*, 90:28.

Shurchkov, O. (2013). Coordination and learning in dynamic global games: experimental evidence. *Experimental Economics*, 16(3):313–334.

Sonnemans, J., Schram, A., and Offerman, T. (1998). Public good provision and public bad prevention: The effect of framing. *Journal of Economic Behavior & Organization*, 34(1):143–161.

Stanca, L. (2009). Measuring indirect reciprocity: Whose back do we scratch? *Journal of Economic Psychology*, 30(2):190–202.

Steinberg, L., Graham, S., O'Brien, L., Woolard, J., Cauffman, E., and Banich, M. (2009). Age differences in future orientation and delay discounting. *Child development*, 80(1):28–44.

Szymanski, S. (2003). The economic design of sporting contests. *Journal of Economic Literature*, 41:1137–1187.

Terracol, A. and Vaksmann, J. (2009). Dumbing down rational players: Learning and teaching in an experimental game. *Journal of Economic Behavior & Organization*, 70(1-2):54–71.

Tufekci, Z. and Wilson, C. (2012). Social media and the decision to participate in political protest: Observations from tahrir square. *Journal of Communication*, 62(2):363–379.

Tversky, A. and Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business*, pages S251–S278.

Van Huyck, J. B., Battalio, R. C., and Beil, R. O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review*, 80(1):234–248.

Van Huyck, J. B., Battalio, R. C., and Beil, R. O. (1991). Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. *The Quarterly Journal of Economics*, pages 885–910.

Van Huyck, J. B., Battalio, R. C., and Rankin, F. W. (2007). Evidence on learning in coordination games. *Experimental Economics*, 10(3):205–220.

Weibull, J. (1995). *Evolutionary Game Theory*. Cambridge: MIT Press.

Young, H. P. (1996). The economics of convention. *The Journal of Economic Perspectives*, 10(2):105–122.

Biography

Aidas Masiliūnas was born on 25 February, 1988, in Vilnius, Lithuania. He graduated from Vilnius Lyceum in 2006. He received a bachelor's degree in Economics from Vilnius University in 2010 after spending one semester at Hogeschool-Universiteit Brussel and another one at the National University of Singapore. He received his master's degree in International Economic Studies (Behavioural Economics) at Maastricht University in 2011. In the same year he joined the PhD programme in Economics at Maastricht University. Part of this time was spent visiting Essex University, Karlsruhe Institute for Technology and Universitat Autonoma de Barcelona. He will continue working as a post-doctoral researcher at Aix-Marseille School of Economics.