# *Decision-making in complex and uncertain environments*
*-*
# *Experimental studies in behavioral economics*

**Dissertation zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft**

**eingereicht an der Fakultät für Wirtschaftswissenschaften der Universität Regensburg**

**vorgelegt von: Johannes Moser**

Berichterstatter: Prof. Dr. Andreas Roider und Prof. Jörg Oechssler, Ph.D.

Tag der Disputation: 22. November 2019

Universität Regensburg

# Dissertation

## Decision-making in complex and uncertain environments
## -
## Experimental studies in behavioral economics

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Regensburg

**Erstbetreuer**: Prof. Dr. Andreas Roider
**Zweitbetreuer**: Prof. Jörg Oechssler, Ph.D.

Vorgelegt im Juli 2019 von:
**Johannes Moser**

# Acknowledgments

First of all, I want to thank my first supervisor Andreas Roider for his guidance and support during the work on my thesis. I always had the opportunity to talk to him about my research and I greatly benefited from his vast experience. Additionally, he also assisted me in a lot of organizational matters, such as applying for conferences or during the revision of a paper. Overall, Andreas always expected a high level of motivation and performance from me, but at the same time he made sure that I could fully concentrate on my research. This made it possible for me to write this thesis in its present form.

I further want to thank my colleagues from the University of Regensburg and my fellow students from the International Graduate Program "*Evidence-Eased Economics*" (EBE). In internal presentations as well as in discussions with my peers, I received great feedback and comments which helped me to further improve my own work. Also the atmosphere within this group was always very pleasant.

At this point, I also want to emphasize how much I profited from my membership in the International Graduate Program "*Evidence-Based Economics*" (EBE) of the Elite Network of Bavaria. The course program of the EBE was structured in a very sensible way and it was very much in line with my own research interests. Through the EBE I got in touch with a lot of interesting people and experts from all over the world. Additionally, I received a generous funding which made it possible for me to attend several national and international conferences, where I got the opportunity to present my own research. Furthermore, because of the funding, I was able to conduct the experiments which I used for my thesis.

Looking back, I can say that working on my thesis was a demanding, but rewarding journey. This thesis is the result of four interesting years of research in the field of behavioral and experimental economics and I hope that this work will give the reader a good insight into this field.

Johannes Moser, July 2019

*Für meine Eltern*

# Contents

# List of Figures

# List of Tables

# 1 Introduction

*Behind all things are reasons. Reasons can even explain the absurd. Do we have the time to learn the reasons behind the human being's varied behavior? I think not. Some take the time.*
– Log Lady, **Twin Peaks**

There is a large body of empirical evidence that people do not always behave according to game theoretic predictions in many economic or social environments. Possible deviations from standard-economic behavior can occur when individuals have either (i) non-standard beliefs, which are systematically biased, (ii) non-standard preferences, such as preferences for fairness, or (iii) when they engage in imperfect utility maximization, for example, because of limited attention and only consider salient alternatives in their choice sets (Rabin, 2002). This thesis addresses issues related to such forms of boundedly rational behavior and non-standard utility maximization. As a general term, for all deviations from standard-economic behavior mentioned above, I will use the expression "*non-standard decision-making*" in this thesis.

The field of behavioral economics deals with questions concerning non-standard decision-making, by combining elements from economics and psychology. This area of economic research has become increasingly important in recent years, which was also reflected in the Nobel Prize for Richard H. Thaler in 2017. Research in this field has been able to explain many empirical findings and economic puzzles, such as charitable giving (Ariely, Bracha, and Meier, 2009), blood donations (Mellström and Johannesson, 2008), overbidding in auctions (Crawford and Iriberri, 2007), or default effects in retirement savings (Madrian and Shea, 2001).

With my own work, I want to contribute to the vast experimental literature in behavioral economics. For this purpose, I will look at three different environments in which non-standard decision-making is commonly observed: common value auctions, public good games, and elections. The main aims of my research in these areas are the following: (i) finding the underlying channels of non-standard decision-making; (ii) investigating how the occurrence of potential cognitive mistakes can be reduced; (iii) checking whether mental shortcuts, or heuristics, are always irrational or whether they can be sometimes beneficial for individuals or groups and even lead to more socially efficient outcomes; (iv) furthermore, I am interested in the reasons and the economic consequences of cooperative and non-cooperative behavior.

With this, I would like to answer questions that have not yet been conclusively clarified in the literature so far. My empirical research strategy relies on laboratory and online experiments. The thesis will consist of four main chapters (Sections 2, 3, 4, and 5) following after the introduction.

The first chapter (Section 2) investigates the relationship between the winner's curse and mistakes in hypothetical thinking in the context of common value auctions.[1] There is evidence that bidders fall prey to the winner's curse because they fail to extract information from hypothetical events - like winning an auction. In this chapter, I investigate experimentally whether bidders in a common value auction perform better when the requirements for this cognitive issue - also denoted by contingent reasoning - are relaxed, leaving all other parameters unchanged. For my underlying research question, I used a lab experiment with two stages. In stage I, the subjects participate in a non-standard common value auction, called the wallet game, in which a naïve bidding strategy can lead to both winner's curse and loser's curse. In stage II, the subjects in the treatment group learn whether their initial bid was the winning bid or not and they get the opportunity to change this bid. In this sense, the bidders face the same decision problems as in stage I again, but the need for hypothetical thinking is reduced in stage II. More generally, I want to answer the question whether models focusing on inconsistent beliefs of individuals, like cursed equilibrium and level-$k$ (Eyster and Rabin, 2005; Crawford and Iriberri, 2007) or approaches concerning contingent reasoning on hypothetical future events (Charness and Levin, 2009; Ivanov, Levin, and Niederle, 2010; Esponda and Vespa, 2014) are more suitable for explaining the winner's curse.
In summary, the main questions I want to answer in this chapter are:

1. Do subjects in a common value auction perform better when they already learn ex ante, before the final payoffs are known, whether their bid is the winning bid or not?

2. Are approaches concerning mistakes in contingent reasoning or belief-based models like cursed equilibrium more accurate to explain the winner's curse?

The second chapter (Section 3) deals with another cognitive mistakes known as correlation neglect.[2] This bias refers to underestimating the degree to which various sources of information may be correlated. Typical examples for this issue are the news media or markets for financial assets. In an online experiment, I tested how the presence of correlation neglect affects the outcome in a collective voting problem. My work aims to supplement the existing experimental literature on correlation neglect by implementing a theoretical work of Levy and Razin (2015a). The experimental setup I used, provides an environment that includes the main features of their theoretical model: subjects with ideological preferences who receive an imperfect signal about the state of the world - which might differ from their

[1] This chapter is based on Moser (2019).
Available under: https://doi.org/10.1007/s11238-019-09693-9

[2] This chapter is based on Moser and Wallmeier (2019).

preferences - and a collective decision via majority vote. So the participants in my experiment are confronted with two conflicting objectives which they have to include in their maximization problem. Since political preferences are difficult to measure accurately, I focused on an approach where the choice of two alternatives depended on the risk preferences of the participants. On the one hand this makes the whole task more abstract, but on the other hand clear statements can be made about the effect of correlation neglect, when subjects are affected by personal preferences. To measure the effect of correlation neglect, the participants either received one signal, two perfectly correlated signals, or two independent signals about a state of the world in a between-subject design.
In summary, the main questions I want to answer in this chapter are:

1. Are subjects prone to correlation neglect in a collective decision problem - even though the correlation of information is presented in an obvious way?
2. What is the effect of correlation neglect on the outcome of an election where voters are influenced by ideological preferences?

The third chapter (Section 4) presents an experiment, where I tested whether leadership can influence the contribution pattern of individuals in a public goods game with growth.[3] In this chapter, I combine existing experiments concerning (i) leadership in static public goods games (see, for example, Moxnes and Van der Heijden, 2003; Güth, Levati, Sutter, and Van Der Heijden, 2007) and (ii) experiments on dynamic public goods games with growth (see, for example, Gächter, Mengel, Tsakas, and Vostroknutov, 2017). In my setup, leadership is associated with role model behavior or leading-by-example. This means that the leader has no formal power, but simply acts as first-mover. Through the dynamic setting, where growth is possible, I can investigate the long run effect of leadership. Additionally, I check whether the behavioral type of the leader has an effect on the success of the group. To determine the cooperation type of each individual, I used a sequential prisoner's dilemma where the participants had to make their decisions via strategy method.
In summary, the main questions I want to answer in this chapter are:

1. Does leading-by-example yield an improvement in a dynamic setting with endowment carryover?
2. Does the behavioral type of the leader affect the group's performance?

The fourth chapter (Section 5) thematically ties in with the third chapter.[4] While, I show in the third chapter that cooperation types have a high predictive

[3]This chapter is based on Eichenseer and Moser (2019b).
[4]The fourth chapter is based on Eichenseer and Moser (2019a).

power concerning the success of a group, I demonstrate in the fourth chapter that different methods for eliciting cooperation types are consistent to a large extent. More preciously, I compare a *one-shot public goods game with strategy method*, using the procedure of Fischbacher, Gächter, and Fehr (2001), with a *sequential prisoner's dilemma*, which was for example used in Miettinen, Kosfeld, Fehr, and Weibull (2017) and Kosfeld (2019). Furthermore, I want to investigate which method yields more valid results depending on the underlying research question.
In summary, the main questions I want to answer in this chapter are:

1. Are two structurally different methods for classifying cooperation types consistent?

2. Which of these methods yields the more valid results?

This thesis concludes with a final review of the conducted experiments, where I summarize the most important results and findings of my work (Section 6).

To answer the underlying questions, I made use of economic experiments, either in the laboratory or as online experiments. Controlled lab and online experiments have a few important advantages compared to field experiments or observational data. For example, (i) subjects can be randomly assigned to different treatment groups to rule out any selection biases and (ii) the researchers can keep specific variables constant in the lab, which is not always possible in the field. This makes it easier to make causal inferences. The lab experiments for the first and third chapter were programmed in zTree (Fischbacher, 2007) and conducted in the Regensburg Economic Science Lab (RESL). The online experiment for the second chapter was programmed in *LimeSurvey* and conducted on the research platform *Prolific*. The online experiment for the fourth chapter was programmed in *LimeSurvey* and conducted on Amazon Mechanical Turk (*MTurk*). For the experiments, I received a generous funding through the International Doctoral Program "Evidence-Based Economics" of the *Elite Network of Bavaria*.

Overall, I hope that my thesis can contribute to the vast literature on non-standard decision-making in behavioral economics by providing some new ideas and motivations which hopefully might further help to identify potential cognitive mistakes and ultimately help to avoid irrational behavioral patterns.

## 2 Hypothetical thinking and the winner's curse[5]

The *winner's curse* is a well-known empirical phenomenon in common value auctions, which was first described by Capen, Clapp, and Campbell (1971). They showed that many oil companies in the 1960's and 1970's had to report a drop in profit rates because of systematic overbidding in auctions for drilling rights. Later experimental evidence for the winner's curse was also found in a large number of lab studies (see, for example, Bazerman and Samuelson, 1983; Thaler, 1988; Charness and Levin, 2009; Ivanov, Levin, and Niederle, 2010).

I show in an experimental setup that bidders are more likely to avoid the winner's curse when they are informed, before submitting a bid, whether their bid is the winning bid or not. By giving the subjects this information, I weaken the requirement for them to condition their bid on winning the auction: winning or losing are now not hypothetical anymore, and the adverse selection issue of winning an auction with a common value for all bidders becomes more salient. The findings of my thesis suggest that mistakes in hypothetical thinking seem to explain a substantial part of the winner's curse. Thus, approaches that focus on this mental process might be more suitable for explaining this phenomenon than belief-based models like cursed equilibrium (Eyster and Rabin, 2005), which state that the winner's curse is mainly driven by inconsistent beliefs. Herewith, this thesis attempts to shed light on the ongoing debate on whether models, focusing on the erroneous belief formation of individuals, like cursed equilibrium and level-$k$ (Eyster and Rabin, 2005; Crawford and Iriberri, 2007)[6] or approaches concerning contingent reasoning on hypothetical future events (Charness and Levin, 2009; Ivanov, Levin, and Niederle, 2010; Esponda and Vespa, 2014) are more suitable for explaining the winner's curse.

Contingent reasoning refers to the ability of thinking through hypothetical scenarios and to perform state-by-state reasoning. There is evidence that people have difficulties engaging in this cognitive task. While this is well documented in the psychological literature (see, for example, Evans, 2007; Nickerson, 2015; Singmann, Klauer, and Beller, 2016), economists devoted little attention to this issue for a long time. However, in the more recent economic literature this topic appears more and more frequently (see, for example, Charness and Levin, 2009; Louis, 2013; Esponda and Vespa, 2014; Ngangoué and Weizsäcker, 2015; Levin, Peck, and Ivanov, 2016;

[5] This chapter is a slightly modified version of Moser (2019).

[6] Both models fall into the category of belief-based models, since the cause manifesting in the winner's curse is seen in the belief formation of individuals. The general assumptions of Bayesian Nash Equilibrium are still fulfilled, in the sense that subjects best-respond to beliefs, but the assumption about the consistency of beliefs is relaxed. In cursed equilibrium, the degree of *cursedness* is given by $\chi \in [0, 1]$, i.e., the belief that with some probability $\chi$ the actions of the opponents do not depend on their types. A value of 0 is equivalent to the usual Bayesian Nash Equilibrium, whereas a value of 1 corresponds to a setting in which the players do not assume any correlation between the actions of a player and his type, which is also denoted as *fully* cursed equilibrium.

Esponda and Vespa, 2016; Li, 2017; Koch and Penczynski, 2018).

In contrast to belief-based models, like cursed equilibrium, the concept of contingent reasoning has still received very little formal treatment. Li (2017) represents the first attempt to capture this mental process formally by introducing the concept of *obviously strategy-proof* (OSP) mechanisms.[7] A major contribution of Li's paper is to explain why subjects perform better in ascending bid auctions compared to sealed bid auctions. However, it cannot explain why common value auctions might be more challenging for the bidders than private value auctions. For this reason, the concept of OSP mechanisms is not sufficient to fully capture the most relevant aspects of the winner's curse.

Can a broader definition of contingent reasoning explain why bidders fall prey to the winner's curse? To answer this question I will go back one step from OSP mechanisms and focus only on the events of winning or losing an auction - which provide information about the true value of a good in common but not in private value auctions. Bidders in common value auctions now face two cognitive hurdles. First, they have to be able to recognize and identify several hypothetical scenarios which might possibly occur (e.g., winning or losing an auction) and second, the bidders have to be able to infer information from such hypothetical events. The question is which of these cognitive tasks is more challenging for subjects and to what extent they affect the likelihood of the winner's curse to occur. There are three possibilities: (i) bidders are perfectly able to perform state-by-state reasoning, but they neglect the informational content of the other bidders' actions and, hence, the informational content of winning an auction as proposed by cursed equilibrium; (ii) bidders take into account that the bids of the other players are correlated with their signals, but they are not able to identify the relevant states to condition on in the first place; (iii) bidders are neither able to perform state-by-state reasoning, nor do they take into account the informational content of winning an auction.

In this chapter, I test experimentally whether subjects in an auction are able to infer information from the events of winning or losing if these are not hypothetical anymore. For this purpose, I constructed a second-price auction in which the bidders learn whether a bid, which was considered optimal ex ante, is the winning bid or not (but without learning their payoff yet) and they receive the possibility to change this bid. This treatment intervention is similar to the *sequential* treatment in Esponda and Vespa (2014) where participants in an election learned whether their votes were

[7] According to Li (2017), a mechanism is OSP if and only if an optimal strategy can be found without the necessity of performing contingent reasoning. However, Li's definition of contingent reasoning is very strict in a game-theoretic sense as it does not only involve conditioning on different broader states of the world or future events (such as conditioning on being the winner in an auction or conditioning on being the pivotal voter in an election), but also conditioning on each possible decision node in a given information set. Additionally, there is only a differentiation between whether a game is OSP or not, but no distinction between "more" or "less" contingent reasoning.

pivotal or not before they had to cast a vote. In contrast to Koch and Penczynski (2018) my approach is weaker, since it does not remove the need for contingent reasoning at all, but only partly in a specific context. However, thereby I am able to show whether such a weak manipulation, on a feedback basis, is sufficient for subjects to improve their bids, without changing the whole structure of the game. Additionally, it has to be noted that even Koch and Penczynski (2018) did not meet the strict requirements of completely eliminating the necessity for *contingent reasoning*, according to the definition of Li (2017). This further emphasizes the urgent need for a consistent definition of this term.

The auction model I used in the experiment is based on a second-price sealed bid auction similar to the *wallet game* proposed by Klemperer (1998) and the model used in Avery and Kagel (1997), which is a non-standard common value auction. The basic idea of the game is the following: two players, indexed by $i = 1, 2$, receive a private *iid* signal, $x_i$, drawn from some commonly known distribution. In a second-price sealed bid auction they bid for an object worth $v = x_1 + x_2$. This game is played in two stages. In stage I, the subjects participate in the wallet game against a random opponent. In stage II, the subjects play the same auctions again, against the decisions of the former opponent, but this time the subjects in the treatment group are informed whether their initial bid was the winning bid or not. In this sense, the bidders receive information about some realized event before they have to come up with a bid.[8] Apart from knowing whether their bid is the winning or losing bid, the subjects face exactly the same decision problem as in stage I.

My design allows a clear distinction between mistakes in contingent reasoning and cursed equilibrium since two crucial assumptions of the latter are that (i) no (or only a partial) correlation between the other players' actions and types is assumed and (ii) bidders still best-respond given their beliefs. Hence, for a "cursed" bidder the information on whether his bid is higher or lower than the bid of the opponent, does not provide him further information which would be relevant for updating his bid. For fully cursed bidders this argument is straightforward, but it also holds for partly cursed bidders because of the best-response assumption. By definition, the bid of a cursed bidder is already evaluated conditional on winning in stage I, albeit a partly cursed bidder implicitly assumes that winning is less informative than in equilibrium. This means, the feedback of winning in stage II provides no further relevant information, since it is already included in the decision of a partly cursed bidder in stage I.

The reason for choosing the wallet game, instead of a more standard model for common value auctions, is that in this game a naïve bidding strategy can lead to both

[8] See also Esponda and Vespa (2016) for a distinction between static and dynamic choice situations.

over- and underbidding, relative to the symmetric equilibrium strategy, depending on whether the private signal is low or high. In this sense, there can be both a winner's and a loser's curse (see also Holt and Sherman, 1994).[9] This property is useful for two reasons. First, I am able to control for psychological explanations, stating that the winner's curse is mainly driven by emotional factors of winning (see, for example, Van den Bos, Li, Lau, Maskin, Cohen, Montague, and McClure, 2008; Astor, Adam, Jähnig, and Seifert, 2013). Second, and more importantly, I am able to check whether bid shading in stage II is due to proper Bayesian updating or just a rule of thumb when learning that a certain bid was the winning bid. In my setup, bid shading is only advisable for low signals, but not for high signals. Thus, the subjects have to differentiate between these two kinds of signals, instead of following the simple decision rule "*decrease your bid, when you learn that your bid was the winning bid*".

The findings of my thesis reveal two important observations:

(i) Bidders are more likely to avoid the winner's curse and the loser's curse in stage II when they are informed whether their bid is the winning bid or not, given that the respective information has a sufficiently high predictive power concerning the opponent's signal. This suggests that the crucial cognitive hurdle for bidders in a common value auction is not forming beliefs about the opponents' behavior, but identifying the relevant states to condition on in the first place.

(ii) Information can also be negative for the bidders, depending on the context. The subjects in my experiment differentiated only imperfectly between situations in which decreasing (increasing) a bid is rational and those in which it is not and they often used simple heuristics instead of making strategic changes. This behavior might be partly explained by an actual *joy of winning* or, to be more precise, a disappointment of losing. When subjects learned that they lost an auction in stage I, most of them increased their bid in stage II, regardless of having a low or high private signal. Conversely, when subjects learned that they won an auction in stage I, they acted more strategically and bids for low signals were decreased at a much higher rate compared to bids for high signals (78.7% vs. 42.3%).

This chapter is organized as follows. Section 2.1 will provide an overview about the current literature closely related to my research topic. Section 2.2 will present the underlying theoretical model for the experiment. Section 2.3 describes the experimental design. Section 2.4 presents and discusses the results of the experiment. Section 2.5 concludes the chapter.

[9]A loser's curse can be understood as losing an auction, but the bidder could have won with a positive payoff.

## 2.1 Literature review

This thesis is the first to investigate the direct effect of the information of winning (or losing) in the context of a common value auction. The novelty of my design is that I use a dynamic choice setting à la Esponda and Vespa (2014) in a sealed bid auction. Similar to them, I want to differentiate between mistakes in hypothetical thinking and problems with extracting information from the opponents' actions. With this setup I am able to test whether potential errors in common value auctions occur due to an inability of performing contingent reasoning or due to inconsistent beliefs as proposed by cursed equilibrium. To the best of my knowledge, none of the recent approaches provided a similar framework. For example, the experiment by Charness and Levin (2009) is a single decision maker problem in an adverse selection environment. Additionally, they measured the effect of contingent reasoning only indirectly by transforming their initial game into a set of simple lotteries. Ivanov, Levin, and Niederle (2010) showed that belief-based models, like cursed equilibrium, might be not that powerful in explaining the winner's curse, but they provided no alternative explanation. The paper of Koch and Penczynski (2018) is also closely related to my work. So far they are the only ones who combined contingent reasoning and belief-based models in their experiment. However, they focused on a different aspect of contingent reasoning, not directly related to the event of winning an auction. Finally, Levin, Peck, and Ivanov (2016) used Dutch auctions and showed that in this format conditioning on winning is more salient compared to a strategically equivalent first-price auction. All of these papers will be discussed in more detail in the following part.

As explained above, the paper most closely related to my own work is the one by Esponda and Vespa (2014). They created a common value voting experiment where a subject interacted with two computers. The main task for the subject was to submit a vote for a ball which was either red or blue. Due to the commonly known voting algorithm of the computers, the vote of the subject was only relevant when the ball was blue and, hence, the optimal choice for the subject was always to vote for blue. Esponda and Vespa (2014) observed that subjects made significantly less errors in a sequential election, where the voters knew whether they were pivotal or not, compared to a simultaneous election, where the voters were not informed about their pivotality. Similarly, in their lab experiment Ngangoué and Weizsäcker (2015) found that traders in a financial market setting performed better in a sequential trading mechanism where no Bayesian updating on hypothetical events was required.

Charness and Levin (2009) conducted an experiment constructed as a simple individual choice problem similar to an *acquiring a company game*, which is based on a lemon market (Akerlof, 1970), to investigate the driving mechanisms behind the winner's curse. The results of their paper revealed that most of the subjects systematically overbid even though the strategy of the computerized seller was known and, hence, there was no need for the subjects to form beliefs about his behavior. This pattern was reduced in a setting where the bidding task was transformed into a set of simple lotteries with no requirement of thinking in hypothetical situations. However, transforming the initial game into a simple lottery task changes the whole structure of the game and so it remains difficult to extract a causal effect of hypothetical thinking.

Ivanov, Levin, and Niederle (2010) used a similar approach as Charness and Levin (2009), but they conducted their experiment in an actual auction context using the *maximal game*.[10] The aim of their experiment was not to look at the effect of contingent reasoning, but rather to disprove that the winner's curse is driven by inconsistent beliefs. Ivanov, Levin, and Niederle (2010) observed significant overbidding which was not reduced in a modified setting of the maximal game where the beliefs of the participants were explicitly formed and, thus, belief-based models, like cursed equilibrium, had little explanatory power. However, Costa-Gomes and Shimoji (2015) criticized Ivanov, Levin, and Niederle (2010) for the misuse of some of the game theoretical concepts, arguing that some of their findings can indeed be explained by belief-based models. Similarly, Camerer, Nunnari, and Palfrey (2016) argued that the observed behavior of the subjects in Ivanov, Levin, and Niederle (2010) can be explained by belief-based models if they are combined with a *quantal response model*. Under this extension the assumption of perfect best-reply behavior in cursed equilibrium and level-$k$ model is relaxed and stochastic choices are allowed. They showed that this extended model fits very well to the data of Ivanov, Levin, and Niederle (2010).[11]

My approach is different from the one in Ivanov, Levin, and Niederle (2010). Rather than showing that bidders do not best-respond, even though their beliefs are explicitly formed, I show that bidders in my setup react to information which is, by definition, not relevant for the updating process of a "cursed" bidder.[12]

Koch and Penczynski (2018) conducted an auction game similar to the one in

[10] Ivanov, Levin, and Niederle (2010) criticized that the acquiring a company game, used in Charness and Levin (2009), represents a lemon market and not a common value auction. Thus, it seems problematic to extend the findings from their experiment to common value auctions in general. They also claimed that it can make a difference whether a subject plays against other people or against a computer. In fact, Ivanov, Levin, and Niederle (2010) also used a computer treatment, but in their case the computer mimicked the subject's own past strategy.

[11] However, Camerer, Nunnari, and Palfrey (2016) also assume that, if the *perfect* best-reply assumption of cursed equilibrium and level-$k$ model is maintained, these models are very bad in predicting the behavior in maximal value games.

[12] In the setup of Ivanov, Levin, and Niederle (2010) one could still argue that the belief formation of a cursed bidder takes place in an isolated "black box" and is not affected by information about the opponent's behavior which is given to such a bidder. This is not an issue in my design.

Kagel and Levin (1986) to investigate how explanations concerning contingent reasoning and belief-based models interact. The authors found that the relaxation of both cognitive requirements had a significant effect on avoiding the winner's curse. To identify the effect of contingent reasoning on the bidding behavior of the subjects, the authors used a transformed version of the original game, where both bidders received the *same* information about an object with a stochastic value, instead of *different* private signals. In this sense, there was no need for the bidders to condition on whether their own signal was relatively low or high. This is different from my approach, where I focus on the direct effect of the information of winning an auction. Additionally, I only changed one parameter in stage II, whereas Koch and Penczynski (2018) used two different games, but with the same best-response functions and equilibria.

Levin, Peck, and Ivanov (2016) used an experiment to investigate the impact of Bayesian updating and non-probabilistic reasoning (referred to as *contingent reasoning* in this thesis) on avoiding the winner's curse. They used common value Dutch and common value first-price auctions based on the model in Kagel, Harstad, and Levin (1987) and compared both versions to quantify the effect of non-probabilistic reasoning. Additionally, they measured the skills of the participants in Bayesian updating and non-probabilistic reasoning through a questionnaire, which the subjects had to answer before they participated in the auctions, and showed that both cognitive skills had a significant effect on avoiding the winner's curse, resulting in higher earnings for the subjects. The authors also showed that subjects performed better in a Dutch auction, where the auction ended when the first subject stopped the clock, compared to a Dutch auction, where the auction ended when the last subject stopped the clock and no subject received any feedback about whether he is the highest bidder. In the first setting, a bidder knows, in the moment of stopping the clock, that he is the highest bidder and, hence, it is more salient for him to condition his bid on winning compared to the second setting with a "silent" clock. The authors concluded that bidders can better handle the winner's curse if the requirements for this form of hypothetical thinking are reduced.[13] Note that this form of a Dutch auction is not comparable to the sequential approach of Esponda and Vespa (2014) because, strictly speaking, only *after* stopping the clock the subject learns that he is the highest bidder.[14] The advantage of my design is that the bidders in stage II receive this information *before* they have to come up with a bid.

[13]As a further robustness check, the authors also used private-value auctions, where they did not find a significant difference between these two kinds of Dutch auctions.

[14]A further problem of Dutch auctions is that the bidders have to make a decision pressed for time. This can lead to various effects on emotional level (see, for example, Adam, Krämer, and Weinhardt, 2012; Adam, Krämer, and Müller, 2015).

## 2.2 The model

In the following section I present a formal description of the auction game used in the experiment. This game is based on the *wallet game* proposed by Klemperer (1998) and the model used in Avery and Kagel (1997). Henceforth, this game will be denoted as the *wallet game*, although it is slightly different from the original one.[15]

There are two players, indexed by $i = 1, 2$. Each player $i$ receives a signal $x_i$ from the set $X = \{0, 1, \ldots, 9, 10, 50, 51, \ldots, 59, 60\}$ ($|X| = 22$), with each value equally likely and with replacement (so there are 11 low and 11 high signals). The players compete for an object worth $v = x_1 + x_2$ in a second-price sealed bid auction. The players are allowed to choose a bid $b_i$ in the range of $[0, 120]$, with only integer values possible. In case of a tie, the player with the higher signal wins. If the signals are also equal, both players receive a payoff of 0. The payoff of player 1 (analogously for player 2) is thus given by:

$$\pi_1 = \begin{cases} x_1 + x_2 - b_2 & \text{if } b_1 > b_2 \\ x_1 + x_2 - b_2 & \text{if } b_1 = b_2 \wedge x_1 > x_2 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The utility function of both players is assumed to be symmetric across $i$ such that $u_i(\mathbf{x}) = v = x_1 + x_2$ with $\mathbf{x} = [x_1, x_2]^T$ (see also Crawford and Iriberri, 2007).

**Proposition 1.** *Given that player $i$ sees a signal $x_i$, the expected value of the object is given by $E[V|X_i = x_i] = x_i + 30$.*

The proof of Proposition 1 is straightforward and, hence, omitted. Henceforth, bidding according to $b_i(x_i) = x_i + 30$ will be denoted as *naïve* bidding, since the bid is not evaluated conditional on winning. It is easy to show that this bidding function cannot be part of a symmetric equilibrium. For example, a naïve bidder with a low signal ($x_i \in \{0, \ldots, 10\}$) only wins when the other naïve player also has a low signal and the price to pay is always above (or equal to) 30, which results in a negative payoff for the winning bidder.

**Proposition 2.** *Bidding $b_1(x_1) = \alpha \cdot x_1$ and $b_2(x_2) = \frac{\alpha}{\alpha-1} \cdot x_2$ are equilibrium strategies for any $\alpha > 1$. For $\alpha = 2$ we have the unique symmetric equilibrium with $b_i(x_i) = 2 \cdot x_i$ for all $i \in \{1, 2\}$.*

The proof of Proposition 2 can be found in Appendix B.[16] The intuitive expla-

[15] A detailed analysis of the general wallet game with $N$ bidders can be found in Eyster and Rabin (2005) and Crawford and Iriberri (2007).

[16] The proof of uniqueness of the symmetric equilibrium is not presented in this thesis, but it can be found in Eyster and Rabin (2005) and Crawford and Iriberri (2007) for the general wallet game.

nation for the symmetric equilibrium is the following: under the assumption that all bidders have the same bidding function $b_i(x_i)$, which is monotonically increasing in $x_i$, a rational player anticipates that he only wins if his signal is at least as high as the signal of the opponent, thus, barely if $x_1 = x_2$ holds. Hence, the optimal bid in (the symmetric) equilibrium is given by $b_i(x_i) = 2 \cdot x_i$. Henceforth, bidding according to $b_i(x_i) = 2 \cdot x_i$ will be denoted as *sophisticated* bidding. Note that these equilibrium strategies are neither affected by risk preferences nor the distribution of the signals $x_1$ and $x_2$ (see also Klemperer, 1998).

**Proposition 3.** *Any bid* $b_i(x_i)$ *outside the interval* $[x_i, x_i + 60]$ *is weakly dominated.*

The proof of Proposition 3 can be found in Appendix B. The intuitive explanation is the following: bidding below the private signal can never be optimal in a second-price auction because the value of the good is at least $x_i$. Bidding above $x_i + 60$ is, likewise, never optimal, since the maximal value of the good is at most $x_i + 60$.

**Proposition 4.** *All equilibria from Proposition 2, except the symmetric one with* $\alpha = 2$*, involve weakly dominated bids for one player for at least some signals.*

The proof of Proposition 4 can be found in Appendix B. Because of this issue, asymmetric equilibria are less plausible, since they involve weakly dominated bids for at least some $x_i$ (for one player). Additionally, some bids resulting from a bidding function $b_i(x_i) = \alpha \cdot x_i$ with $\alpha > 2$ are not even feasible because the bidding range is restricted on the interval $[0, 120]$. Consequently, throughout the analysis of this thesis, I will only focus on the symmetric equilibrium, with $b_i(x_i) = 2 \cdot x_i$, as a benchmark for sophisticated bidding. Hence, we have the following benchmark bidding functions for naïve (NVE) and sophisticated (BNE) bidding:

$$b_i^{NVE}(x_i) = x_i + 30 \tag{2}$$

and

$$b_i^{BNE}(x_i) = 2 \cdot x_i. \tag{3}$$

At this point it seems important to explain the reasons for choosing a modified model of the wallet game, with discrete signal space and a gap in the middle, instead of a continuous range between 0 and 60. First, in the original wallet game, the naïve and sophisticated bidding functions get closer, the more one reaches the middle of the signal space. For the value in the middle, which would be 30 in a continuous range, both bidding functions are even identical. Thus, in the original wallet game, it gets more difficult to distinguish between naïve and sophisticated bidding for these intermediate values. Second, in my setup the sophisticated bidding strategy is also a best-response for the naïve bidding strategy. So even if a sophisticated player assumes that his opponent bids naïvely, he still best-responds by using the sophisticated bidding function. The intuition is simple: if the other player uses the naïve bidding strategy, the best-response is to always win for high signals and to always lose for low signals. This is given when following the sophisticated bidding rule. This makes $b_i^{BNE}(x_i) = 2 \cdot x_i$ a stronger and more credible benchmark for sophisticated bidding, although it is still not a weakly dominant strategy.

As a concluding remark it is important to note that bidding $b_i(x_i) = x_i + 30$ can be explained by both: mistakes in contingent reasoning and cursed equilibrium or level-$k$ model (Eyster and Rabin, 2005; Crawford and Iriberri, 2007).[17] This elucidates a general dilemma in behavioral and experimental economics: even though a model fits well to the data, it is not clear whether it has indeed explanatory power. Hence, it remains important to disentangle competing theories.

## 2.3 Experimental design

### 2.3.1 Implementation

The experiment was conducted in the Regensburg Economics Science Lab (RESL) in February 2017. For the technical implementation the software zTree was used (Fischbacher, 2007) and for the recruitment of participants the online recruitment system ORSEE was used (Greiner, 2004). In total, 5 sessions were conducted with overall 72 participants (mostly undergraduate students from various fields). For each session I had between 10 and 18 participants (always an even number).

The experimental currency unit (ECU) used in the experiment were *Taler*. All signals and bids in the experiment were expressed in terms of Taler. The exchange rate was 1 Euro = 10 Taler. The participants were payed out in Euro at the end of the experiment. The average payment was 16.02 Euro. The sessions lasted between 60 and 75 minutes.

At the beginning of the experiment each participant was endowed with 50 Taler. At the end of the experiment, the participants received their initial endowment plus (minus) their generated earnings (losses) in both stages of the experiment (in total 6 rounds were payoff-relevant). Additionally, a show-up fee of 4 Euro was payed to each subject which was guaranteed no matter what decisions the subject made during the experiment. So each subject earned at least 4 Euro. If the losses exceeded 50 Taler, the participants only received their show-up fee. 4 out of 72 participants suffered from higher losses.

[17] A bidder who does not condition his bid on winning, ignores the adverse selection issue inherent in this kind of auction and only considers his private signal and a "cursed" bidder implicitly assumes that the opponent will bid independently of his signal which makes $b_i(x_i) = x_i + 30$ a best-response.

### 2.3.2 Basic setup

The experiment is divided into *two stages*, both of which are payoff-relevant. The subjects are informed that there is a second stage, but they receive the details only after finishing stage I. In stage I, the players participate in 15 rounds of the wallet game (i.e., each subject receives successively 15 random signals drawn from the set $X = \{0, 1, \ldots, 9, 10, 50, 51, \ldots, 59, 60\}$ *with replacement*).[18] Each subject gets randomly matched with another subject of the group (e.g., subject $k$ and subject $l$). In this sense, the first signal of subject $k$ is matched with the first signal of subject $l$, the second signal of $k$ is matched with the second signal of $l$, and so on. The subjects are aware that they play against a fixed stranger for the course of the 15 rounds.[19] The subjects receive no immediate feedback after submitting their bids, but only learn their payoff at the very end of the experiment (i.e., after finishing stage II). So there should be neither endowment effects nor learning effects through explicit feedback. Three randomly selected rounds are payoff-relevant. A typical decision screen of stage I is shown by Figure A.1 in Appendix A.

Before starting with the actual task, all participants are asked to answer eight control questions (see Appendix D) and to participate in five testing rounds of the wallet game without monetary payoff, but with immediate feedback about their hypothetical payoff after each bid, to give them a practical understanding of the game. However, the subjects receive no feedback about the bid or the signal of the opponent. The opponent in the testing rounds is represented by a computer who uses the bidding strategy $b_j(x_j) = x_j + 30$. The participants are not explicitly informed about the strategy of the computer and they only learn that the bidding function of the computer is monotonically increasing in his signal.[20]

Stage II is, from a theoretical point of view, a repetition of stage I. All subjects receive the same 15 signals as in stage I, in the same order. In stage II, the subjects play against a computerized opponent who mimics the behavior of their former opponent from stage I. Subject $k$ plays against the decisions of subject $l$ in stage I and vice versa. Each subject therefore faces exactly the same decision

[18]The actual explanation in the experiment is that each player receives an envelope with a random amount of money inside (see *Instructions* in Appendix C).

[19]Since the subjects received no feedback after submitting their bids, I did not use a random matching approach in which the opponent would change after each round.

[20]The reason for using the naïve bidding function as a strategy for the computer in the testing rounds is that deviating from the sophisticated bidding function $b_i^{BNE}(x_i) = 2 \cdot x_i$ is much more harmful when the other player uses the naïve bidding function. If the computer had used the sophisticated bidding function, the subjects would have been less likely to realize that deviating from this strategy is a bad idea. In the presented setup $b_i^{BNE}(x_i) = 2 \cdot x_i$ is a best-response for both the naïve and the sophisticated bidding function.

problems as in stage I if we abstract from social preferences.[21] As in stage I, the bids and the signals of the opponent are not observable. The rules for bidding and winning are the same as in stage I and the same three randomly selected rounds are again payoff-relevant. In stage II, the subjects are randomly (and with equal likelihood) assigned to either treatment INF (*information*) or NOINF (*no information*).

### 2.3.3 Treatments

**INF** (*Information*)

The subjects who receive treatment INF are able to see for each signal whether their initial bid from stage I was HIGHER or LOWER than the respective bid of the opponent.[22] For example if a subjects sees that her initial bid of $b_i^I = \bar{z}$ was HIGHER than the bid of his opponent, she knows that submitting a bid of $b_i^{II} = z \geq \bar{z}$ results in winning the auction for sure. Conversely, if a subjects sees that his bid $b_i^I = \underline{z}$ was LOWER than the bid of his opponent, he knows that submitting a bid of $b_i^{II} = z < \underline{z}$ results in losing the auction for sure. In this sense, there is no requirement anymore to condition on the hypothetical event of winning (or losing) for a certain range of bids, especially for the bid which was considered as optimal in stage I. An example of a typical screen is given by Figure A.2 in Appendix A.

**NOINF** (*No information*)

The subjects in treatment NOINF face exactly the same situation as those in treatment INF, except that they do not get any information about the bid of their opponent. Instead of HIGHER or LOWER they only see *???* on their screen. However, all subjects are informed about both treatments (i.e., the subjects in NOINF know how treatment INF looks like and vice versa). An example of a typical screen is given by Figure A.3 in Appendix A. The general structure of the treatments is illustrated by Figure 2.1.

[21] In contrast to stage I, the decisions do not affect the payoff of the opponent anymore. So if a subject has preferences concerning the other player's payoff, the decision problem might be different for him.

[22] If the bids are equal, the subjects also get the message "LOWER", so LOWER means lower or equal.

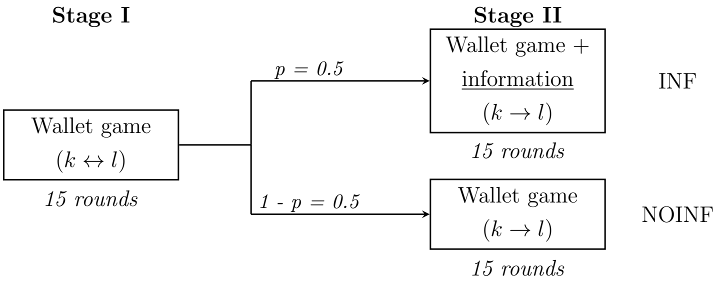



Figure 2.1: Illustration of the treatments

### 2.3.4 Behavioral predictions

The following behavioral predictions are based on the assumption of a boundedly rational agent who is cognitively limited in the sense that he is not able to infer information from hypothetical events, like winning or losing an auction. Such an agent will be denoted as *naïve*.[23] A naïve bidder will behave as if the auction is a private value auction since he ignores the adverse selection in the event of winning (for low signals) and the positive selection in the event of losing (for high signals) and will choose a bid that is based on the expected value of the good if he is risk-neutral. Hence, he will bid $b_i(x_i) = x_i + 30$ for any given signal.

If all bidders behave like this, there will always be a winner's curse for the bidders who win with a low signal ($x_i \in \{0, \ldots, 10\}$), since they only win when the other player also has a low signal and the price to pay is always above (or equal to) 30. Conversely, all bidders who lose with a high signal ($x_i \in \{50, \ldots, 60\}$) fall prey to a loser's curse, since they could have won the auction at a profitable price.

How will such a cognitively limited agent react, when he receives the information HIGHER or LOWER - i.e., when he learns whether his ex ante, as optimally considered bid, is the winning bid or not? If such an agent is able to correctly extract information form *observed events*, which implies that he has consistent beliefs about the behavior of the opponent, he should update his bids at least for some constellations of signal and information. Consider therefore Table 2.1 and two players, 1 and 2. Player 1 is a naïve bidder who bids according to $b_1(x_1) = x_1 + 30$ in stage I. Player 2 is the opponent of player 1. As a minimum requirement for rationality, player 2 does not use any weakly dominated bids.[24] The first columns show the four

[23]Note that this thesis will not provide a generalizable formal model of naïve behavior as it can be found, for example, in Li (2017).

[24]This means for each $x_2$ he may choose any bid in a range of $[x_2, x_2 + 60]$. So, for low signals the bid can be in a range of $[0, 70]$ and for high signals in a range of $[50, 120]$.

different constellations of signal and information that can occur in the experiment from the perspective of player 1. The constellations are abbreviated by the form $s^T$, with $s \in \{ls, hs\}$ (low or high signal) and $T \in \{L, H\}$ (information LOWER or HIGHER). The last column shows the prediction of player 2's signal when player 1 receives the respective information and given that he believes that his opponent does not use any weakly dominated bids.

**Proposition 5.** *If a player bids according to $b_i(x_i) = x_i + 30$ in stage I and the other player does not use any weakly dominated bids, he can predict in stage II whether the signal of the other player is low or high when (i) he wins with a low signal or (ii) he loses with a high signal.*

| *Constellation* | **Signal** | **Information** | **Prediction of other's signal** |
|---|---|---|---|
| $ls^H$ | Low signal | HIGHER | *Low signal* |
| $ls^L$ | Low signal | LOWER | *Low or high signal* |
| $hs^H$ | High signal | HIGHER | *Low or high signal* |
| $hs^L$ | High signal | LOWER | *High signal* |

Table 2.1: Prediction of other player's signal for different constellations

The proof of Proposition 5 is provided in Appendix B. We can see that for constellations $ls^H$ and $hs^L$ the respective information provides unambiguous hints about the opponent's signal for player 1. While for these constellations the weak assumption of an opponent who does not use any weakly dominated bids is sufficient to correctly predict his signal (low or high)[25], this is not the case for constellations $ls^L$ and $hs^H$ where the probabilities for a low or high signal depend on the specific bidding function of the opponent. So for the latter two constellations a sophisticated computation of probabilities is required to predict the expected value of the opponent's signal correctly. Based on this, clear predictions can be made for constellations $ls^H$ and $hs^L$, but not for constellations $ls^L$ and $hs^H$.

- In constellations $ls^H$ and $hs^L$, bidders will update their bids in stage II, resulting in bids closer to the Nash prediction and higher earnings for the subjects.
- For constellations $ls^L$ and $hs^H$, the behavior of the bidders can be ambiguous since the expected value of the opponent's signal depends on the specific beliefs of the bidders and the updating process requires non-trivial computations.

[25] To be more precise: for constellations $ls^H$ and $hs^L$ it is sufficient to assume that player 2 does not bid more than $x_2 + 60$ for low signals (no weakly dominated overbidding) and not less than $x_2$ for high signals (no weakly dominated underbidding). Since 91.94% of all bids in stage I fall into this category, this is actually a plausible belief.

As a concluding remark, it is important to note that cursed equilibrium cannot explain an updating of bids in stage II, since a "cursed" bidder implicitly assumes that the opponent chooses a bid which is independent of his signal, hence, the respective information would not be useful for such a bidder.[26]

## 2.4 Results

In this section, the results of the experiment will be presented. First, I will give a descriptive overview about the overall bidding pattern in stage I, before the subjects are assigned to a treatment group. Second, I will present the effects of the *information* treatment (INF) in terms of bidding behavior and the resulting profits.[27] Finally, I will provide evidence that the observed behavior is to a not negligible extent driven by an actual updating of the opponent's signal, when receiving information which is not hypothetical anymore, and not because the subjects followed a simple decision rule like "always decrease when you see HIGHER and always increase when you see LOWER". Additionally, I will also provide an analysis on the subject level, where I classify subjects into different categories based on their bidding behavior.

All bids and profits in the following part are expressed in terms of *experimental currency units* (ECU) with an exchange rate of 1 Euro = 10 ECU.

### 2.4.1 Overall bidding pattern in stage I

Figure 2.2 presents the mean and median bids for low and high signals in stage I. The average bids for low signals are *above* the Nash prediction and the average bids for high signals are *below* the Nash prediction in stage I. For high signals the mean and median bids are fitted very well by the naïve bidding function $b_i^{NVE}(x_i) = x_i + 30$. The results of a Wilcoxon sign rank test show that for *high signals* the hypothesis that the actual bids are equal to bids resulting from the naïve bidding function cannot be rejected ($p = 0.624$).[28]

[26]Eyster and Rabin (2005) argued that cursed equilibrium is basically not defined for sequential games and that players in sequential games might be less "cursed" than in simultaneous games. However, in my setup the players only observe whether their bid was higher or lower than the bid of the opponent, but they do not observe the specific action of the other player, as, for example, in an ascending bid auction. Additionally, it can be argued that the concept of "cursedness" would be a rather weak one if "cursedness" would suddenly vanish in the moment of revealing the other player's action (see also Deversi, Ispano, and Schwardmann (2018)).

[27]When I compare the effects between INF and NOINF, for different constellations of signal and information, I look at those subjects in treatment NOINF who *would have* received the respective information if they had been in treatment INF, in order to get an appropriate control group.

[28]Graphs reporting all bids in stage I and II can be found in Appendix A (*Distribution of bids*).

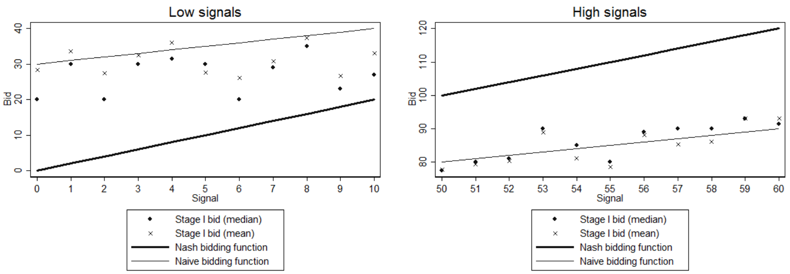



Figure 2.2: Median and mean bids in stage I

Overall, we can observe a high occurrence of the winner's curse for low signals (16.67% in stage I) and of the loser's curse for high signals (17.02% in stage I). *Conditional on winning*, the rate for the winner's curse increases to 59.31% for low signals and *conditional on losing*, the rate for the loser's curse increases to 56.47% for high signals. The rate of the winner's curse is very high, especially when considering that the auctions were conducted as *second-price* auctions. This shows clearly that the problem of irrational bidding behavior is a considerable one. A further conclusion is that emotional factors of winning seem to play a minor role in stage I, since the rates for winner's and loser's curse are very similar.

| | **Low** (LOST) | **Low** (WON) | **Low** | **High** (LOST) | **High** (WON) | **High** |
|---|---|---|---|---|---|---|
| No curse | 338<br>91.11% | 59<br>40.69% | 397<br>76.94% | 74<br>43.53% | 378<br>95.94% | 452<br>80.14% |
| Winner's curse | 0<br>0.00% | 86<br>59.31% | 86<br>16.67% | 0<br>0.00% | 16<br>4.06% | 16<br>2.84% |
| Loser's curse | 33<br>8.89% | 0<br>0.00% | 33<br>6.40% | 96<br>56.47% | 0<br>0.00% | 96<br>17.02% |
| **Total** | 371<br>100.00% | 145<br>100.00% | 516<br>100.00% | 170<br>100.00% | 394<br>100.00% | 564<br>100.00% |

Table 2.2: Winner's curse and loser's curse in Stage I for different constellations of signals. Winner's curse: won, but with a negative payoff. Loser's curse: lost, but could have won the auction with a positive payoff.

### 2.4.2 Bidding behavior in stage II

In this subsection, I present an analysis of the difference in bidding behavior between stage I and II and across treatments. First, I will illustrate my findings by providing cumulative distribution functions (CDFs) of the bids in stage I and II for different constellations of signal and information for both treatments INF and NOINF. Second, I will support my claims by means of a regression analysis.

As explained in Section 2.2, the symmetric Nash bidding function $b_i^{BNE}(x_i) = 2 \cdot x_i$ serves as a benchmark for sophisticated bidding. However, since it might be too much to expect that subjects follow exactly the symmetric Nash bidding rule, I will define a broader corridor for sophisticated bids which I call *Nash range*. Hereby, I lower the requirements for sophisticated bidding, by only checking whether the subjects are able to reason on the level of *signal categories* (low and high signals). Bids are in the Nash range for low signals as long as $b(x_i) \in [x_i, x_i + 10]$ holds. In this sense all bids will be classified as *sophisticated*, where a subject with a low signal realizes that she should bid as if the other player has a low signal as well. For high signals, bids are in the Nash range as long as $b(x_i) \in [x_i + 50, x_i + 60]$ holds. Equivalently, all bids for which the subject realizes that with a high signal, he should bid as if the other player also has a high signal, are classified as sophisticated bids. The lower and upper bounds of bids in the Nash range are, hence, 0 and 20 for low signals, and 100 and 120 for high signals, respectively. These are also the lower and upper bounds of the symmetric Nash bidding function $b_i^{BNE}(x_i) = 2 \cdot x_i$, which is embedded in the corridor of the Nash range. This pattern is also illustrated by Figure A.7 in Appendix A. The bids which are not in the Nash range are defined as *unsophisticated* (see also Figure 2.3).[29]

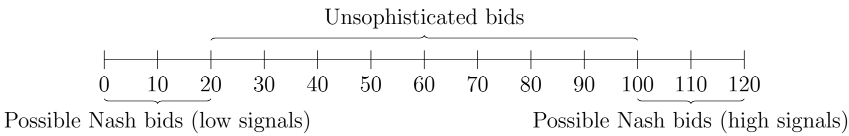



Figure 2.3: Range of unsophisticated and sophisticated bids

In the following graphs, presenting the distribution of bids in stage I and II, the vertical line marks the upper bound of Nash bids for low signals ($b_i = 20$) and the lower bound of Nash bids for high signals ($b_i = 100$), respectively.

Most noticeable are the changes in bids in constellations $ls^H$ and $hs^L$ for the subjects in treatment INF (Figures 2.4 (left side) and 2.5 (right side)). We can see that in those constellations, the mass of bids in stage II is shifted towards the Nash prediction and that a substantially larger proportion of the bids in stage II is now

[29] Note here that not all bids between 0 and 20 or between 100 and 120, respectively, are always in the Nash range. E.g., if a player has a signal of 5, only bids between 5 and 15 are in the Nash range. All other bids would be classified as *unsophisticated* in this context.

actually within the Nash range, compared to stage I. For the other constellations in treatment INF ($ls^L$ and $hs^H$), the effect is reverse, albeit the changes are less distinct (see Figures 2.4 (right side) and 2.5 (left side)). Here the mass of bids is shifted to the right for low signals and to the left for high signals - hence, in both cases further away from the Nash prediction. For constellation $ls^L$ we see that the bids in the Nash range substantially decreased in stage II. The bids in the NOINF treatment were also changed, although the subjects received no information about their bids, but here the pattern is less distinct and uniform.

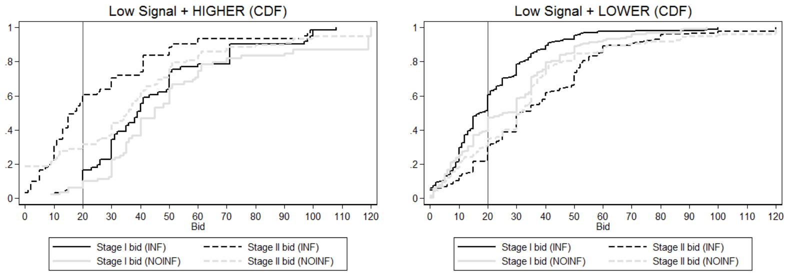



Figure 2.4: Cumulative distribution functions - Low signals

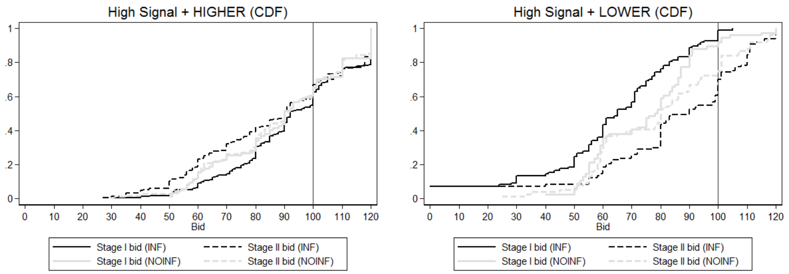



Figure 2.5: Cumulative distribution functions - High signals

In the left part of Figure 2.4 we can see that there is an accumulation of bids below 20 in stage II for low signals, when receiving the information HIGHER (treatment INF). 60.66% of all bids in this constellation are lower or equal to 20. If we only look at the subset of bids that were actually decreased, 72.92% of these newly selected bids are below or equal to 20. This in fact indicates that the subjects who decreased their bids seem to realize that if they win with a low signal, the other player has most likely a low signal as well and, hence, they bid exactly for this case (if both players have low signals, the *maximal* value of the good is 20). Equivalently, for high signals paired with the information LOWER there should be an accumulation of bids above 100, when the subjects realize that the other player has most likely a high signal, when losing with a high signal (if both players have high signals, the *minimal* value of the good is 100). For this constellation, a slight accumulation is still noticeable even though it is less distinct. Here only 42.53% of the newly selected bids, which are higher than before, are above or equal to 100.

I also used a Kolmogorov-Smirnov test to check whether the distributions of bids in stage I and II are equal for the different constellations. The results of the test showed that in treatment INF the null hypothesis of equal distributions was always rejected ($ls^H$: $p = 0.000$, $ls^L$: $p = 0.000$, $hs^H$: $p = 0.002$, $hs^L$: $p = 0.000$). On the other hand, for treatment NOINF, the null hypothesis was only rejected for constellation $ls^H$, based on a significance level of 5% ($ls^H$: $p = 0.020$, $ls^L$: $p = 0.136$, $hs^H$: $p > 0.999$, $hs^L$: $p = 0.069$).[30]

| | Bid is in Nash range in stage II (YES [1] or NO [0]) | | | | | | |
|---|---|---|---|---|---|---|---|
| | $ls^H$ | $ls^L$ | $hs^H$ | $hs^L$ | $H$ | $L$ | *Overall* |
| INF | 0.361$^{***}$ | -0.142 | -0.046 | 0.072 | 0.047 | -0.074 | -0.012 |
| | (0.123) | (0.088) | (0.059) | (0.110) | (0.046) | (0.070) | (0.046) |
| Constant | 0.114 | 0.293$^{***}$ | 0.153$^{***}$ | 0.145$^{*}$ | 0.141$^{***}$ | 0.246$^{***}$ | 0.192$^{***}$ |
| | (0.075) | (0.079) | (0.049) | (0.076) | (0.038) | (0.059) | (0.038) |
| Observations | 140 | 376 | 391 | 173 | 531 | 549 | 1080 |
| Subjects | 49 | 70 | 70 | 50 | 70 | 70 | 72 |
| $R^2$ | 0.162 | 0.029 | 0.005 | 0.008 | 0.004 | 0.008 | 0.000 |

Notes: Cluster-robust standard errors (on the subject-level) are in parentheses. Constellation $ls^H$ stands for low signals paired with the information HIGHER. Constellation $L$ stands for all signals paired with the information LOWER. The other constellations are defined analogously.
$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table 2.3: Regression Table - Bid in Nash range (Stage II)

To support the claim that bidders in treatment INF significantly improved their

[30] It is important to note that by using the standard Kolmogorov-Smirnov test, I cannot say whether one distribution first-order stochastically dominates (FOSD) the other one. For this purpose, I would have to consider the test of Barrett and Donald (2003) (special thanks to an anonymous reviewer who pointed this out). For the further analysis, I will, however, focus on an OLS regression model in order to provide a more rigorous comparison of the bidding behavior between the two treatments across stage I and II.

bids in some constellations, compared to the control group NOINF, I ran an OLS regression analysis with clustered standard errors on the subject-level, where I tested whether the bids in treatment INF are more often in the Nash range in stage II, compared to NOINF. The results are presented in Table 2.3.

As we can see there is only a significant effect for constellation $ls^H$. This effect is quite large, indicating that the number of bids in the Nash range in treatment INF is by around 36.1 percentage points larger compared to NOINF. As a placebo test, I ran the same regression for the bids in stage I, before the treatment intervention (see Table A.1 in Appendix A). Here we see no significant difference between the treatments, except for constellation $hs^L$, where we find a small, but significantly negative coefficient. This effect might be explained by small sample properties.

Additionally, Table A.2 in Appendix A reports a further regression, where the dependent variable indicates whether a bid was shifted from one range to the other across the stages (shifted from Nash to unsophisticated range ($-1$); not shifted from one range to the other (0); shifted from unsophisticated to Nash range (1)). The results are in line with the findings from Table 2.3. Furthermore, we see that receiving the information LOWER has an overall negative effect concerning the changing of bids, i.e., more bids are shifted from Nash to unsophisticated range than vice versa.

This leads to the following conclusions. If we only look at those constellations in which the respective information provides a relatively unambiguous hint about whether the opponent's signal is low or high ($ls^H$ and $hs^L$), the *information* treatment (INF) leads to a significant improvement of the bids in constellation $ls^H$ but not $hs^L$, where the coefficient is also positive, but not significant. For the other constellations there seems to be a reverse effect, which is, however, (i) not significant for $hs^H$, and (ii) for $ls^L$ only significant in the specification where I look at the shifting of bids across the bidding ranges (see Table A.2).

### 2.4.3 Profits

The profits of the bidders are calculated as defined by equation (1) in Section 2.2. The change in profits from stage I to stage II for individual $i$ in round $m$ is then given by

$$\Delta\pi_{im} = \pi_{im}^{II} - \pi_{im}^{I} \tag{4}$$

Hence, a positive value of $\Delta\pi_{im}$ implies that the subject's payoff increased from stage I to stage II, whereas a negative value implies the opposite. Table 2.4 reports OLS regressions with clustered standard errors, testing whether the *information* treatment (INF) has a significant effect on the changes in profits from stage I to stage II for different constellations of signal and information.

Overall, we can see that the change in average profits follows the same pattern as the deviation from the Nash bid: if the subjects bid closer to the Nash prediction in stage II, their profits increase, and vice versa. However, the effects, although pointing into the right direction, are only significant for constellations $hs^H$ and $hs^L$ and here only at the 10%-level. The coefficient in constellation $hs^L$ indicates that the subjects in treatment INF increased their profits per auction from stage I to stage II by 5.115 ECU, on average, when they learned that they lost an auction with a high signal, compared to the subjects in treatment NOINF who did not get the information of losing. Conversely, in constellation $hs^H$, we can see that the profits per auction decreased from stage I to stage II by 1.414 ECU, on average, when learning that the bid for a high signal was the winning bid.

| | Change in profits | | | | | | |
|---|---|---|---|---|---|---|---|
| | $ls^H$ | $ls^L$ | $hs^H$ | $hs^L$ | $H$ | $L$ | *Overall* |
| INF | 3.773 | -1.971 | -1.414* | 5.115* | -0.469 | 0.246 | -0.158 |
| | (4.907) | (1.327) | (0.758) | (2.644) | (1.482) | (1.284) | (1.001) |
| Constant | 2.063 | -1.793* | -0.205 | 0.566 | 0.498 | -1.046 | -0.251 |
| | (4.534) | (0.982) | (0.207) | (1.388) | (1.396) | (0.796) | (0.856) |
| Observations | 140 | 376 | 391 | 173 | 531 | 549 | 1080 |
| Subjects | 49 | 70 | 70 | 50 | 70 | 70 | 72 |
| $R^2$ | 0.014 | 0.010 | 0.015 | 0.043 | 0.001 | 0.000 | 0.000 |

Notes: Cluster-robust standard errors (on the subject-level) are in parentheses. Constellation $ls^H$ stands for low signals paired with the information HIGHER. Constellation $L$ stands for all signals paired with the information LOWER. The other constellations are defined analogously.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.4: Regression Table - Change in profits

A major problem in second-price auctions, however, is that the magnitude of the profits is affected by the behavior of the opponents. Hence, the very same winning bid can lead to different profits, depending on the bid of the opponent.[31] For this reason, I also used an alternative and more robust measure for profitability in which

[31]Consider the following example. Player 1 has a signal of 10 and player 2 has a signal of 5. Player 1 submits a bid of 40 in the stage I and player 2 submits a bid of 35. Player 1 wins and his payoff would be −20 in this scenario. If player 1 decreases his bid to 20 in stage II, he receives a payoff of 0 and improves his payoff by 20 ECU. Now suppose player 2 would have chosen a bid of 25 in stage I, with everything else exactly as before. Now player 1 improves his payoff by only 10 ECU, when he decreases his bid to 20 in stage II. Hence, in the first scenario the very same decreasing strategy is twice as profitable.

any positive payoff is transformed into 1 and any negative payoff is transformed into $-1$. In this sense there is only a distinction between whether an auction was won profitable, unprofitable or lost. Henceforth, I will call this *adjusted profit*. In contrast to the actual profit, the magnitude of the adjusted profit is not affected by the opponent's bid which makes it a cleaner measure for sophisticated bidding. The adjusted profit has also a second interpretation: switching auctions with negative payoff into auctions with zero payoff in stage II refers to overcoming a winner's curse. Switching auctions with zero payoff into auctions with positive payoff in stage II refers to overcoming a loser's curse. Thus, the change in the adjusted profit from stage I to stage II can actually be interpreted as the (net) rate of switching unfavorable auctions into favorable auctions. Table 2.5 reports OLS regressions with clustered standard errors, testing whether the *information* treatment (INF) has a significant effect on the changes in adjusted profits from stage I to stage II, for different constellations of signal and information.

| | Change in adjusted profits | | | | | | |
|---|---|---|---|---|---|---|---|
| | $ls^H$ | $ls^L$ | $hs^H$ | $hs^L$ | $H$ | $L$ | *Overall* |
| INF | 0.306** | -0.071* | -0.050** | 0.190** | 0.025 | 0.011 | 0.016 |
| | (0.152) | (0.041) | (0.025) | (0.092) | (0.041) | (0.040) | (0.030) |
| Constant | 0.038 | -0.061** | -0.006 | 0.026 | 0.008 | -0.033 | -0.012 |
| | (0.115) | (0.029) | (0.006) | (0.051) | (0.035) | (0.024) | (0.022) |
| Observations | 140 | 376 | 391 | 173 | 531 | 549 | 1080 |
| Subjects | 49 | 70 | 70 | 50 | 70 | 70 | 72 |
| $R^2$ | 0.074 | 0.012 | 0.019 | 0.039 | 0.001 | 0.000 | 0.000 |

Notes: Cluster-robust standard errors (on the subject-level) are in parentheses. Constellation $ls^H$ stands for low signals paired with the information HIGHER. Constellation $L$ stands for all signals paired with the information LOWER. The other constellations are defined analogously. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2.5: Regression Table - Change in adjusted profits

Here we can see a significant effect for all constellations except for the last three, where low and high signals are pooled, and for the whole sample. For constellations $ls^H$ and $hs^L$ there is a significant and positive effect of the *information* treatment (INF) on the change in adjusted profits. For example, the coefficient of 0.306 in the first column can be interpreted in two ways. (i) In constellation $ls^H$, the subjects in treatment INF have an overall net rate of switching unfavorable auctions into favorable auctions in stage II, which is by 30.6 percentage points larger than that of the subjects in treatment NOINF. Or equivalently, (ii) in constellation $ls^H$, the subjects in treatment INF increased their adjusted profits per auction from stage I to stage II by 0.306 units, on average, compared to the subjects in treatment NOINF. The other coefficients can be interpreted analogously. For constellations $ls^L$ and $hs^H$, there is a negative effect of the *information* treatment (INF) on the change in adjusted profits, making the subjects in treatment INF worse off compared to those in treatment NOINF. However, the positive effects of the *information* treatment (INF) are greater in absolute terms than the negative ones and the coefficient for constellation $ls^L$ is only significant at the 10%-level. But since constellations $ls^L$ and $hs^H$ occur, by construction of the game, more often than $ls^H$ and $hs^L$, the overall effect of treatment INF is not significantly different from the one of treatment NOINF. This indicates that, in total, the subjects who received information about their bids are not better off than those who received no information.

### 2.4.4 Differentiation between low and high signals

By construction of the game and the observed bidding behavior, it is profitable for the bidders to increase the initial bid for high signals and to decrease the initial bid for low signals, in most of the cases. Thus, when receiving either the information HIGHER or LOWER it is important to distinguish between those two kinds of signals instead of following the simple decision rule "decrease when you receive HIGHER and increase when you receive LOWER". The results show that when receiving the information HIGHER, the subjects in treatment INF strongly differentiate between low and high signals. The decrease rate for low signals is 78.7% and only 42.3% for high signals ($p = 0.000$, Fisher's exact test). For the information LOWER, the subjects in treatment INF do not differentiate between low and high signals. The increase rate for low signals is 83.0% and 89.7% for high signals. However, the difference is not statistically significant ($p = 0.168$, Fisher's exact test). After receiving the information LOWER, it seems to be very tempting to increase the initial bid. This can be an indicator of an actual joy of winning (or rather disappointment of losing), but this is also in line with theories about hypothetical thinking. If the initial bid was higher, it was actually the relevant bid. If the initial bid was lower, this is not the case and the subjects have to engage in hypothetical thinking again before coming up with a new bid.[32] For example, a naïve Bayesian updater might correctly assume that the value of the good is higher than expected when receiving the information LOWER and increases his initial bid. The problem is that the new bid is not chosen conditional on winning - this can result in an even

[32] An alternative explanation for this behavior might be a general misunderstanding of second-price auctions. However, this point seems less likely, since in the instructions the mechanism of a second-price auction was explained in detail and the control questions could only be answered when this mechanism was understood correctly.

more frequent appearance of the winner's curse for low signals.

Figure 2.6 reports the fractions of decreasing, increasing and retaining the initial bids in stage II for different constellations of signal and information for the subjects in treatment INF.

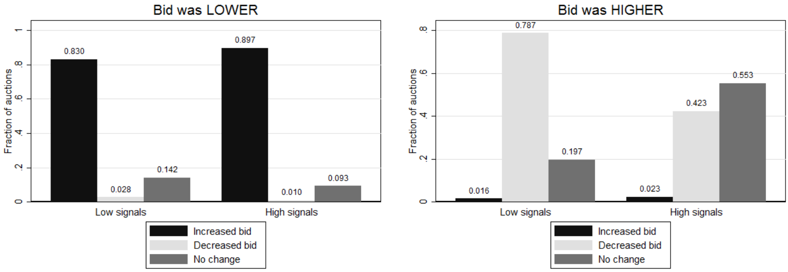



Figure 2.6: Changing of bids in stage II (treatment INF)

We have already seen that most of the bidders in treatment INF, in constellation $ls^H$, seem to realize that winning implies that the opponent also has a low signal and, hence, they select a bid below 20 in stage II. Combined with the pattern that those subjects seem to differentiate between low and high signals when receiving the information HIGHER, it seems to be very plausible to assume that the changing behavior in stage II is to a substantial extent driven by an actual updating of the opponent's signal. This is also in line with the answers of a questionnaire, which was provided after the actual experiment. 79.49% (31 out of 39) of the participants in treatment INF answered that the information they received helped them indeed to get a better estimate of the opponent's signal.

Besides the effect of contingent reasoning, some of the updating in stage II could be explained by spite preferences. In second-price auctions, some bidders might tend to overbid in order to decrease the surplus of the opponent. This is especially the case when the private valuation is low and, hence, the probability of winning the auction (see, for example, Morgan, Steiglitz, and Reis, 2003; Cooper and Fang, 2008). In stage II, the participants compete against a computer mimicking the behavior of their former opponent, hence, in stage II, spite plays no role anymore for the subjects. The prediction therefore would be that the subjects decrease their bids in stage II, especially for low signals - this is the pattern which I observe. In the control group NOINF, 25.5% of the bids with low signal were decreased, but only 13.5% of the bids with a high signal. While this would be in line with an explanation about spite preferences, it has to be noted that both values are considerably smaller compared to treatment INF. I.e., if spite would be the main driver behind the observed behavior, the magnitude of decreased bids in both treatments should be similar. Additionally, it can be argued that spite cannot explain the high increasing rate for low signals paired with the information LOWER in stage II, and the frequent occurrence of underbidding for high signals in stage I.

### 2.4.5 Cursed equilibrium

A further conclusion that can be derived from the previous results is that the overall behavior of the subjects in stage II, especially of those in treatment INF, cannot be explained by cursed equilibrium since a "cursed" bidder assumes by definition that the bids and signals of the opponent are not correlated. Hence, such a bidder should not react on the information of winning or losing because he implicitly assumes that the bid of the opponent provides no valid information about the true value of the good. Additionally, we have the best-response assumption, inherent in cursed equilibrium, stating that the bid of a "cursed" player is already evaluated conditional on winning in stage I. Hence, there would be no need to change the bid in stage II for such a bidder. As already explained in the beginning, this reasoning holds for both fully and partly cursed bidders.

I observed that 97.44% of the subjects in treatment INF changed their bids at least once and that 66.67% of them changed their bids in more than 10 rounds. In treatment NOINF the changing rate is slightly smaller: 90.91% of the subjects changed their bids at least once and only 51.52% of them changed their bids in more than 10 rounds. If we only look at bids that were changed by at least 10 units, the difference between INF and NOINF is more distinct (see Figure 2.7).[33]

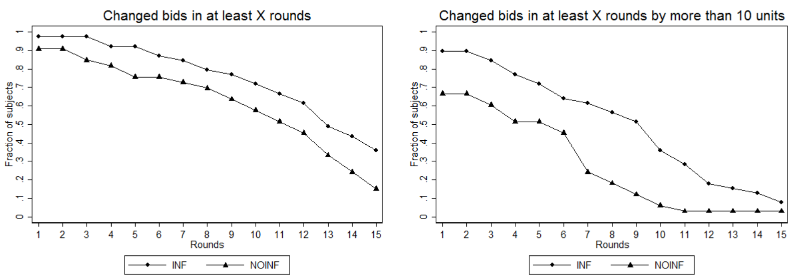



Figure 2.7: Changing of bids across the rounds

[33]These graphs report the fractions of subjects who changed their bids in at least X rounds (left) and in at least X rounds by more than 10 units (right).

So the best-response assumption is violated for the majority of subjects in both treatments. If the bid in stage I would have been an optimal bid, given the subject's beliefs, there would be no need to change the bid in stage II, since the decision problem in a given auction is the same in stage II. For the subjects in treatment INF, we can additionally use the even stronger argument that those subjects actually recognized the informational content of winning or losing an auction. This pattern casts doubts on whether the bidding behavior of the majority of subjects can be explained by cursed equilibrium, unless one assumes that the bidders are suddenly less cursed in stage II or that a bidder can suffer from both: a cursed system of beliefs and the inability of thinking in hypothetical situations. However, it is not clear how or whether the effects of both cognitive mistakes add up. So far, Koch and Penczynski (2018) are the only ones who looked at both combined in a lab setting, but more research is needed especially concerning the interaction of both cognitive mistakes. In general, my findings support the claim of Ivanov, Levin, and Niederle (2010) who stated that bidders in common value auctions might act "as if" they have cursed beliefs.[34]

One obvious concern about the changing behavior of the subjects could be that some subjects changed their bids out of boredom or some experimenter demand effects. To rule out boredom effects over the course of the 15 rounds, I also checked whether the fraction of changed bids increased towards the end of the bidding period in stage II. As we can see in Figure A.8 (Appendix A) there is no increasing trend within stage II, but the fraction of changed bids remains, on average, constant.

### 2.4.6 Analysis on subject level

In this section I will provide a classification of the subjects on an individual level. The key point of interest is to clarify whether there is a heterogeneity in bidding behavior. For this purpose, I define the following five bidder types.

**Sophisticated:** *At least 8 of 15 bids are in the Nash range both in stage I and II, and there are no switches from Nash to unsophisticated range or reverse. I.e., these subjects are sophisticated from the beginning and do not benefit from further information.*

**Improved:** *More bids in the Nash range in stage II compared to stage I. Too few bids in the Nash range in stage I to be qualified as sophisticated. I.e., these subjects actually profited from the information they received.*

**Balanced:** *Same number of bids in Nash range in stage I and II, but at least two switches between Nash and unsophisticated range across the stages. I.e., these*

[34] Ivanov, Levin, and Niederle (2010) were the first to claim that bidders in common value auctions might just act "as if" they have cursed beliefs, since they observed seemingly cursed behavior in a context where belief-based models had few explanatory power.

*subjects sometimes improved and sometimes worsened their bids.*
**Worsened:** *Fewer bids in the Nash range in stage II compared to stage I. Too few bids in the Nash range in stage II to be qualified as sophisticated. I.e., these subjects were worse-off after receiving information.*
**Ignorant:** *Same number of bids in Nash range in stage I and II, and there are no switches between Nash and unsophisticated range across the stages. Too few bids in the Nash range in stage I to be qualified as sophisticated. I.e., these subjects mostly used unsophisticated bids in stage I and did not change their behavior.*

| Bidder type | Treatment | |
|---|---|---|
| | *NOINF* | *INF* |
| *Sophisticated* | 1 | 1 |
| *Improved* | 8 | 14 |
| *Balanced* | 0 | 4 |
| *Worsened* | 9 | 17 |
| *Ignorant* | 15 | 3 |
| ***Total*** | 33 | 39 |

Table 2.6: Number of different bidding types across treatments

Table 2.6 reports the number of subjects in each class for both treatments. As we can see, there are many subjects who either improved or worsened their bids in stage II in treatment INF (31 out of 39). In the control group NOINF a large fraction of subjects used unsophisticated bids from the beginning on and did not change their bids in stage II. This is consistent with the finding that especially the subjects in treatment INF changed their bids. In the next step, I check how the different bidder types performed for different signals.

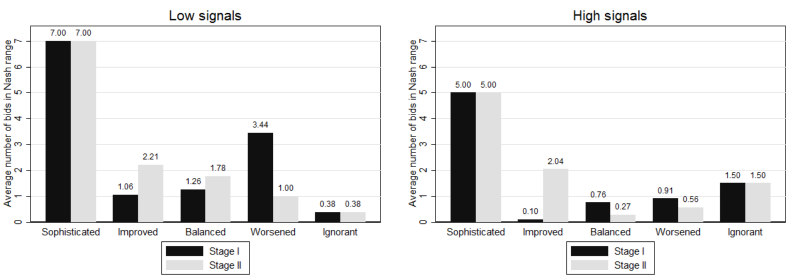



Figure 2.8: Bids in Nash range by bidder type

Figure 2.8 reports the average number of bids in the Nash range in stage I and II of all bidder types for low and high signals (only treatment INF). We can see that those subjects in the *Improved* class, benefited mostly from changing their bids for high signals, i.e., by increasing bids which were too low ex ante. On the other hand, those subjects in the *Worsened* class, incurred the most harm for low signals, i.e., by increasing bids above the optimal level.

This illustrates a potential dilemma: by choosing low bids in stage I, which are potentially in the Nash range, the probability of receiving the information LOWER increases. As we can see, those subjects in the *Worsened* class are the second highest performers in stage I, after the *sophisticated* ones (at least for low signals). However, after receiving the message LOWER, subjects have a strong tendency of increasing their bids, as already shown in Section 2.4.4. This leads to the conclusion that choosing moderate bids in stage I can be harmful, when subjects are tempted to increase their bids after losing an auction without differentiating between low and high signals.

## 2.5 Discussion

This chapter investigates whether subjects in a common value auction perform better when they already learn ex ante, before the final payoffs are known, whether their bid is the winning bid or not - an information bidders in a sealed bid auction usually receive only at the very end of the auction.

First, the overall results show that the majority of subjects indeed reacted on the respective information they received and changed their bids in stage II. From a theoretical point of view this is surprising, since we have to assume that a bid, which was chosen in stage I, was considered as optimal in this context. So in the framework of a second-price auction, there would be no need to change such a bid in stage II, even when learning whether this bid is the winning bid or not. In other words, if a bid was chosen conditional on winning in stage I, there would be no need to change it in stage II.

Second, we have seen that at least in some constellations ($ls^H$ and $hs^L$), the subjects profited from the information they received. In other constellations, however, this effect is reverse and, overall, there is no difference between the subjects in treatment INF and NOINF when we look at average profits. Therefore, additional information can be even negative for the bidders.[35] A crucial problem is that

[35] In related studies like Charness and Levin (2009) and Koch and Penczynski (2018), the optimal behavior was mainly given by choosing a bid as low as possible. In my design, the bidders have to differentiate between low and high signals, and decreasing a bid is not always optimal. This additional hurdle shows that more information can be negative for the bidders when they differentiate only imperfectly between situations in which decreasing (increasing) a bid is rational and those in which it is not.

subjects still often used simple heuristics instead of making strategic choices after receiving feedback about their bid. On an individual level there is evidence that bidders who choose moderate bids in stage I, face some kind of a ‘curse’, since the probability of receiving the message LOWER is higher for them. This in turn leads to a more frequent occurrence of the winner’s curse in stage II.

Finally, there is evidence in the data that at least for the information HIGHER, the changing of bids is not just a rule of thumb, but it rather occurs due to Bayesian updating since there is a much higher decreasing rate for low signals than for high signals. This claim is also supported by the pattern that, for low signals, after receiving the information HIGHER, most of the newly selected bids are below 20. This suggests that the bidders, who won with a low signal, indeed realized that the other bidder has most likely also a low signal.

As a concluding remark, there is to say that mistakes in hypothetical thinking seem to explain a substantial part of irrational bidding behavior in common value auctions. However, even without the necessity of conditioning on winning, there still exists a significant deviation from optimal behavior which remains unexplained.

In the next chapter, I will investigate the presence and the consequences of another cognitive bias, known as correlation neglect, in voting decisions. Contrary to mistakes in hypothetical thinking, I will show in the following chapter that mistakes due to correlation neglect might lead to more efficient outcomes in some environments.

# 3 Correlation neglect and voting[36]

In many voting decisions people have to weigh their subjective preferences about the alternatives against objective information they receive. This assessment is aggravated since the latter may be subject to a double-counting problematic when people treat correlated information sources as independent. This thesis is the first to test the phenomenon of correlation neglect as a potentially welfare-*enhancing* bias in a collective voting decision empirically. The results of my online experiment suggest that (i) information aggregation is improved by this cognitive bias, (ii) the results vary substantially with the risk-attitude of the subjects, and (iii) correlation neglect is present in an environment with perfectly correlated signals and very high salience of the correlation.

Two highly preference-driven elections on globally relevant outcomes dominated the media coverage over the past years. In the *Brexit* referendum, voters had ideological preferences concerning "*leave*" or "*remain*", but at the same time they also received information about the overall pros and cons of each of the options, such as the economic consequences, for example through the media. Similarly, Donald Trump's presidential campaign "*Make America Great Again*" served patriotic preferences of the voters who had to weigh them against public evaluations of the protectionist policies. In both cases a huge of information sources was constantly available. At the same it was nearly impossible to identify common origins of seemingly independent news. This made the voters prone to overestimating the informational content.

Correlation neglect refers to underestimating the degree to which various sources of information may be correlated. Typical examples in the literature are the news media or markets for financial assets (see, for example, Ortoleva and Snowberg, 2015; Eyster and Weizsäcker, 2016; Enke and Zimmermann, 2017).[37] There is an emerging empirical evidence that individuals are prone to this cognitive bias when making decisions in environments where they receive information from various sources which might be correlated. This in turn can lead to highly biased beliefs and inefficient economic outcomes, as for example shown by the experiments of Eyster and Weizsäcker (2016) and Enke and Zimmermann (2017).

However, in a theoretical paper, Levy and Razin (2015a) argued that correlation neglect might have positive effects on the outcome of *collective decision problems*, like an election, due to improved information aggregation. So for example, in the presence of correlation neglect, individuals who participate in a collective decision

[36]This chapter is a slightly modified version of Moser and Wallmeier (2019).

[37]For further evidence on correlation neglect consider, for example, Levy and Razin (2015b), Eyster, Rabin, and Weizsäcker (2015), Ellis and Piccione (2017), Levy, Moreno de Barreda, and Razin (2018), and Hossain and Okui (2018).

might focus more on their available information from various sources and less on their ideological preferences. This can lead to more efficient outcomes, even if the various sources of information are not independent. So an important point, when evaluating the effects of correlation neglect on economic outcomes, seems to be the distinction between individual and collective decisions.

I provide an experimental design, where subjects can either vote for a safe payment or for a risky lottery with the same expected payoff. Hence, the choice for one of the options depends on their risk preferences. Additionally, the subjects receive a bonus for matching an unknown state of the world and a penalty for matching the wrong state, respectively. In my setup the correct state of the world is either a "safe" or a "risky" state. From this perspective my setup is very similar to the theoretical model of Levy and Razin (2015a), where voters have (i) ideological preferences on a left-right spectrum and, additionally, (ii) they receive an imperfect signal about an unknown state of the world which is either *left* or *right*. In my setup, subjects receive either a single signal, two perfectly correlated signals, or two independent signals, which provide information about the unknown state of the world. When the voter's preference and the signal(s) are not aligned, she faces a non-trivial decision problem.

I test the presence of correlation neglect in a between-subject design by comparing the share of voters that follow the signal contrary to their preferences in the treatments with a single signal and two perfectly correlated signals. In both cases the signal transmits the identical informational content such that differences may be only attributed to correlation neglect. In addition, I also estimate the magnitude of the bias by comparing the treatments with two perfectly correlated and two independent signals. The latter presents the case where individuals update the informational content rightfully which serves as the upper bound for correlation neglect.

My overall findings suggest that correlation neglect is also present in environments where groups face a collective decision problem. Subjects who received perfectly correlated information followed more often this information compared to subjects who only received one signal, even though the correlation of signals was clearly recognizable and no calculations were required. Still, I observe a slight difference between the treatment where subjects received two correlated and the one where they received two independent signals as the fraction of subjects following their signals is highest with independent signals.

To the best of my knowledge, this thesis is the first one that investigates the effect of correlation neglect experimentally in a collective decision problem. In contrast to other experimental studies, I include two dimensions in the maximization problem of the subjects. One the one hand, individuals have preferences about one of two options and, on the other hand, they also receive a payoff for matching a specific state of the world. Hence, subjects are not only paid for estimating a certain state of the world exactly as possible, but they also receive an utility from their private preferences which can potentially deviate from the state of the world.

The theoretical basis of this work is the paper by Levy and Razin (2015a). However, the aim of my thesis is not to provide a direct theory testing of their model. Instead, I want to create an environment that includes their main features, since their voting context resembles scenarios of general interest (e.g., Brexit referendum, 2016 United States presidential election) which have not been conclusively investigated in the experimental literature so far. In contrast to Levy and Razin (2015a), I do not have a continuum of voters. Hence, each voter in my experiment is pivotal with a non-zero probability. Implications resulting from this issue will be discussed in Section 3.2. A further difference is that in my setup there can be individuals with extreme preferences who would not vote for the state of the world even for the hypothetical case where the state is perfectly observable. This in not the case in Levy and Razin (2015a) where all individuals prefer to implement the state of the world if the state would be perfectly observable. Additionally, I do not assume any specific utility function, but focus on a revealed preferences approach instead.

## 3.1 Experimental design

### 3.1.1 General setting

In the case of real world elections one can think of numerous preferences which may influence the voting decision and may or may not be aligned with the socially efficient choice. An example are environmental preferences that would be in favor of or against a change to renewable energy. While a voter may have preferences on this topic in general, it is not clear which policy is the efficient alternative for the society at the moment. People have to rely, for example, on estimates about the consequences of fossil fuels for the climate and the costs of the change.

For my experiment, I have two requirements for the application. I need an environment where individuals have preferences which are sufficiently easy to measure and which can be incentivized. Since I have a second dimension, social efficiency, which should influence the vote, I also need an immediate consequence from the vote on the society.

For this purpose, I focused on an approach with risk preferences, since an attempt of measuring political preferences with monetary incentives can lead to highly inaccurate results. The advantage of risk preferences is that we have a sufficiently high variation between individuals which are easy to measure in an online experiment. In addition, the choice of a more (less) risky payment scheme can be directly implemented by generating a payoff for the group. Using risk preferences makes the voting task more abstract, but at the same time clear statements can be made about the effect of correlation neglect, when subjects are affected by personal preferences - which refers to my main research question.

The experiment is divided into two parts. For each subject, one of the parts is randomly chosen to be the payoff-relevant part. In part 1, subjects have to make a choice for either a safe payment or a risky lottery with the same expected payoffs. Hence, the purpose of part 1 is to elicit the risk preferences of the subjects using an approach relying on revealed preferences. In part 2, subjects have to make the same choice between a safe and a risky option, but now they receive an additional payoff for matching a specific state of the world. To measure the prevalence and the magnitude of correlation neglect in part 2, the participants either received one signal (treatment *OneSignal*), two perfectly correlated signals (treatment *TwoCorr*), or two independent signals (treatment *TwoInd*) which gave an imperfect hint about a state of the world. Each subject only participated in of the treatments (between-subject design). Detailed instructions can be found in Appendix F.

### 3.1.2 Part 1

In part 1, subjects are randomly allocated in groups of five people. The task of a group is to implement a payment scheme $W$, with $W \in \{S, R\}$, by a majority vote (i.e., if at least three people vote for the same payment scheme, this option will be implemented). $S$ generates a safe payment of 1 GBP for each group member and $R$ is a risky lottery that generates 0.25 GBP with 90% probability and 7.75 GBP with 10% probability - which will be drawn individually and independently for each group member. Both payment schemes, $S$ and $R$, have the same expected value of 1 GBP. Since this may not be obvious to the participants, I included a control question before the vote where they had to indicate if one of the alternatives yield a higher expected payoff than the other.[38] In this way, I am able to obtain information on the understanding of the mechanisms and to bring all participants to the same level.

| **Majority decision** | **Total payoff** |
|---|---|
| S | 1.00 |
| R | 0.25 (90%)<br>7.75 (10%) |

Table 3.1: Payoff structure in part 1

[38] To facilitate an intuitive understanding of the question, I did not ask for the expected payoff which requires calculations. Instead, I asked if they expect to earn more by sticking to one alternative or the other if they would participate 100 times in this game.

The individuals make their voting decision in private. The purpose of part 1, is to elicit the individual preferences of each subject about the two schemes. Individuals voting for $S$ have preference $p^S$ and individuals voting for $R$ have preference $p^R$.

### 3.1.3 Part 2 (Treatments)

In part 2, the groups again face a voting decision. The basic setup is like in part 1, but now the groups also have an unknown *group state* $G$, with $G \in \{s, r\}$, and $P(s) = P(r) = 0.5$. If the group implements $S$ ($R$) and the state of the group is $s$ ($r$), each group member earns an additional bonus of 0.25 GBP. If the group does not match the state, each group member pays a penalty of 0.25 GBP. The base payoff of both options, $S$ and $R$, is the same as before. Each group participates in one of the treatments *OneSignal*, *TwoCorr*, or *TwoInd*.

**OneSignal** The subjects are not aware of the group state, but each subject receives an *iid* signal, $\sigma_s$ or $\sigma_r$, about the state of the group. The signal is correct in two out of three times, hence, we have $P(s|\sigma_s) = P(r|\sigma_r) = \frac{2}{3}$. After receiving their signal, the subjects are asked to submit a vote for either $S$ or $R$.

**TwoCorr** Same as in *OneSignal*, but now the subjects receive two signals which are perfectly correlated. So either they receive $(\sigma_s^1, \sigma_s^2)$ or $(\sigma_r^1, \sigma_r^2)$. The available information from the signals is the same as in *OneSignal*, however, subjects suffering from correlation neglect might interpret the two signals as independent draws.

**TwoInd** Same as in *OneSignal*, but now the subjects receive two signals from two independent draws. So either they receive $(\sigma_s^1, \sigma_s^2)$, $(\sigma_r^1, \sigma_r^2)$, $(\sigma_s^1, \sigma_r^2)$, or $(\sigma_r^1, \sigma_s^2)$. Here we have $P(s|\sigma_s^1, \sigma_s^2) = P(r|\sigma_r^1, \sigma_r^2) = \frac{4}{5}$ and $P(r|\sigma_s^1, \sigma_r^2) = P(r|\sigma_r^1, \sigma_s^2) = \frac{1}{2}$ (a formal proof can be found in Appendix E).

The payoff structure, which is the same in all treatments, is summarized in Table 3.2.

### 3.1.4 Post-experimental questionnaire

After part 2, the subjects were asked to answer a short post-experimental questionnaire. This involved self-reported risk preferences and two questions about the mechanisms of the game.

For the self-reported risk preferences, I used the approach of Falk, Becker, Dohmen, Huffman, and Sunde (2016), where subjects could indicate their willingness to take risks on scale from 0 to 10, with a 0 meaning "completely unwilling to

| Majority decision | Group state | Base payoff | Bonus/Penalty | Total payoff |
|---|---|---|---|---|
| S | s | 1.00 | 0.25 | 1.25 |
| S | r | 1.00 | −0.25 | 0.75 |
| R | r | 0.25 (90%)<br>7.75 (10%) | 0.25 | 0.50 (90%)<br>8.00 (10%) |
| R | s | 0.25 (90%)<br>7.75 (10%) | −0.25 | 0.00 (90%)<br>7.50 (10%) |

Table 3.2: Payoff structure in part 2

take risks" and a 10 meaning "very willing to take risks". Eliciting the risk preferences in this way, additional to the vote in part 1, has two advantages. First, I can deduce from the correlation of the vote for a risky payment and the self-reported risk whether the participants understood the mechanism and actual risk preferences influence the decision in my setting. Second, I have a measure that offers a larger variation in the preferences compared to the binary choice. This additional information is helpful to identify the role of the size of the risk preferences. For example, if an individual is extremely risk-loving, then even if she is prone to correlation neglect, a signal suggesting a bonus for implementing the safe alternative may not be sufficient to make her switch to "safe". Conversely, if she was leaning towards a risky option only moderately, the signal may be sufficient for her to vote for "safe". Therefore, relying only on the chosen option in part 1 maybe too imprecise.

### 3.1.5 Implementation

**Framing** For my research question, I need preferences to be as clearly identifiable as possible. Therefore, I chose not to use any context-specific framing. Although this could be helpful to increase the understanding of the instructions (see, for example, Alekseev, Charness, and Gneezy, 2017), it has the potential downside that I measure a mixture of preferences. For example, I could use a hypothetical real world context, say a choice between two environmental policies - one with a risky outcome and one with a safe outcome. In that case the participant may have an idea what is the more beneficial strategy in the real-world. This could (i) be not aligned with her general risk preferences and (ii) might lead her to neglect the identical expected profit in this experiment.

Also, I aimed at reducing the experimenter demand effect that could arise from using the terms "safe" and "risky". Therefore, I chose to name the alternatives "YELLOW" and "PURPLE", while the states of the world are "yellow" and "purple", respectively (see also Appendix G).

**Calibration** I have three calibration choices to make for this experiment. The first is the spread of the risky payment. I aimed for a spread that roughly splits the sample in half. If the spread is too close to the safe option, I might observe a large share of individuals who choose the risky option although they consider themselves as rather risk-averse. Alternatively, if the spread is too steep, the opposite may be the case where even risk-loving persons still prefer the safe option.

The second choice concerns the size of the bonus (penalty), i.e., the strength of the incentive to follow the signal. If the bonus (penalty) is too high, it dominates the preferences of the electorate such that nobody would vote against their signal. In that case, correlation neglect would not show up in the behavior since everybody votes according to the signal for any informational content. On the other hand, an incentive which is too low would lead voters to ignore the signal and to just rely on their preferences. I used two pilot sessions to calibrate the parameters concerning the risk spread and the size of the bonus.

Third, I need to choose the spread of signal strength for the cases when receiving one signal and two independent signals, respectively. The relatively narrow range between 66.67% and 80% is driven by the simple and intuitive mechanism I used to provide the individuals with the signals.

**Control questions** I used control questions during the instructions to keep track of the individuals' understanding of the mechanisms. To get a precise estimate of the understanding, (i) I made the answers optional, (ii) I provided at least three answer alternatives, and (iii) I always set the default to "no answer" to reduce the likelihood of randomly correct answers. Instead of employing an iterative mechanism that allows the participant only to proceed once she answered correctly, I provided a detailed answer page that explained the correct solution. While the first alternative could be solved by trying all options randomly, the latter aims at bringing all participants to the same level of understanding.

**Procedural details** The experiment was programmed in *LimeSurvey* and conducted on the research platform *Prolific*. Overall, I had 600 participants. 200 of the subjects participated in treatment *OneSignal*, 200 of them in treatment *TwoCorr*, and 200 participated in treatment *TwoInd*. The average completion time was approximately 10 minutes and the subjects earned, on average, 2 GBP including a participation fee of 1 GBP.

## 3.2 Properties of the game

In this section, I will provide propositions about the properties of underlying game. This includes considerations about the optimal behavior in the voting game, the effect of information aggregation, and assumptions concerning pivotality.

**Proposition 6.** *Given that the group state G would be known, it is always optimal to implement the respective group state, under the assumption of risk-neutrality.*

The proof of Proposition 6 is straightforward and, hence, omitted. To see that matching the state of the world is optimal, the expected values of the total payoffs in Table 3.2 can be compared. Subjects voting according to their preference $p^W$ and against the signal $\sigma_G$ can do this for two reasons: (i) the preference for one of the options is so strong, such that the individual still prefers the initial option chosen in part 1 - even if the signal would be perfect; (ii) the belief in the signal is not large enough, such that the individual prefers to stick with his initial vote in part 1.

With my design, I cannot disentangle between the motives (i) and (ii). However, the fraction of subjects who vote against the signal for reason (i) should, on average, be identical across the treatments since the choice of these subjects is not affected by any signals. What I am interested in is the fraction of subjects who deviate from their signal because of reason (ii) across the treatments to measure the effect of correlation neglect.

Given that the subjects do not suffer from any correlation neglect, there should be no statistical difference between *OneSignal* and *TwoCorr*, because the received information is the same in both cases. Only in the case of correlation neglect we should observe a significant difference. I am also able to measure the magnitude of correlation neglect. Under the assumption of full correlation neglect for all subjects, there should be no difference between the treatments *TwoCorr* and *TwoInd*. However, if only some subjects are affected by correlation neglect, the fraction of subjects, who stick to their preference and vote against the signals in *TwoCorr*, should be larger compared to *TwoInd*, but still smaller than in *OneSignal*.

**Proposition 7.** *In a group of five members, where every member receives an iid signal $\sigma_G$ with $P(G|\sigma_G) = \frac{2}{3}$, and every member votes according to her signal, the state of the group G will be matched with ∼79% probability, when the group state is implemented with simple majority.*

The proof of Proposition 7 is provided in Appendix E. Groups where individuals receive imperfect - but independent - signals are more likely to match a given state of the world via a majority vote than a single individual who receives only one signal of the same quality. This is due to effect of information aggregation which is also mentioned in Levy and Razin (2015a). A group of five can increase the probability of matching the state of the world from 66.67% to around 79%, when all group members vote according to their signal.

As indicated in the beginning of this chapter, my voters are pivotal with non-zero probability in contrast to Levy and Razin (2015a). However, in my setup being pivotal does not provide further information about the state of the world under some mild assumptions which are stated below. Hence, the voters in my experiment can opt for one option or the other independently of the behavior of the other group members. This is summarized in Proposition 8.

**Proposition 8.** *Conditional on being pivotal, the probability for one state or the other does not change, given that all other group members (i) vote according to their signal or (ii) vote according to their preference.*

The proof of Proposition 8 is provided in Appendix E. When conditions (i) or (ii) are violated, cases might exist where being pivotal is informative. Consider for example an individual who has preference $p^S$ and believes that most of the other group members also have the same preference. If this individual further believes that some, but not all, of the other players with preference $p^S$ would vote for $S$ even when the signal is $\sigma_R$, being pivotal can now indicate that most of the other group members received a signal $\sigma_R$ and, hence, that state $r$ is more likely. Because if the state would have been $s$, most of the players would have received a signal $\sigma_S$, and those with preference $p^S$ will definitely vote for $S$. Hence, being pivotal is more unlikely in this scenario. However, even under such conditions, there are no differences between the treatments, unless the subjects have specific beliefs concerning the presence of correlation neglect of the other group members. This in turn would imply that the subjects are sufficiently sophisticated and able to recognize the problem of correlation neglect. For these reasons, I assume in my analysis that any differences between treatment *OneSignal* and *TwoCorr* cannot be explained by considerations regarding pivotality.

## 3.3 Results

In this section, the results of the experiment are presented. First, I report the ex ante preferences of all subjects before the treatment intervention. Second, I will exhibit behavioral patterns in my data which indicate that a substantial fraction of the subjects in my setup are prone to correlation neglect, treating perfectly correlated signals as independent draws. Here, I will also provide a subgroup analysis by distinguishing between the subjects with preference $p^S$ and those with preference $p^R$. As I will show, there is quite a heterogeneity between these types, since those with preference $p^S$ vote considerably more often according to their received signal(s) compared to those with preference $p^R$. However, the degree of correlation neglect is similar for both groups. Here, I also looked at the relationship between self-reported risk preferences and the preferences elicitated in part 1 of the experiment. Finally, I will provide some additional analysis for those subjects in treatment *TwoInd* who receive mixed signals (i.e., either $(\sigma_s^1, \sigma_r^2)$ or $(\sigma_r^1, \sigma_s^2)$). Overall, I observe patterns of correlation neglect in my data, however, I am not able to find strong statistical evidence for this issue.

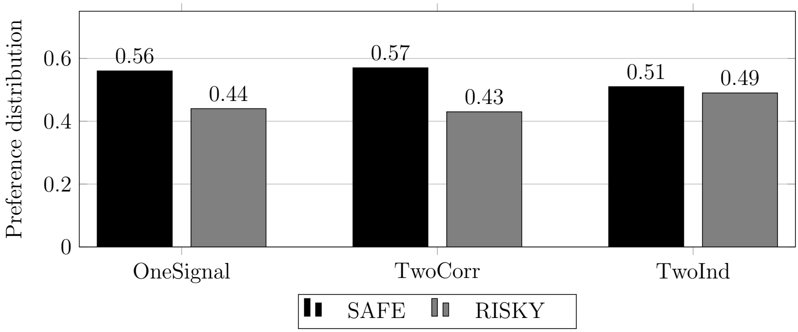



Figure 3.1: Preferences across treatments

Figure 3.1 reports the preferences of the subjects across treatments. Overall, I can observe that the preferences for $S$ and $R$ are quite balanced among the subjects with a slightly more frequent appearance of preference $S$. Between the treatments there is no significant difference in ex ante preferences (Fisher's exact test, pairwise: OS vs. TC, OS vs. TI, TC vs. TI). Comparing these results to the self-reported risk preferences offers two insights. First, the mean values across treatments (5.40 for *OneSignal*, 5.26 for *TwoCorr*, and 5.24 for *TwoInd*) show also moderate risk attitudes and they are consistent with the votes in part 1. The mean risk attitude for the subjects with preference $p^S$ is 4.75 and the mean risk attitude for the subjects with preference $p^R$ is 5.96 ($p = 0.000$, two-sample $t$-test). Second, a Kolmogorov-Smirnov test reports no significant difference between the respective distributions. This means potential treatment differences are not driven by imbalanced subject groups with respect to the shape of the risk preferences.

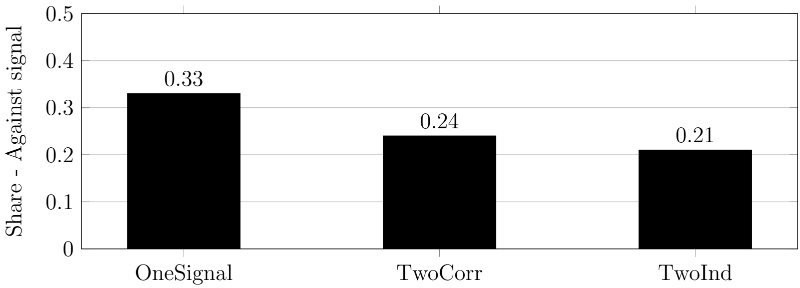



Figure 3.2: Voted against signal when signal is different to preference

Figure 3.2 reports the share of subjects who voted against their signal and according to their preference, split up by treatment. Consequently, I restricted my sample to those subjects whose signal $\sigma_G$ (or signals $(\sigma_G^1, \sigma_G^2)$) and preference $p^W$ were different. In treatment *TwoInd*, I further restricted my attention to the cases, where the subjects received the same two signals, i.e., either $(\sigma_s^1, \sigma_s^2)$ or $(\sigma_r^1, \sigma_r^2)$, to make the results comparable. As a consequence, the number of applicable observations for my analysis is reduced to 98 in *OneSignal*, 97 in *TwoCorr*, and 57 in *TwoInd*.

In Figure 3.2, we see that in the case of only one signal, approximately 33% of the subjects vote against their signal and in favor of their preference. In the case of two perfectly correlated signals, this rate decreases to 24%. When receiving two independent signals the rate goes further down to 21%. The differences between the treatments, although pointing into the right direction, are not statistically different (based on a significance level of 5%) when using a 1-sided Fisher's exact test (OS vs. TC, $p$-value: 0.110; OS vs. TI, $p$-value: 0.086; TC vs. TI, $p$-value: 0.432).

In conclusion, I can say that I am not able to report a statistical treatment effect based on a significance level of 5%. I observe patterns of correlation neglect in my data, however, further empirical evidence would be needed to speak of a clear effect.

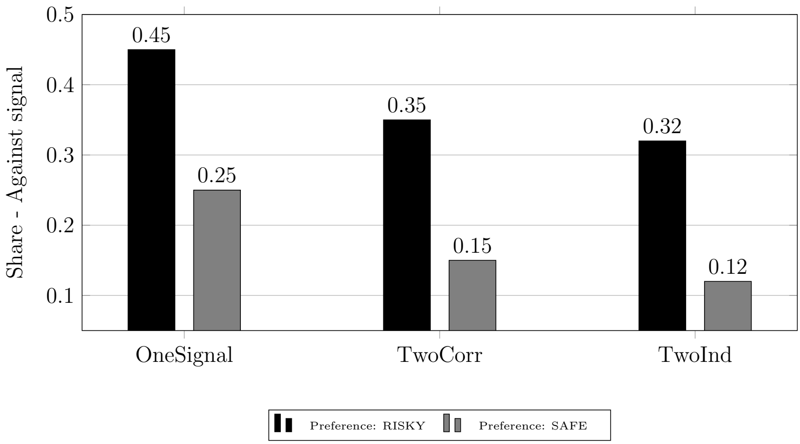



Figure 3.3: Votes by treatment when signal is different to preference

Figure 3.3 reports the share of subjects who voted against their signal and according to their preference, split up by preference and treatment. Here we can observe a substantial heterogeneity between the subjects with preferences $p^R$ and $p^S$, respectively. Subjects with preference $p^R$ vote considerably more often against their signal compared to those with preference $p^S$. In the case of one signal, 25% of the subjects, with a preference for the safe payment scheme $S$, vote against their signal. In the group of those with a preference for the risky payment scheme $R$, 45%

vote against their signal. Overall, we can observe that for both preferences, $p^S$ and $p^R$, the rate of voting against the signal decreases when subjects receive two perfectly correlated signals instead of one single signal (to 15% when having preference $p^S$ and to 35% when having preference $p^R$). In the case of two independent signals, this rate goes further down to 12% when having preference $p^S$, and to 32% when having preference $p^R$.

Here again, the differences between the treatments are not statistically significant when using a 1-sided Fisher's exact test (Preference RISKY, OS vs. TC, $p$-value: 0.249; Preference RISKY, OS vs. TI, $p$-value: 0.228; Preference RISKY, TC vs. TI, $p$-value: 0.512; Preference SAFE, OS vs. TC, $p$-value: 0.131; Preference SAFE, OS vs. TI, $p$-value: 0.126; Preference SAFE, TC vs. TI, $p$-value: 0.517).

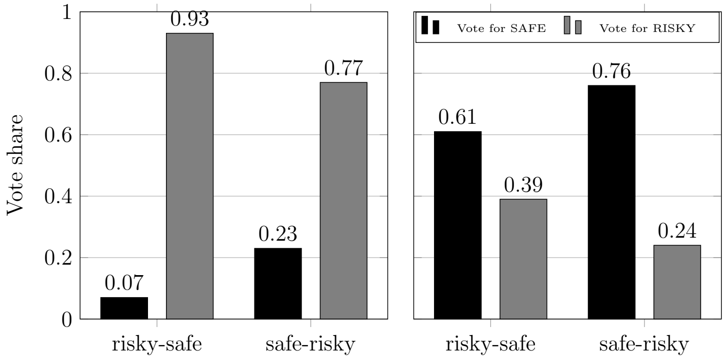
To get a deeper insight about the effect of risk preferences on the choice on one of the options, I also looked at the mean values of self-reported risk attitudes split up by preference and voting decision (see Table 3.3). The overall pattern is consistent with my previous findings. In general, the subjects with preference $p^S$ report, on average, a more risk-averse attitude compared to the subjects with preference $p^R$. Furthermore, we see that subjects who deviate from their preference and vote according to their signal(s) are (i) relatively more *risk-loving* in case of preference $p^S$ (4.19 vs. 5.00; $p = 0.056$, two-sample $t$-test) and (ii) relatively more *risk-averse* in case of preference $p^R$, respectively (6.30 vs. 5.65; $p = 0.097$, two-sample $t$-test). This strengthens my claim that the choice for one option or the other is actually affected by the risk preferences of the subjects and that subjects with more extreme risk attitudes vote more often against their signal(s).

| **Risk preference** | **Mean self-reported risk** | |
|---|---|---|
| | **Vote for preference** | **Vote for signal(s) $\neq$ preference** |
| *Preference SAFE* | 4.19 | 5.00 |
| *Preference RISKY* | 6.30 | 5.65 |

0 = completely unwilling to take risks, 10 = very willing to take risks

Table 3.3: Self-reported risk preferences

A remaining question is, whether correlation neglect can be actually beneficial in my experimental setup. By construction of the game, groups where the individuals vote more often according to the signal(s), generate, on average, higher earnings. Hence, from the perspective of expected earnings it is beneficial for the subjects when they overestimate the informational content of two perfectly correlated signals and vote more often according to them. This leads to fewer cases of uniformed voting and, as a result, the state of the world is matched more often. In this sense, I have shown that, similar to Levy and Razin (2015a), it is possible to create an environment where correlation neglect can have a welfare-enhancing effect.




(a) Preference: RISKY (b) Preference: SAFE

Figure 3.4: Votes by preference when signals are mixed

As a last step of my analysis, I look at those subjects in treatment *TwoInd* who received mixed signals. While this is not directly related to my initial research questions, I am interested at this point whether I observe the tendency that one of the signals is more salient for the voters and, hence, has a stronger influence on their voting decision. Figure 3.4 reports the fractions of subjects who voted for $R$ or $S$ when receiving mixed signals in treatment *TwoInd*. I can observe the pattern that most of the subjects stick with their preference, especially those with preference $p^R$, but when deviating, the first signal has a higher impact on their choice. Also here, it should be noted that the differences are not statistically significant.

## 3.4 Discussion

I provided an experiment where individuals participated in a collective voting decision and faced a trade-off between personal preferences and matching a state of the world. Subjects received either one signal, two perfectly correlated signals, or two independent signals which gave an imperfect hint about the state of the world. I find weak evidence that subjects are biased due to correlation neglect, even though the voting problem is very simple and no sophisticated calculations are required. This in turn leads to the pattern that subjects in *TwoCorr* match the state of the group more often compared to those in *OneSignal*, which consequently leads to higher earnings for the subjects in this group by the design of the game.

A general implication that correlation neglect is beneficial in elections with majority vote should, however, be taken with caution. The positive effect strongly depends on the structure of the correlated signals. In my setup the number of correlated signals was identical for both states of the world and there were no asymmetries. A typical problem of correlated information in real-life scenarios is, however, that some sources of information are more numerous than other sources. For example, before the Brexit referendum articles in newspapers supporting "*Leave*" were more spread compared to articles supporting "*Remain*".[39] For a common voter it is barely possible to observe the entire correlation structure of such news and to calculate how much weight she should put on each of the articles. In this context, a heuristic of simply following the most widely used newspapers is not necessarily optimal.

Overall, I can confirm the findings of related studies concerning correlation neglect, by showing that this cognitive bias is very robust and does also occur in environments where the correlation of signals is presented in an obvious way. Since I do only find weak evidence for my claims, further research might be necessary to fully confirm the patterns I observed.

In the next chapter, I will deviate from the topic of cognitive biases and heuristics and, instead, I will present research concerning leadership and preferences for cooperation within the framework of a dynamic public goods game. Hence, I leave the area of non-standard beliefs and focus on non-standard preferences now.

[39] UK newspapers' positions on Brexit (2016, May 23). Retrieved May 12, 2019, from http://www.ox.ac.uk/news/2016-05-23-uk-newspapers-positions-brexit

# 4 Leadership in dynamic public goods games[40]

Are leaders important for the progress of societies? Jones and Olken (2005) find evidence for this claim regarding national leaders and show that they are an important determinant for economic growth. In general, economic growth in developed countries is often linked with cooperation processes, such as the development and spread of new technologies and knowledge, which requires, in many cases, the joint efforts of several actors. A key question here is whether leadership expressed through "leading-by-example" (Hermalin, 1998) fosters stable cooperation in such partnerships. This is particularly relevant for current urgent problems, such as climate change, and the countries' joint activities concerning environmental protection.

In experimental economics, a common tool to analyze cooperation problems are static public goods games (PG games henceforth). Most of the time, cooperation, however, is not a one-time affair. Instead, stakeholders commonly interact many times with each other and dynamic dependencies exist such that contributions in former rounds have an impact on the public good supply in the current round. *Endowment carryover* is an example for a dynamic dependency which generates scope for endogenous growth and inequality. In such an environment, it is possible for individuals to reinvest cooperation profits from the previous period into further cooperation. For example, public infrastructure can ensure that higher profits are made and those profits, as a consequence, can be reinvested. A similar pattern holds for companies' R&D investments or, in a broader context, the evolution of societies (Gächter, Mengel, Tsakas, and Vostroknutov, 2017).

I am interested whether *leading-by-example* yields an amelioration in public good supply considering such a dynamic environment with endowment carryover. In my setup, leading-by-example means that a group member decides first and the other group members can observe his or her decision right before they make their own decision. Hence, the leader has the opportunity to precede by setting a (good) example. I use a plain setting with a random subject being determined as a leader (exogenous leadership). The implementation that is probably most close to mine is that of Güth, Levati, Sutter, and Van Der Heijden (2007). In their "fixed treatment", one of four group members is randomly selected to be the leader and remains in that position. This exogenous procedure offers a simple and clear-cut treatment.[41] I designed the dynamic dependency in a way that is close to that of Gächter, Mengel, Tsakas, and Vostroknutov (2017), having the following features: (i) there is no consumption until the last round, i.e., the entire wealth can be reinvested at the beginning of a new

[40]This chapter is a slightly modified version of Eichenseer and Moser (2019b).

[41]Furthermore, at least in a static game, there seems to be no differences with respect to contributions between a fixed and a rotating leader (see Güth, Levati, Sutter, and Van Der Heijden, 2007).

round and (ii) endogenous endowments are determined by previous contributions. As a result, this variant of a public goods game allows for both, endogenous growth and endogenous inequality.

With my experiment, I contribute both to the literature on leadership in public goods experiments (see Eichenseer, 2019, for a review) and the literature on dynamic public goods games (e.g., Sadrieh and Verbon, 2006; Battaglini, Nunnari, and Palfrey, 2016; Rockenbach and Wolff, 2017; Gächter, Mengel, Tsakas, and Vostroknutov, 2017). My results indicate that a positive effect of leading-by-example is present in a dynamic setting with endowment carryover. Moreover, the presence of a leader also significantly reduces the within group inequality. Regarding the literature on leadership in public goods games, I provide a further contribution by establishing a link between behavioral types of the leader and leadership success, which complements existing research (e.g., Gächter, Nosenzo, Renner, and Sefton, 2012). Using a sequential prisoner's dilemma for classification (Kosfeld, 2019; Eichenseer and Moser, 2019a), I show that an important predictor for the success of a group is the behavioral type of the leader.

The remainder of this chapter is organized as follows. Section 4.1 offers a brief review of the related literature followed by Section 4.2 which describes the experimental design. Section 4.3 gives an overview on hypotheses and research questions whereas Section 4.4 presents and discusses the results of the experiment. Section 4.5 concludes the chapter.

## 4.1 Literature review

Many (textbook) examples of public goods have in common that they take years to accumulate and provide streams of benefits in the long run, however, at the same time, they also require consecutive expenditures either to improve or to maintain their levels (Battaglini, Nunnari, and Palfrey, 2016). Consequently, it is important to know whether major results of static public goods also hold for dynamic variants, which are often closer to reality. In experimental economics, an emerging literature has attempted to fill this gap in the past years. Sadrieh and Verbon (2006) analyze a setting where social output today determines production possibilities tomorrow (endowment carryover), and social output is distributed unequally. Noussair and Soo (2008) study a dynamic public goods game with a MPCR (*marginal per capita return*) carryover: the return on contributions is a function of decisions in previous periods. This characteristic leads to the effect that, in most cases, the usual pattern of declining contributions over time does not turn up. Cadigan, Wayland, Schmitt, and Swope (2011) investigate whether subjects' behavior in a PG game with a stock level carryover comes close to the qualitative predictions of Markov Perfect

Equilibria. Gächter, Mengel, Tsakas, and Vostroknutov (2017) and Rockenbach and Wolff (2017) both investigate the effects of punishment in a dynamic public goods game with endowment carryover. While punishment in a static setting is usually very successful, there are drawbacks in a dynamic setting. Punishment results in fewer resources available for the punished ones, which could be invested in further cooperation. In consequence, it reduces the potential for future cooperation gains (Rockenbach and Wolff, 2017). To sum up, it is not obvious that standard recipes to improve cooperation do really work well in a dynamic setting.

This question also holds for *leading-by-example*. To the best of my knowledge, there exists no paper in the experimental literature that looks at the effect of leadership in a dynamic public goods game. Restricting my attention to experiments in which a leader is exogenously determined and does not enjoy formal power or superior information, I can summarize previous results in static public goods games as following: exogenous leadership significantly increases contributions to the public good in Moxnes and Van der Heijden (2003), Güth, Levati, Sutter, and Van Der Heijden (2007), Levati, Sutter, and van der Heijden (2007), Pogrebna, Krantz, Schade, and Keser (2011), Dannenberg (2015), and McCannon (2018) while insignificant effects are reported in Haigner and Wakolbinger (2010), Gächter and Renner (2018), Sahin, Eckel, and Komai (2015), and Gürerk, Lauer, and Scheuermann (2018). A significant negative effect is reported in Rivas and Sutter (2011). I can therefore state that in the majority of cases, exogenous leadership yields an improvement in public good contributions.[42]

Leaders who set a good example have a positive impact on followers. A common observation from static PG experiments is that leaders' and followers' contributions are correlated to a large degree (see, for example, Arbak and Villeval, 2013). In addition, leaders also greatly shape the expectations of the followers (Gächter and Renner, 2018). However, the literature also shows that followers systematically contribute less than leaders (see, for example, Güth, Levati, Sutter, and Van Der Heijden, 2007; Haigner and Wakolbinger, 2010; Rivas and Sutter, 2011; Arbak and Villeval, 2013; Dannenberg, 2015; Cappelen, Reme, Sørensen, and Tungodden, 2016; Gächter and Renner, 2018). As a result of this selfish-biased conditional cooperation (Neugebauer, Perote, Schmidt, and Loos, 2009), leaders are relatively worse off giving rise to a 'leader's curse' (Gächter and Renner, 2018). Regarding the question of who makes a good leader, Gächter, Nosenzo, Renner, and Sefton (2012) try to give an answer by classifying subjects in a one-shot strategy method setting. They

[42] Larger effects of leadership usually show up when it is endogenous (Rivas and Sutter, 2011; Haigner and Wakolbinger, 2010), when the leader is endowed with formal power (Güth, Levati, Sutter, and Van Der Heijden, 2007; Levati, Sutter, and van der Heijden, 2007; Gürerk, Irlenbusch, and Rockenbach, 2009; Sutter and Rivas, 2014), or when the leader has an informational edge (Potters, Sefton, and Vesterlund, 2005; Levati, Sutter, and van der Heijden, 2007).

elicit within-subject leader and (conditional) follower contributions as well as their beliefs. The main finding is that subjects who would behave conditionally cooperative as followers, give significantly more if they are leaders, even after controlling for optimism. As a result, groups perform best when led by weak or strong conditional cooperators in this strategy method setting.

## 4.2 Experimental design and procedures

My experiment consists of two parts. Part I, which is a sequential prisoner's dilemma, is the same for all subjects. I used this decision task to classify subjects' cooperation types. In part II of the experiment, subjects performed a dynamic public goods game which was either played *simultaneously* without leader (*NOLEAD*) or *sequentially* with a leader who moved first (*LEAD*). I used a between-subject design, hence, subjects either participated in *NOLEAD* or in *LEAD*.

### 4.2.1 Part I

**Eliciting conditional cooperation types**

Before entering one of the two main treatments (*NOLEAD* or *LEAD*), I elicit conditional cooperation types of all subjects by a sequential prisoner's dilemma.[43] I matched the subjects in groups of two people and both players are endowed with 1 Euro.

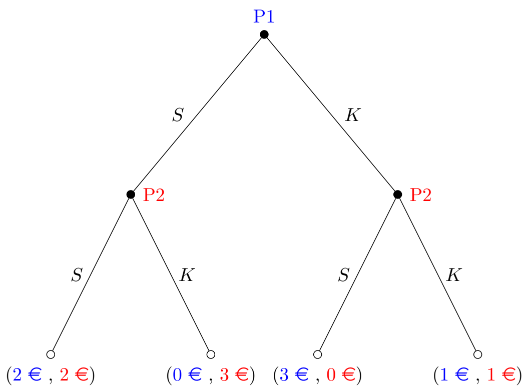



Figure 4.1: Payoff structure of the sequential prisoner's dilemma

First, Player 1 decides whether to send his Euro to Player 2 (choose action $S$) or keep the Euro for himself (choose action $K$). After observing his choice, Player 2 can

[43]This method is used for example by Miettinen, Kosfeld, Fehr, and Weibull (2017) and described in further detail in Kosfeld (2019) and Eichenseer and Moser (2019a).

decide, conditionally on the choice of Player 1, whether to send 1 Euro (action $S$) or keep it (action $K$). If the Euro is sent to the other player, it doubles. Therefore, a social optimum is reached when both players send their Euro. However, maximizing their own payoffs means that none of the two players send their Euro. Hence, the decision situation of part I resembles a sequential prisoner's dilemma. Figure 4.1 depicts the payoff structure of this game.

All subjects state their decisions for being either Player 1 or 2 (strategy method) and they are randomly allocated to one of these roles at the end of the experiment and paid accordingly. In addition, I asked them how many percent of participants they consider to choose the cooperative action when Player 1 cooperates/defects to elicit their beliefs about the other players' behavior.[44] The set of strategies, $X_i$, in this game for Player 2 is given by $X_i = \{SS, KK, SK, KS\}$. Based on the participants' conditional second mover's choice, I can classify subjects as *altruists* (unconditional cooperators), *conditional cooperators* (cooperate only if the first-mover cooperates), *free-riders* (never cooperate) and *mismatchers* (play the opposite of the other player) as shown by Table 4.1.[45] [46]

| ***Cooperation type*** | ***Behavior & Strategy*** |
|---|---|
| Conditional cooperator (CC) | Cooperates only if the other player also cooperates; $SK$ |
| Selfish (SF) | Never cooperates (free-rider); $KK$ |
| Altruist (AL) | Always cooperates; $SS$ |
| Mismatcher (MM) | Does the opposite of the other player; $KS$ |

Table 4.1: Cooperation types

### 4.2.2 Part II

**Simultaneous treatment (*NOLEAD*)**

In the main part of my experiment, subjects play a dynamic public goods game in one of two treatments (between-subject design), either *NOLEAD* or *LEAD*. The earnings of a given round serve as the endowment for the next round.

| **Period** | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| **Round** | 1 2 3 4 5 6 7 | 1 2 3 4 5 6 7 | 1 2 3 4 5 6 7 | 1 2 3 4 5 6 7 |

Table 4.2: Structure of periods and rounds

[44] I did not incentivize this guessing task in order to keep cognitive load at a moderate level.

[45] In my experiment, I only had one subject which I classified as *mismatcher*. Since this subject responded to cooperation with selfish behavior, I also classified it as a *selfish* type for my further analysis.

[46] A balancing table, reporting the distribution of types across treatments, can be found in Appendix H (see Table H.1).

In both conditions, subjects are randomly assigned to groups of four people, which stay the same for one period consisting of seven rounds. After each period, subjects are randomly rematched. In total, the game is played for four periods (see Table 4.2). In round 1 of every period, each participant $i$, with $i \in \{1, 2, 3, 4\}$, is endowed with $E_i^1 = 20$ Taler (exchange rate 1 Euro = 10 Taler), which he can either keep for himself in his "private account" or contribute to the public good labeled as a "group account" ($g_i$ denotes the individual contribution to the group account). The MPCR for the group account is 0.375 which means that the group account has a return of 1.5.

Hence, the earnings at the end of round 1 are given by:

$$E_i^2 = E_i^1 - g_i^1 + \frac{1.5}{4} \sum_{j=1}^{4} g_j^1$$

which serve as the endowment in round 2. Consequently, the endowment for round $t$ is given by:

$$E_i^t = E_i^{t-1} - g_i^{t-1} + \frac{1.5}{4} \sum_{j=1}^{4} g_j^{t-1}$$

and the final payoff of a period is given by the endowment after round 7, which is:

$$\pi_i = E_i^7 - g_i^7 + \frac{1.5}{4} \sum_{j=1}^{4} g_j^7.$$

In each round of a period, subjects simultaneously make their decisions without knowledge of other participants' contributions prior to taking their own contribution decision in *NOLEAD*. After each round, they are informed how many Taler the other group members contributed individually and about their and the others' new endowment. After finishing a period, subjects are randomly rematched to new groups of four. One of the four periods is randomly chosen for payment (random incentive mechanism).

**Sequential treatment (*LEAD*)**

Treatment *LEAD* is similar to treatment *NOLEAD*, except that one subject is randomly allocated to the role of a leader at the beginning of a period. This participant (leader) moves first in all rounds of a period. In each round, the other subjects (followers) are informed about the leader's contribution and subsequently decide simultaneously upon their own contributions. In every period a new leader is randomly selected. Everything else is identical to *NOLEAD*.

### 4.2.3 Participants and procedures

The experimental sessions were conducted in the Regensburg Economic Science Lab (RESL) in February 2018 using zTree (Fischbacher, 2007) for programming and Orsee (Greiner, 2004) for recruitment. 92 participants (45 men and 47 women; mean age: 23), most of them enrolled in business administration, economics, or a related subject, took part in four experimental sessions with a minimum of 20 and a maximum of 24 subjects per session. Before entering the lab, participants were randomly assigned to a cabin with a computer. For both parts of the experiment, I provided participants with written instructions as well as a verbal summary that was read aloud (translated instructions are available in Appendix I). Subjects were not aware of the content of part II of the experiment before finishing part I. I payed participants in Euro in private at the end of the experiment. In total, the experiment lasted about 75 minutes and generated average earnings of about 14.12 EUR per subject (including a show-up fee of 4 EUR).[47]

## 4.3 Hypotheses and research questions

### 4.3.1 Wealth

Under the assumption of selfish, payoff-maximizing players, the unique equilibrium of the simultaneous game, *NOLEAD*, is that all players contribute zero in each round (consider the online appendix of Gächter, Mengel, Tsakas, and Vostroknutov, 2017, for a formal proof). A simple backwards induction argument also renders zero contributions in each round as the unique equilibrium of the sequential game *LEAD* (hence, I abstain from providing a formal proof here). Therefore, zero contributions, and thus, no difference in final payoffs should emerge in both treatments assuming only selfish players.

However, in the presence of a substantial proportion of subjects whose patterns of behavior can be described as reciprocal types that exhibit conditional cooperation (Keser and Van Winden, 2000; Fischbacher, Gächter, and Fehr, 2001), this does not necessarily hold anymore. Ambrus and Pathak (2011) set up such a model and adopt it to a static public goods game. Their main idea is that in an environment where both selfish and reciprocal types exist, the selfish types can induce reciprocal types to choose non-zero contributions. As conditional cooperators are backward-looking in determining their contributions, selfish players can influence future contributions of reciprocal players positively by contributing a large amount. The more rounds that are left, the higher the influence of these contributions. It is therefore in the interest of selfish players to contribute, especially in early rounds of a period,

[47]In part I of the experiment, payoffs were given in EUR. For part II, I used Taler as experimental currency unit (ECU) with an exchange rate of 1 Euro = 10 Taler.

as this causes a positive reaction of the reciprocal players affecting the remaining periods. In the course of a period, this incentive for selfish players diminishes, which leads to a decreasing pattern of their contributions. The intuition behind this mechanism is the assumption of selfish players who maximize their material payoffs and of reciprocal players whose payoffs additionally depend on a concave reciprocity function. The function itself is non-decreasing in other players' contributions and it is embedded in the payoff function of a reciprocal player. This payoff function is maximal when a specific target contribution is reached - which depends on the other players' contributions and the reciprocity function. Additionally, there is the assumption of no overreciprocation, which means that one unit of contribution by any player does increase the value of the reciprocity function by not more than 1. This in turn, leads to a decreasing marginal impact of contributions over the course of the game.[48] This driving force in the model of Ambrus and Pathak (2011) is also reasonable in my dynamic setting with endowment carryover. In addition, higher contributions by selfish players have a positive impact on contribution capabilities of the reciprocal types, thus, constituting a second incentive for them to contribute in early rounds.

The sources of a possible positive impact of leadership are twofold. *First*, due to their exposition, the initial contribution of a leader offers a chance for amelioration: it gives clues for everyone else about the distribution of types and the degree of optimism the leader has over the occurrence of cooperation. Reciprocal types are no longer solely dependent on their expectations about the contributions of others, instead, the leader's first contribution is setting a salient example. The leader acts as a 'belief manager' (Gächter and Renner, 2018). As early contributions determine later contribution capabilities, the (positive) signal that the leader can give here might have a big impact. *Second* and likewise, the leader can give a positive signal inducing conditional cooperators to reciprocate in other periods by his particular visibility. This is even true for the very last round of a given period: a selfish leader has incentives to make a positive contribution in the last period, since reciprocal types may respond to his contribution. In the simultaneous game this is not the case, where selfish types should choose a contribution of zero in the last round. Taken together, this gives rise to my first hypothesis:

**Hypothesis 1.** *Assuming a substantial fraction of conditional cooperators, I expect non-negative contributions in both treatments. In addition, I expect that contributions in LEAD are larger than in NOLEAD, resulting in higher final earnings.*

[48]Consider the appendix of Ambrus and Pathak (2011) for a formal proof.

### 4.3.2 Inequality

As described in Section 4.1, followers typically contribute less than leaders. Followers adopt an imperfect matching strategy and donate systematically less compared to the leaders, displaying signs of imperfect or selfish-biased conditional cooperation. In my dynamic setting, this effect would intensify round by round, so that the leader would be impoverished relative to the followers. Compared to *NOLEAD*, larger inequality would be the result. This leads to the following hypothesis:

**Hypothesis 2a.** *LEAD leads to* ***higher*** *average within group inequality compared to NOLEAD.*

An alternative view would insist that due to the leaders influence (Gächter and Renner, 2018), the followers may have a more homogeneous contribution pattern than subjects in *NOLEAD*. Despite the discrepancy between a leader's and his followers' contributions, this would give rise to reduced inequality compared to *NOLEAD* which consequently yields the following alternative hypothesis:

**Hypothesis 2b.** *LEAD leads to* ***lower*** *average within group inequality compared to NOLEAD.*

So in summary, I have two competing hypotheses. Accordingly, it is undecided to me a priori whether leadership (*LEAD*) leads to an increase or decrease in inequality within a group compared to *NOLEAD*.

### 4.3.3 Leader types

Similar to the one-shot-game scenario of Gächter, Nosenzo, Renner, and Sefton (2012), I also expect cooperative types to be the better leaders with respect to average final earnings of a group. Due to a kind of "false consensus effect" and their social preferences, I expect them to make higher initial contributions. In addition, I especially expect conditional cooperators to promote cooperation by giving the group the right signals for cooperation. In the spirit of the model of Ambrus and Pathak (2011), I expect strategic leadership of selfish types with a declining pattern of their contributions.

**Hypothesis 3.** *Within LEAD I expect that groups work best when led by cooperatively inclined individuals.*

## 4.4 Results

In this section, I present the results of my experiment. First, I focus on the effect of leadership by comparing both treatments. This mainly involves answering the question of whether leadership has an impact on the wealth of group members and inequality within groups. For each of the two questions, I also consider the influence of the behavioral type of the leader.

Furthermore, I want to know whether leadership reduces the number of individuals that are likely to find themselves in a sucker position, not benefiting from any cooperation gains. For this, I make use of a slightly modified version of the *payoff dominance* criterion, which is used in game theory. According to my definition, we have *payoff dominance* within a group and round as long as *all* group members *strictly* improve their earnings at the end of a round compared to the previous round, in absolute terms.[49] More formally: there is *payoff dominance* at the end of round $t-1$, as long as $E_i^t > E_i^{t-1}$ holds for all $i \in \{1, 2, 3, 4\}$. A violation of this pattern naturally occurs, when some members of the group contribute relatively few money to the group pot, while others spend more (free-rider behavior), or when all group members stop contributing money to the group pot. In my analysis, I find that leadership leads to a delayed occurrence of the first violation of the payoff dominance condition.

In the last part of the analysis, I focus on treatment *LEAD* only, to get a deeper insight about the underlying mechanisms of successful leadership. My aim is to find out whether the followers' behavior is also dependent on their behavioral types, with regard to their reaction on a leader's contribution. Furthermore, it is my intention to take a closer look at different leader types and to show why some of them are more successful than others. In particular, I investigate whether I can observe the predicted pattern of decreasing relative contributions of selfish leaders, and thus strategic leadership behavior. I conclude with illustrative examples of successful and less successful leadership.

Overall, my results suggest that leadership is an important driver for the success of a group. Especially the behavior of the leader in the first round(s) is crucial for the further course of the group. Additionally, I find that the cooperation type of the leader matters a lot: groups led by cooperatively inclined types have, on average, a significantly greater final endowment. The leader himself, however, is, on average, not better off than the participants in the simultaneous treatment *NOLEAD*. So it is mainly the followers who profit from leadership.

[49] In this sense, my definition of *payoff dominance* can also be seen as a stricter version of the Pareto criterion, since I require a *strict* improvement of *all* players.

### 4.4.1 Wealth

I start with the question of whether group members are better off in the end, when they are in a group with a leader. Comparing final earnings of participants at the end of round 7, I find a significant treatment effect ($p$=0.025, clustered two-sided two-sample $t$-test). Participants in the sequential treatment *LEAD* earn, on average, about 89 Taler compared to 65 Taler average earnings in *NOLEAD* at the end of round 7 - which is an increase of around 37%. Figure 4.2 (*left*) depicts average endowments at the end of each round by treatment.

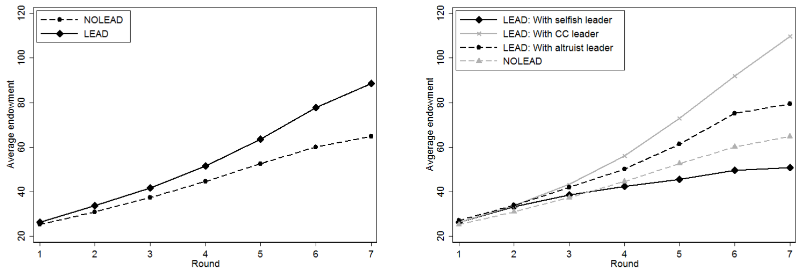



Figure 4.2: Average endowment at the end of each round

Table 4.3 reports random-effects models, where I estimated the treatment effect of *LEAD* on final earnings in different specifications (period controls are included in all regression models). In this regression table, I used the information from part I of the experiment, regarding the cooperation type of the leader, and I also investigated the effect of a violation of the payoff dominance criterion already after round 1. It becomes apparent that the leader's cooperation type is essential (see also Figure 4.2 (*right*)). Leadership with a conditional cooperator as first-mover improves earnings a lot (more than an altruist) while a selfish leader worsens the outcome (although not significantly).[50] I will discuss this in more detail in Section 4.4.4.

If there are no mutual benefits of cooperation in round 1, this has persistent effects on final outcomes (see column (3) of Table 4.3). Subjects in groups, without payoff dominance after round 1, suffer from large and negative effects on their final earnings. The coordinating role of leadership is also reflected in the fact that the presence of a leader delays the first violation of the payoff dominance condition, as I will show in the further part of the analysis (see Section 4.4.3). This is an important aspect which can explain why a leader can be valuable for the group.

[50] I used a Wald test for comparing the coefficients.

| | Endowment (end of round 7) | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| *LEAD* | 23.847** | -10.780 | -33.224*** |
| | (10.511) | (8.342) | (6.712) |
| *LEAD* X CC leader | | 54.339*** | 57.315*** |
| | | (13.505) | (12.562) |
| *LEAD* X AL leader | | 24.365** | 23.365* |
| | | (11.929) | (13.159) |
| No PAYOFF DOMINANCE after R1 | | | -65.028*** |
| | | | (6.485) |
| Constant | 60.439*** | 58.545*** | 103.191*** |
| | (8.297) | (8.406) | (8.063) |
| Period controls | YES | YES | YES |
| Observations | 368 | 368 | 368 |
| Subjects | 92 | 92 | 92 |
| $R^2$ overall | 0.024 | 0.075 | 0.217 |

*Note:* Random-effects regression with period controls. Cluster-robust standard errors (on the subject-level) are in parentheses. Panel variable: *subject*; time variable: *period*. In columns (2) and (3) the variable "*LEAD*" can be interpreted as *LEAD* treatment times a *selfish leader*.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.3: Regression Table - Final endowment

I have now investigated whether the presence of a leader has benefits for the final wealth of the group members. In addition, we saw how the leader can already shape the game in round 1. My results can be summarized as follows:

**Observation 1.** *My findings suggest that leadership has a positive impact on final wealth. In addition, the behavioral type of the leader has a major impact. The members of groups led by a conditional cooperator are best off, on average. Furthermore, if there are no mutual benefits of cooperation in round 1, this has persistent negative effects on final outcomes.*

### 4.4.2 Inequality

A question that naturally arises in connection with a dynamic public goods game with endowment carryover is that of inequality. In my setting, inequality can endogenously arise through different contributions of the group members to the public good.

According to Hypothesis 2a, leaders suffer some kind of curse. Previous research, concerning static public goods games, has shown that leaders earn, on average, less than followers (see, for example, Güth, Levati, Sutter, and Van Der Heijden, 2007). They are, so to speak, exploited by the followers. This pattern should lead to rising inequality. However, as reflected in Hypothesis 2b, previous research also states that leaders have a great influence on the expectations and contributions of the followers (see, for example, Gächter and Renner, 2018). Thus, if a leader has a big impact on the followers, it can be expected that the contributions to the public good will become more even. More even followers' contributions would result in less within-group inequality. In the further course of this chapter, I will show that it is mainly the second effect which predominates here.

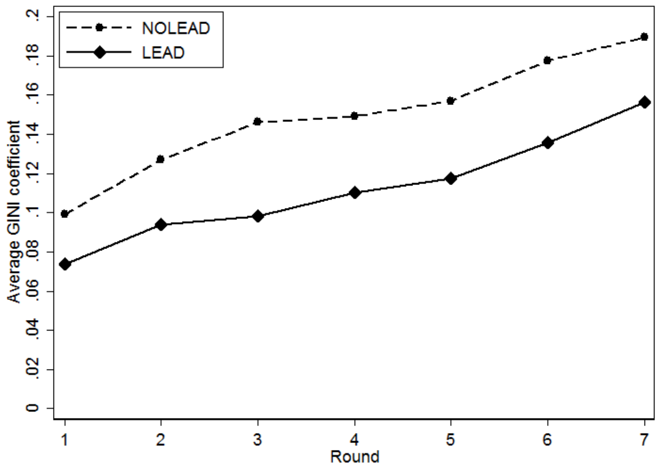



Figure 4.3: Average Gini coefficient at the end of each round by treatment

A common measurement for inequality is the Gini index that I compute for every group and round (a value of 0 refers to complete *equality*, whereas a value of 1 refers to complete *inequality*). I used a random-effects regression where a uniquely identifiable group serves as panel variable and the round serves as time variable (see Table 4.4). Figure 4.3 further illustrates my findings graphically, by plotting the average Gini coefficient per round, split up by treatment. We see that the average inequality in *LEAD* is lower compared to *NOLEAD* and, in general, I observe that inequality rises with each round in both treatments of the dynamic public goods game.

| | GINI coefficient | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| *LEAD* | -0.037*** | 0.003 | 0.011 |
| | (0.013) | (0.018) | (0.018) |
| *LEAD* X CC leader | | -0.058*** | -0.058*** |
| | | (0.020) | (0.019) |
| *LEAD* X AL leader | | -0.041 | -0.040 |
| | | (0.026) | (0.026) |
| No PAYOFF DOMINANCE after R1 | | | 0.028** |
| | | | (0.013) |
| Constant | 0.099*** | 0.100*** | 0.081*** |
| | (0.013) | (0.013) | (0.016) |
| Period controls | YES | YES | YES |
| Round controls | YES | YES | YES |
| Observations | 644 | 644 | 644 |
| Groups | 92 | 92 | 92 |
| $R^2$ overall | 0.193 | 0.246 | 0.277 |

*Note:* Random-effects regression with period and round controls. Cluster-robust standard errors (on the group-level) are in parentheses. Panel variable: *group*; time variable: *round*. In columns (2) and (3) the variable "*LEAD*" can be interpreted as *LEAD* treatment times a *selfish leader*.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.4: Regression Table - GINI coefficient

As shown by column (1) in Table 4.4, I can report a significant negative treatment effect with respect to within group inequality. Again the leader's type does matter a lot. A conditional cooperator as a leader has a significant effect in reducing inequality as measured by the Gini index, while the effect for other leader types is insignificant (see column (2) in Table 4.4). I also see that groups that achieve no payoff dominance in round 1, have a significantly higher inequality. I summarize as follows:

**Observation 2.** *Leadership has a positive impact on reducing inequality within groups as measured by the Gini index. This effect is mainly driven by the conditional cooperators. If there are no mutual benefits of cooperation in round 1, this increases inequality significantly.*

### 4.4.3 Payoff dominance accross rounds

As we have seen in the previous parts, having at least one group member who does not strictly improve his or her endowment already in round 1, has a persistent effect, resulting in lower final earnings for *all* group members. In this section, I show that leadership can be a useful tool in preventing an early violation of the payoff dominance condition.

As Figure 4.4 indicates, I find evidence that leadership plays a pivotal role in establishing stable cooperation for multiple rounds. In more than 60% of the cases at least one individual does not strictly improve in *NOLEAD* after the first round. By contrast, in *LEAD* only in around 30% of the cases the criterion for no payoff dominance is met after the first round, highlighting the leader's role in coordinating others' contributions. On average, at least one subject within a group does not strictly improve for the first time after 2.75 rounds in *NOLEAD*, compared to 3.75 rounds in *LEAD* ($p$=0.0289, two-sided two-sample $t$-test). The sequential treatment therefore delays the first violation, on average, by one round.

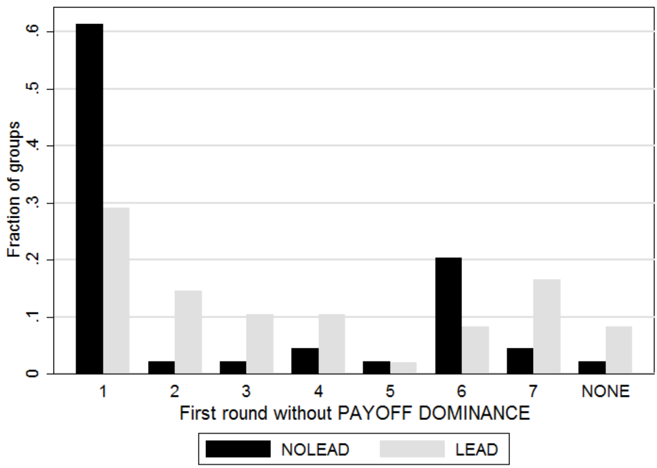



Figure 4.4: First emergence of no PAYOFF DOMINANCE

We can further see that in treatment *NOLEAD*, there is a second peak in round 6, whereas in *LEAD* this second peak occurs only in round 7. This indicates that typical last round effects, driven by free-rider behavior, occur earlier in *NOLEAD* compared to *LEAD*.

In Table 4.5 we can see that the sequential treatment *LEAD* has a significant effect on deferring the first appearance of a situation where at least one subject of the group has no benefits from cooperation (the dependent variable is the round in which at least one group member does not strictly improve for the first time). Controlling for the type of the leader (column (2) of Table 4.5), I find that it is the conditional cooperators that have a positive impact when selected to be leaders.

| | Round of first violation of PD | |
|---|---|---|
| | (1) | (2) |
| *LEAD* | 1.000*** | 0.186 |
| | (0.351) | (0.384) |
| *LEAD* X CC leader | | 1.243*** |
| | | (0.388) |
| *LEAD* X AL leader | | 0.689* |
| | | (0.405) |
| Constant | 2.609*** | 2.571*** |
| | (0.361) | (0.360) |
| Period controls | YES | YES |
| Observations | 368 | 368 |
| Subjects | 92 | 92 |
| $R^2$ overall | 0.055 | 0.088 |

*Note:* Random-effects regression with period and round controls. Cluster-robust standard errors (on the subject-level) are in parentheses. Panel variable: *subject*; time variable: *period*. In column (2) the variable "*LEAD*" can be interpreted as *LEAD* treatment times a *selfish leader*. PD is an abbreviation for payoff dominance.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.5: Regression Table - Violation of payoff dominance

**Observation 3.** *Leadership can be a useful tool in preventing an early violation of the payoff dominance condition. The sequential treatment LEAD delays the first occurrence of the event, that at least one individual does not strictly improve, on average, by one round. The behavioral type of the leader is again crucial in this context.*

### 4.4.4 Further analysis of the *LEAD* treatment

In this section, I want to take a closer look behind the mechanisms of effective leadership and investigate why some groups in *LEAD* are more successful than others. I begin by analyzing the followers' reaction and then proceed by looking at the behavior of the leaders.

**Follower behavior**

The followers' reaction is decisive for successful leading-by-example. Only if they respond adequately, it will be ensured that leadership (i) has success in terms of the final endowments for the group members, and additionally, (ii) that the leader does not fare worse either.

| | Contribution | |
|---|---|---|
| | (1) | (2) |
| Leader contribution (LC) | 0.822*** | |
| | (0.056) | |
| LC X SF type | | 0.560*** |
| | | (0.095) |
| LC X CC type | | 0.837*** |
| | | (0.056) |
| LC X AL type | | 0.909*** |
| | | (0.085) |
| Constant | 1.526 | 1.830* |
| | (1.235) | (1.087) |
| Period controls | YES | YES |
| Round controls | YES | YES |
| Observations | 1008 | 1008 |
| Subjects | 47 | 47 |
| $R^2$ | 0.781 | 0.791 |

*Note:* OLS regression with period controls. Cluster-robust standard errors (on the subject-level) are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

| | Relative Contribution | |
|---|---|---|
| | (1) | (2) |
| Rel. leader contribution (RLC) | 0.835*** | |
| | (0.027) | |
| RLC X SF type | | 0.639*** |
| | | (0.101) |
| RLC X CC type | | 0.853*** |
| | | (0.024) |
| RLC X AL type | | 0.869*** |
| | | (0.080) |
| Constant | 0.021 | 0.033 |
| | (0.037) | (0.036) |
| Period controls | YES | YES |
| Round controls | YES | YES |
| Observations | 1008 | 1008 |
| Subjects | 47 | 47 |
| $R^2$ | 0.701 | 0.711 |

*Note:* OLS regression with period controls. Cluster-robust standard errors (on the subject-level) are in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.6: Regression Table - Matching of leader contribution (absolute and relative)

I consider the reactions to a leader's contribution both relative to a subject's own endowment and in absolute terms. Focusing on absolute contributions first, Table 4.6 (*left*) indicates that for every Taler, a leader gives in a round, followers give on average 0.822 Taler to the public good. Additionally, follower type heterogeneity leads to different reactions. Although all types react positively to higher contributions of the leader, this reaction is particularly pronounced for followers that were classified as conditional cooperators and altruists, as the coefficients in Table 4.6 reveal. Selfish follower types match a Taler given by the leader by only 0.56 Taler, which is much less compared to the amount given by conditional cooperators and altruists. A similar pattern emerges for relative contributions (see Table 4.6 (*right*)).

My result is summarized in Observation 4. This high rate of matching a leader's contribution further illustrates why the behavior of the leader matters a lot, since even selfish types react positively on the amount spent by the leader.

**Observation 4.** *In general, followers react positively to a leader's contribution, matching it to a large degree, both in absolute and in relative terms. The type of the follower plays a major role as selfish followers exhibit a much smaller reaction.*

**Leader behavior**

I have already mentioned two results on leadership. First, I find evidence that it is important that they set a good example - best already in the beginning. Additionally, the behavioral type of the leader also matters, as Figure 4.2 (*right*) illustrates. In the next step, I will go more into detail to explain the mechanisms behind successful leadership.

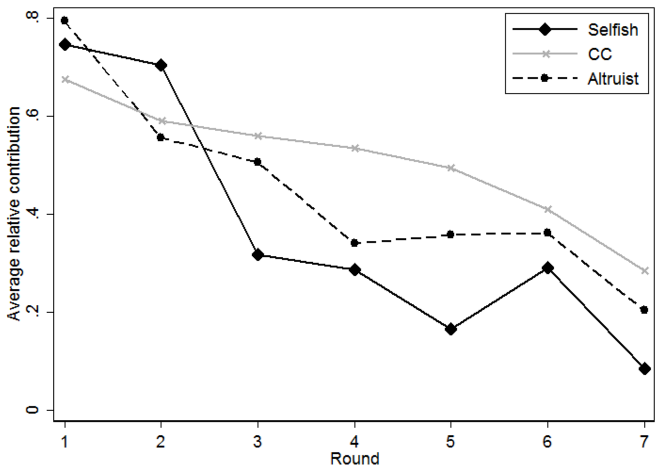



Figure 4.5: Leader's relative contribution by behavioral type and round

We have seen that conditional cooperators prosper as leaders with respect to their group members' final earnings. A rather naïve guess that they - driven by some kind of false consensus effect - contribute a lot to the group pot at first, does not meet the mechanism. As Figure 4.5 shows, conditional cooperators have the lowest average initial leader contribution of all types. Rather, as Figure 4.5 also shows, it is their persistence that makes CC types successful.

A behavior that can probably be explained by a false consensus effect, is the one shown by those subjects which I classified as altruists. When they are leaders, they start with very high relative contributions in the beginning, but display a faster decline than the CC types afterwards. By contrast, I discover signs of strategic leadership by selfish types. They start with relatively high contributions in the first two rounds, revising them sharply afterwards. What came to my mind was a strategy the New York Times (Gleick, 1986) described as a 'tranquilizer strategy' for the repeated prisoner's dilemma: to lull the opponent for a few moves and then try to exploit him.

I find evidence for this claim in my data as selfish types contribute more in the beginning when they are leaders. In round 1 of *LEAD*, 61.54% of the leaders, which I classified as "selfish", contributed their whole endowment of 20 Taler. However, only 21.74% of the selfish followers also contributed this amount. In treatment *NOLEAD*, by comparison, 25.00% of the selfish types contributed 20 Taler in round 1.

| | Endowment (end of round 7) | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Is leader | -26.540** | -25.703** | 24.404 |
| | (12.829) | (12.884) | (27.012) |
| Leader's first contribution | | 7.566*** | 8.351*** |
| | | (1.332) | (1.433) |
| Leader's first contr. X Is Leader | | | -3.567 |
| | | | (2.539) |
| Constant | 80.073*** | -15.345 | -26.520* |
| | (11.066) | (13.489) | (15.160) |
| Period controls | YES | YES | YES |
| Observations | 192 | 192 | 192 |
| Subjects | 48 | 48 | 48 |
| $R^2$ overall | 0.029 | 0.194 | 0.196 |

*Note:* Random-effects regression with period controls. Cluster-robust standard errors (on the subject-level) are in parentheses. Panel variable: *subject*; time variable: *period*.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.7: Regression Table - Final endowment (only *LEAD*)

Overall, a high early contribution by the leader in round 1 of a period has a large impact on final earnings in round 7 of the respective period. In columns (2) to (3) of Table 4.7 we can see that a leader's first contribution has a large effect on final wealth. E.g., in column (2) we see that for each Taler the leader gives in round 1, the final endowment of every group member (including the leader himself) increases by around 7.6 Taler, on average.

The leader himself in turn, does not lose money by contributing more in round 1, but he profits less from each Taler compared to the other group members as it can be seen by the interaction effect in column (3). However, as shown by Figure 4.6 (*left*), the relationship between first leader's contributions and final wealth is not necessarily linear, since I only observe distinct peaks for contributions of 15 and 20 Taler. Additionally, in the right part of Figure 4.6, we can observe that there is a large heterogeneity between leader types for the 20 Taler bracket. This provides further evidence that the success of CC leaders is not driven by their initial contributions, but rather by their persistence.

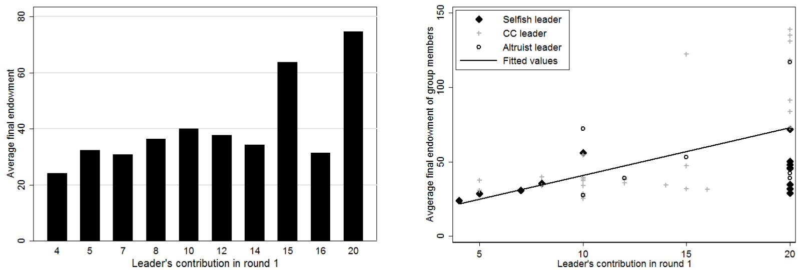



Figure 4.6: Relationship between leader's first contribution and final endowment

I have already reported that group members are more successful when they have a conditional cooperator as leader. Yet, being the leader is not necessarily good for the own payoffs as column (1) of Table 4.7 reveals. Leaders face some kind of curse as their earnings are lower than those of other group members. This result is not very surprising and can be explained by incomplete or selfish-biased conditional cooperation by followers. Figure 4.7 illustrates this circumstance graphically. However, leaders do not fare worse than average participants in *NOLEAD* in the end. This, too, is depicted in Figure 4.7.

**Observation 5.** *The leader type has a significant effect on final earnings of the group members. The results suggest that leaders classified as conditional cooperators are successful because they give persistently a larger fraction of their endowment over the course of a period. High first contributions by the leader, yield a high return concerning final earnings. The leader herself is slightly better off than an average subject in NOLEAD, but not significantly.*

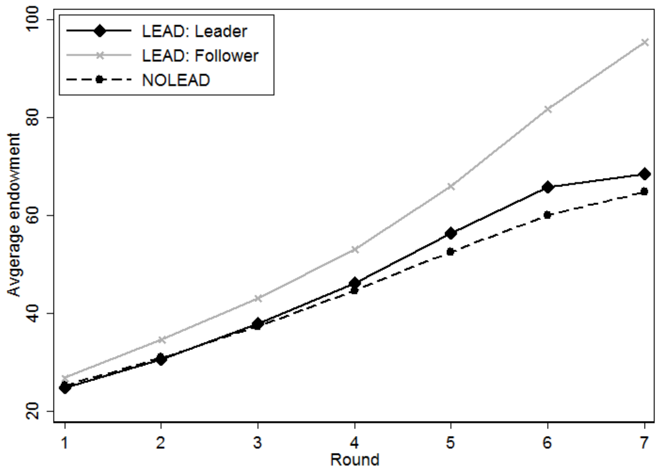



Figure 4.7: Average endowment by treatment + leader/non-leader

**Examples of good and bad leadership**

To illustrate the effect of good and bad leadership, I present some exemplary groups in this section. The black line in the graphs represents the leader contributions, while the grey lines represent the follower contributions.

Figure 4.8 shows two examples of bad leadership. In the left part we see that the leader stops contributing after three rounds, although the followers matched him before (at least partly). This results in a breakdown of cooperation and almost no further growth after round 3. In the right part we can observe a leader who started with a medium contribution in the first round. After this, the leader decreases his contribution in every subsequent round and the followers mimic this behavior, resulting in a low final average endowment.

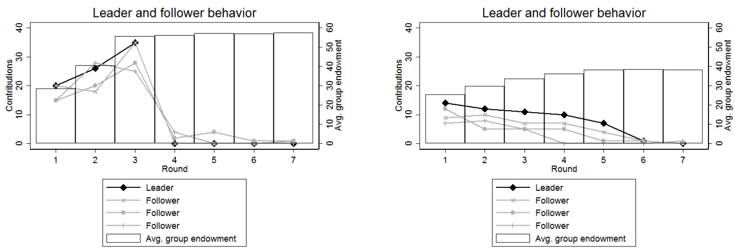



Figure 4.8: Bad leadership

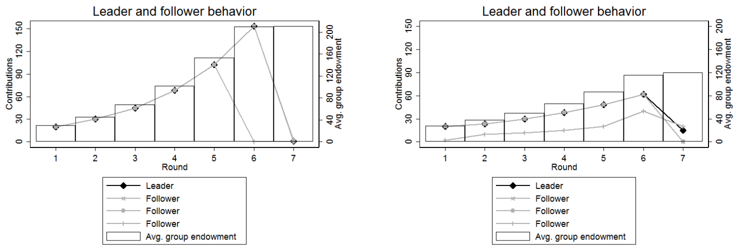



Figure 4.9: Good leadership

Figure 4.9 shows two examples of good leadership. In both cases the leader starts with a full contribution of 20 Taler in the beginning. In the left part we can see that the leader only deviates in the last round from full contribution. The same is true in the graph on the right part, where the leader sticks to his plan of full contribution until round 6, although one group member is a free-rider who only matches the leader partly.

## 4.5 Discussion

In this chapter, I provided an experiment investigating the effect of leadership in a dynamic public goods game with endowment carryover and, hence, endogenous growth. In summary, the analysis of my experiment shows that leadership has a positive impact on final wealth of the participants. The results of static public goods games continue to hold and, consequently, also in my setting, exogenous leadership yields an improvement of group members' earnings. The leaders, however, benefit less. They contribute more to the group pot than the followers, but receive lower payouts than them. However, the average leader in *LEAD* is not worse off than an average player in *NOLEAD*. From a welfare perspective, it can be argued that *leading-by-example* leads to a Pareto improvement, when we consider the average earnings.

I also find that the contribution of the leader in the first round, has a high impact on the final results across all groups. For the groups, it pays off if the leader prefaces by setting a good example in the form of a large initial contribution. The leader himself, however, profits less from being very cooperative in the beginning.

In addition, leadership has other positive effects. I observe a significantly lower inequality in groups with a leader compared to *NOLEAD*. As a measurement for inequality, I used the Gini index which also refers to a utilitarian social welfare function that integrates individual inequity aversion (Schmidt and Wichardt, 2018). From this perspective, I can report a further welfare improvement through leadership which, of course, is not generally valid. Linked to this is my result that leadership delays the point of time when at least one individual no longer benefits from cooperation.

Based on a sequential prisoner's dilemma, I elicited types for conditional cooperation in part I of the experiment. This classification of types allows me to go into further detail. Regarding the question of who is a good leader in such a dynamic game, the results obtained indicate that it is especially the conditional cooperators who stand out. The mechanism is quite interesting: it is apparently less a false consensus effect that makes conditional cooperator types successful, with regard to group members' final earnings, but rather their perseverance in setting a good example. While I see signs of strategic leadership in the selfish types' behavior, with very high contributions in the first two rounds followed by a sharp crash, conditionally cooperative types change their contribution relative to income only gradually. They are not the ones who give the highest contributions in round 1, on average. Nevertheless, they have endurance and contribute a relatively high proportion of their income for a long time. The subjects classified as altruists are those for whom leadership by a false consensus effect could best be used as an argument. They make very high contributions at the beginning, but then quickly - speculatively because of disappointment - reduce their contributions. Regarding the reaction of the followers, I can report that followers who are either classified as conditional cooperators or altruists, match a leader's contribution to a higher degree than selfish types.

As a concluding remark, I can state that *leading-by-example* seems to be a very useful tool in public good environments. I find that it has a positive effect on both final wealth across groups and reducing inequality within groups. In a dynamic public goods game setting with endowment carryover, I can further show that the benefits of leadership especially come into play in the long run.

In the next chapter, I will take a closer look at different methods for classifying cooperation types. My aim here is to investigate whether two structurally different procedures are consistent.

# 5 Consistency of cooperation types[51]

One of the main contributions of behavioral economics is to establish the behavioral relevance of another type beyond the purely payoff-maximizing "homo oeconomicus", named "homo reciprocans", who represents a large fraction of the population.[52] If a researcher needs to determine behavioral types of subjects in the lab, there are essentially two methods available to him. On the one hand, he can use the method introduced by Fischbacher, Gächter, and Fehr (2001) which relies on a conditional contribution vector elicited by the strategy method in a one-shot public goods game (*FGF* hereafter).[53] This method is typically based on a set of 22 questions.[54] On the other hand, a simple sequential prisoner's dilemma (*SPD* hereafter), for which only three questions are sufficient, can be used for type classification as well (Miettinen, Kosfeld, Fehr, and Weibull, 2017; Kosfeld, 2019; Eichenseer and Moser, 2019b). For a researcher, the question arises whether using the simpler method is sufficient for type classification as it may save time and reduce cognitive load for the participants. To the best of my knowledge, there exists no systematic comparison of classification congruence between these two procedures.

Consequently, the aim of this paper is to assess the stability of classifications across games thereby contributing to the literature on the within subject stability of cooperation preferences (Blanco, Engelmann, and Normann, 2011; Volk, Thöni, and Ruigrok, 2012). To this end, I compare the types assigned by *SPD* to those assigned by *FGF* in its latest refinements (Fallucchi, Luccasen, and Turocy, 2018; Thöni and Volk, 2018). The remainder of this paper will be as follows: Section 5.1 describes the experimental design and procedures. Section 5.2 presents and discusses my results. Section 5.3 provides as short summary and concludes.

## 5.1 Design and procedures

### 5.1.1 Protocol

The experiment was conducted on Amazon Mechanical Turk (MTurk henceforth) in December 2018 using a sample of MTurk experienced US residents. In total, 232 participants took part in the experiment earning $2.85 on average with an average completion time of approximately 13 minutes. About half of the subjects

[51] This chapter is a slightly modified version of Eichenseer and Moser (2019a).

[52] See, for example, Fehr and Gächter (2000), Dohmen, Falk, Huffman, and Sunde (2009), and Kosfeld (2019).

[53] This method is by now the most commonly used one and, for example, labeled as 'P-Experiment' in Fischbacher and Gächter (2010).

[54] As a second-mover, subjects are typically asked to specify their contribution conditional on the other players' average contribution for integers in the interval $[0, 20]$. This results in 21 questions plus an unconditional contribution question for the role as first-mover.

(120) played *SPD* first, while the other half (112) was doing the *FGF* task first. Subsequently, the participants completed a short questionnaire on age, gender, and education. Instructions for the experiment can be found in Appendix A.

### 5.1.2 Sequential Prisoner's Dilemma (SPD)

In the *SPD* we have two players, indexed by $i = 1, 2$. Each player can choose between actions *SEND* (S) and *KEEP* (K). Choices are elicited by using the strategy method such that Player 2 can condition his choice on the action of Player 1. Figure 5.1 depicts the structure of the game in extensive form including the resulting final payoffs in POINTS (worth \$0.05 each). The social optimum is reached when Player 1 chooses $S$ and Player 2 responds with action $S$ as well. However, maximizing their own payoffs means that Player 2 will choose action $K$ at both decision nodes and Player 1, who anticipates this behavior, chooses $K$ at the beginning. This is the unique subgame-perfect equilibrium of this game. Hence, the decision situation resembles a sequential prisoner's dilemma.

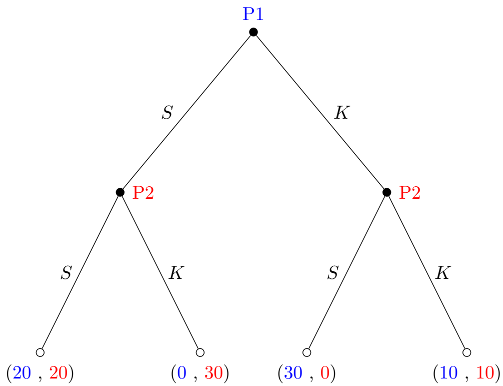



Figure 5.1: Payoff structure of the sequential prisoner's dilemma

All subjects state decisions for both being Player 1 and 2 (strategy method). They are randomly allocated to one of these roles at the end of the experiment and paid accordingly. The set of strategies, $X_i$, in this game for Player 2 is given by $X_i = \{SS, KK, SK, KS\}$.[55] Based on the participants' conditional second mover's choices, I can classify subjects as *altruists* (unconditional cooperators), *conditional cooperators* (cooperate only if the first-mover cooperates), *free-riders* (never cooperate), and *mismatchers* (counteract the other player) as depicted in Table 5.1.

[55] The first action is played when Player 1 chooses *SEND* and the second action is played when Player 1 chooses *KEEP*.

| ***Cooperation type*** | ***Strategy*** |
|---|---|
| Conditional cooperator (CC) | $(SEND, KEEP)$ |
| Selfish (SF) | $(KEEP, KEEP)$ |
| Altruist (AL) | $(SEND, SEND)$ |
| Mismatcher (MM) | $(KEEP, SEND)$ |

Table 5.1: Cooperation types in *SPD*

### 5.1.3 Sequential Public Goods Game (FGF)

For the conditional contributions task in *FGF*, I used an adapted version of the procedure of Fischbacher, Gächter, and Fehr (2001). Four players, indexed by $i = 1, 2, 3, 4$, play a sequential public goods game in which one player makes his contribution after observing the other three players' rounded average contribution when they were moving simultaneously beforehand. The resulting payoff of player $i$ with initial endowment $y_i = 20$ POINTS is given by:

$$\pi_i = y_i - g_i + \alpha \sum_{j=1}^{4} g_j$$

where $g_i \in [0, 20]$ denotes individual contributions and $\alpha = 0.4$ is the marginal per capita return (MPCR) of the public good. Choices are elicited by using the strategy method such that every player $i$ makes a choice both for being one of the three first-movers (*unconditional contribution*) and being a second-mover (*contribution table*). As a second-mover, subjects condition their contribution $g_i$ on the average contribution (rounded to the next integer) of the first-movers which results in a conditional contribution path. Subjects are randomly assigned roles of first- and second-movers at the end of the experiment. For the type classification, only the *contribution table* of a subject is considered. The classification of Fischbacher, Gächter, and Fehr (2001) results in four types: a *conditional cooperator* whose contributions increase with other players' contributions, a *selfish type* who never cooperates, a *triangle cooperator* with hump-shaped contributions, and the *remaining subjects* who do not fit either one of the classifications.

Recently, there have been two proposals to refine the classification based on Fischbacher, Gächter, and Fehr (2001): (i) the method of Thöni and Volk (2018), which is based on the Pearson correlation coefficient and (ii) the method of Fallucchi, Luccasen, and Turocy (2018), which is based on hierarchical clustering. I will describe the behavioral types resulting from both refinements in Section 5.2.2. They have in common that they entail a behavioral type whose description comes close to the altruist in *SPD*: the unconditional cooperator (UC) in Thöni and Volk (2018) and the unconditional high type (UHC) in Fallucchi, Luccasen, and Turocy (2018).

## 5.2 Results

### 5.2.1 Contribution paths in *FGF* by *SPD* type

As a first step in my data analysis, I provide a visual inspection to see whether there is a systematic relationship between behavioral types in *SPD* and contribution paths in *FGF* which follow from the subjects' conditional contributions. Figure 5.2 depicts contribution paths in *FGF* by *SPD* type.[56]

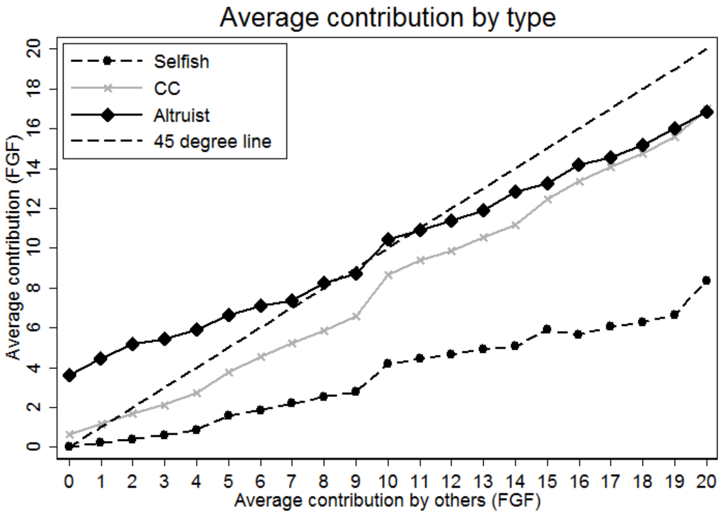



Figure 5.2: Contribution paths by *SPD* classification in *FGF*

There are considerable differences between types. Compared to subjects classified as 'selfish' in *SPD*, contributions of 'conditional cooperators' (CC) have a decisively steeper slope in the contributions of others, i.e., they match others' contributions to a larger degree. In addition, subjects classified as 'altruist' in *SPD* have the highest intercept which reflects that they give most when others give nothing. In line with expectations, 'selfish' types have, on average, the lowest conditional contributions for every level of average contributions of others.

In Table 5.2, I examine whether this visual interpretation can be supported statistically. Columns OLS(1) and Tobit(1) assume a common slope of all types in the *average contribution of others* (ACO) - and only different intercepts - whereas OLS(2) and Tobit(2) take different slopes for different *SPD* types into account. The Tobit regressions consider observations censored at 0 and 20. Both regressions

[56]I excluded the mismatcher type in this graph, since it is a rare empirical phenomenon (9 of 232 subjects) whose behavior is difficult to interpret.

OLS(2) and Tobit(2) indicate that conditional cooperators show a significantly larger reaction to others' contributions compared to the reference category of selfish types. This corresponds to the graphical findings reported in Figure 5.2. Moreover, the coefficient of the intercept - the unconditional contribution - is largest for the altruist type and significantly different from the reference category of selfish types in the regressions OLS(1), Tobit(1), and Tobit(2).

| | Contribution | | | |
|---|---|---|---|---|
| | OLS(1) | OLS(2) | Tobit(1) | Tobit(2) |
| Conditional cooperator | 4.573*** | 0.174 | 7.688*** | 3.394*** |
| | (0.499) | (0.350) | (1.003) | (1.021) |
| Altruist | 6.426*** | 3.781 | 10.299*** | 8.432** |
| | (1.798) | (2.410) | (2.796) | (3.639) |
| Mismatcher | 3.545*** | 2.339* | 6.970*** | 7.415*** |
| | (0.999) | (1.278) | (1.455) | (2.142) |
| Avg. contr. of others (ACO) | 0.665*** | 0.394*** | 0.970*** | 0.720*** |
| | (0.028) | (0.045) | (0.034) | (0.061) |
| Conditional cooperator X ACO | | 0.440*** | | 0.387*** |
| | | (0.055) | | (0.074) |
| Altruist X ACO | | 0.265* | | 0.148 |
| | | (0.142) | | (0.182) |
| Mismatcher X ACO | | 0.121 | | -0.075 |
| | | (0.144) | | (0.188) |
| Constant | -3.071*** | -0.358** | -10.007*** | -7.115*** |
| | (0.359) | (0.170) | (1.017) | (0.933) |
| Observations | 4872 | 4872 | 4872 | 4872 |
| Subjects | 232 | 232 | 232 | 232 |
| $R^2$ | 0.483 | 0.518 | | |
| Pseudo $R^2$ | | | 0.114 | 0.118 |

*Note:* Cluster-robust standard errors (on the subject-level) are in parentheses. Tobit regressions account for 1,646 left-censored and 346 right-censored observations. ACO stands for "average contributions of others". The 'selfish' type serves as a reference category.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5.2: Regression Table - Contribution paths

### 5.2.2 Relationship between classification methods

I now investigate the relationship between the discrete behavioral types classified by *SPD* and *FGF* in the refinements of Thöni and Volk (2018) and Fallucchi, Luccasen, and Turocy (2018). The refinement of Thöni and Volk (2018) of *FGF* (*FGF-**T*** hereafter) resembles a theory-driven approach and is based on the Pearson correlation coefficient. It distinguishes the five behavioral types depicted in Table 5.3.

In my sample, I can categorize 184 out of 232 subjects (79.3%) as conditional cooperators (CC) using the *FGF-**T*** refinement.[57] Conditional cooperators also constitute

[57] This is close to the 80.6% CC share reported in the US sample of Kocher, Cherry, Kroll,

| ***Type*** | ***Behavior*** |
|---|---|
| Free-rider (FR) | Zero contributions. |
| Conditional cooperator (CC) | Monotonically increasing pattern in others' contributions. |
| Unconditional cooperator (UC) | Constant contributions irrespective of what others do. |
| Triangle cooperator (TC) | "Hump-shaped" contributions. |
| Other | Undefined contribution pattern. |

Table 5.3: Cooperation types in Thöni and Volk (2018)

the largest group in *SPD* with a share of 57.8%. The second largest group are selfish types that account for 33.6% of all subjects in *SPD* and 13.8% in *FGF-**T***. In both games, these two categories cover the vast majority of subjects. Table 5.4 reports the number and percentage of subjects falling into each possible combination of the two methods in a contingency table.

*Behavioral type* $FGF - T$

| *Behavioral type* $SPD$ | | $FR$ | $CC$ | $UC$ | $TR$ | $Other$ | $Total$ |
|---|---|---|---|---|---|---|---|
| | $Selfish$ | **27 (11.64%)** | **44 (18.97%)** | 1 (0.43%) | 5 (2.16%) | 1 (0.43%) | 78 (33.62%) |
| | $CC$ | **4 (1.72%)** | **125 (53.88%)** | 2 (0.86%) | 2 (0.86%) | 1 (0.43%) | 134 (57.76%) |
| | $Altruist$ | 1 (0.43%) | 8 (3.45%) | 2 (0.86%) | 0 (0.00%) | 0 (0.00%) | 11 (4.74%) |
| | $Mismatcher$ | 0 (0.00%) | 7 (3.02%) | 2 (0.86%) | 0 (0.00%) | 0 (0.00%) | 9 (3.88%) |
| | $Total$ | 32 (13.79%) | 184 (79.31%) | 7 (3.02%) | 7 (3.02%) | 2 (0.86%) | 232 (100.00%) |

Table 5.4: Types in *SPD* and *FGF* (Refinement of Thöni and Volk, 2018)

Comparing the classification of *SPD* and *FGF-**T***, we see that slightly more than half of all subjects (125 of 232) are classified as CC according to both methods, while 11.6% are classified as selfish types in both games (27 of 232). Overall, only around 13.8% of the subjects (32 of 232) are classified in a category different from selfish or CC according to at least one of the methods. The results of a $\chi^2$-test suggests that the characteristics of both methods are not independent ($p < 0.001$). Hence, I can reject the null-hypothesis that there is no relationship between the two classification methods.

Conditional relative frequencies allows me to get a better picture of the type stability across games. About 93.3% of the subjects who are classified as CC in *SPD*, are also classified as CC according to *FGF-**T***. However, individuals classified as 'selfish' in *SPD*, are classified as 'selfish' according to *FGF-**T*** only in around 34.6% of the cases. This indicates that *SPD* performs well in identifying subjects who have a consistent pattern of conditional cooperation across games, while this

Netzer, and Sutter (2008).

does not hold for selfish types.

Conversely, starting from *FGF*, subjects classified as CC according to *FGF-**T***, are in around 67.9% of the cases also CC in *SPD*, and those who are classified as 'selfish' according to *FGF-**T*** are in around 84.4% of the cases also 'selfish' in *SPD*. This means that *FGF* is better suited to identify types who are classified as 'selfish' in both games compared to *SPD*.

| ***Type*** | ***Behavior*** |
|---|---|
| Own maximizers (OWN) | Zero contributions. |
| Strong conditional cooperators (SCC) | Match others' contributions exactly. |
| Weak conditional cooperators (WCC) | Increasing contributions, but less than one-for-one. |
| Unconditional high contributors (UCH) | Contribute fully irrespective of what others do. |
| Other | Undefined contribution pattern. |

Table 5.5: Cooperation types in Fallucchi et al. (2018)

These findings are robust when changing to the refinement of Fallucchi, Luccasen, and Turocy (2018), which is based on hierarchical clustering and resembles a data-driven approach (*FGF-**F*** hereafter). The *FGF-**F*** categorization splits the CC category and distinguishes between weak conditional cooperators (WCC) and strong conditional cooperators (SCC). The type classification of *FGF-**F*** is depicted in Table 5.5. In my experimental sample, there has not been a distinct cluster of 'Other' types and, hence, I only consider four behavioral types.

*Behavioral type $FGF - F$*

| *Behavioral type SPD* | *OWN* | *WCC* | *SCC* | *UCH* | *Total* |
|---|---|---|---|---|---|
| *Selfish* | **31 (13.86%)** | **37 (15.95%)** | **10 (4.31%)** | 0 (0.00%) | 78 (33.62%) |
| *CC* | **11 (4.74%)** | **45 (19.40%)** | **74 (31.90%)** | 4 (1.72%) | 134 (57.76%) |
| *Altruist* | 1 (0.43%) | 2 (0.86%) | 6 (2.59%) | 2 (0.86%) | 11 (4.74%) |
| *Mismatcher* | 0 (0.00%) | 6 (2.59%) | 3 (1.29%) | 0 (0.00%) | 9 (3.88%) |
| *Total* | 43 (18.53%) | 90 (38.79%) | 93 (40.09%) | 6 (2.59%) | 232 (100.00%) |

Table 5.6: Types in *SPD* and *FGF* (Refinement of Fallucchi et al., 2018)

Table 5.6 presents the contingency table of types. Again, a $\chi^2$-test shows that the type classifications are not independent ($p < 0.001$), indicating a significant relationship between the two methods. If we look at the conditional relative frequencies, we see that conditional on being classified as CC type in *SPD*, the relative frequency is 88.8% to be classified as either WCC or SCC according to *FGF-**F***. By contrast, a subject classified as selfish in *SPD* is only selfish in 39.7% of the cases according to *FGF-**F***.

Starting from *FGF*-***F***, a subject sorted in the group of selfish types according to *FGF*-***F***, is also selfish in *SPD* in 72.1% of the cases. By contrast, the relative frequency of being CC in *SPD* is only 65.0% when being classified as either WCC and SCC according to *FGF*-***F***. When distinguishing between WCC and SCC, I observe that in the group of those who are classified as WCC according to *FGF*-***F***, only 50% are also classified as CC in *SPD*, whereas in the group of those who are classified as SCC, almost 80% are classified as CC in *SPD*. Thus, the distinction between WCC and SCC predicts the relative frequency of being CC in *SPD* quite well. Likewise, the relative frequency of being selfish in *SPD* is highest for OWN maximizers, followed by WCC and SCC types.

Figures 5.3 and 5.4 illustrate the respective intersections between *SPD* and *FGF* - for CC and selfish types - graphically by using Venn diagrams. The solid circles represent the respective sets of CC and selfish types according to *SPD*, while the dashed circles represent these types according to the *FGF* classification. The intersection of both circles illustrates the set of subjects who are of the same type according to both methods. In Figure 5.4 (*left)*, the WCC and SCC types are pooled as conditional cooperators.

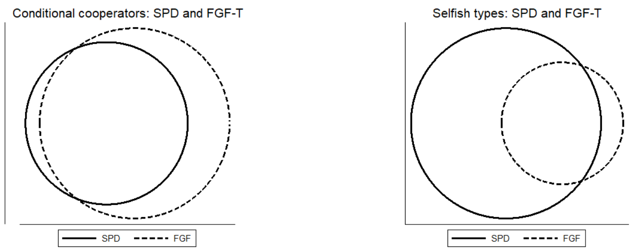



Figure 5.3: Venn diagrams of SPD and FGF-T (Thöni and Volk, 2018)

The fact that the overlap between selfish types in *SPD* and *FGF* is quite small leaves room for further research. One hypothesis would be that the *FGF* method underestimates the share of selfish types. Confused types, who do not understand the rules of the game completely, may act as if they were cooperative types in *FGF* (see Detemple, Kosfeld, and Kröll, 2019). Assuming that the *SPD* imposes fewer cognitive load on subjects would allow for the hypothesis that the share of confused types is lower in this game and, consequently, the share of selfish types should be higher in *SPD* compared to the *FGF* method. This might explain why many of the selfish types in *SPD* behave cooperatively in *FGF*.

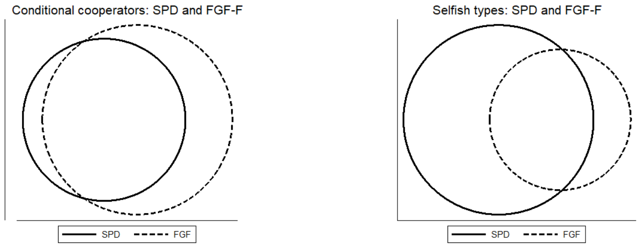



Figure 5.4: Venn diagrams of SPD and FGF-F (Fallucchi et al., 2018)

## 5.3 Summary and discussion

I provided an online experiment, in which I investigated the consistency of two methods for classifying different cooperation types. With regard to discrete behavioral types, my results indicate that *SPD* performs very well in identifying subjects with a stable pattern of conditional cooperation. Given that a subject is of CC type in *SPD*, the probability is 93.3% to be classified as CC as well according to *FGF-**T*** (refinement of Thöni and Volk, 2018) and 88.8% according to *FGF-**F*** (refinement of Fallucchi, Luccasen, and Turocy, 2018), respectively. I further observe that the distinction between WCC and SCC is helpful for identifying CC types in *SPD* more precisely, since the likelihood for being 'selfish' in *SPD* is considerably higher for WCC types compared to SCC types.

Considering contribution paths in *FGF*, subjects classified as conditional cooperators in *SPD* match others' contributions to a significantly larger degree compared to selfish types. This is captured in the significantly larger slope of their conditional cooperation path.

On this basis, I can conclude that if a researcher's objective is to identify those subjects in a group who are, with a high probability, conditional cooperators in both games, the simple method of the *SPD* is well suited for this task. If, on the other hand, the focus is on identifying selfish types, I cannot offer a clear conclusion. I observe many 'selfish' subjects in *SPD* who show cooperative behavioral patterns in *FGF*. However, based on the hypothesis that there is a larger share of confused types in *FGF*, who act as if they were CC types, the simpler game (*SPD*) is not necessarily a weak tool for identifying selfish types, but may be more accurate in measuring the true fraction of selfish types in the population (see Detemple, Kosfeld, and Kröll, 2019).

# 6 Conclusion

In this thesis, I presented four experiments in the field of behavioral economics to investigate the causes and consequences of non-standard decision-making in several economic and social environments. Two of the experiments were conducted in the lab and two of them were conducted as online experiments. The aim of my research was to take a closer look at the underlying channels of deviations from standard-economic behavior with a focus on cognitive biases.

In Section 2, I have shown that mistakes in hypothetical thinking often lead to suboptimally high or low bids in auctions. The overall pattern of my data suggests that the problem of irrational over- and underbidding can be weakened by giving the subjects ex ante feedback about their bid, but unlike related studies I also find negative effects of additional information. Hence, I showed that the effect of the provided information strongly depends on the context and, overall, I can conclude that through information subjects do not suddenly become more sophisticated, but rather they adapt their heuristics.

In Section 3, I demonstrated that the cognitive bias of correlation neglect is present in a voting environment where individuals received correlated signals. However, in a setup where groups can reach beneficial results through information aggregation, this mental bias can actually lead to an improvement of the voting results, because voters suffering from correlation neglect vote less often according to their ideological preferences and focus more on their available information instead. This in turn leads to the pattern that groups receiving two perfectly correlated signals instead of only one signal are, on average, more successful - even though they have no information advantage.

In Section 4, I have shown that groups in a public goods game, with a first-mover who contributes first, are substantially more successful compared to groups without a leader. This is expressed in both a higher income and a lower inequality. Additionally, I demonstrated that the type of the leader matters a lot, since groups lead by cooperatively inclined individuals generate significantly higher earnings.

In Section 5, I demonstrated that two well-known methods in the literature for eliciting cooperation types are consistent to a large extent. Furthermore, I have shown that each of the methods has its own strengths and weaknesses and that the choice of the proper method depends strongly on the underlying research question.

As a concluding remark there is to say, that there are many more research opportunities in the field of behavioral and experimental economics and a lot more work is needed to solve further urgent puzzles. The humble attempt of this thesis is to close some of the most important gaps in the literature. I hope that I have succeeded with my work.

# References

Adam, M. T., J. Krämer, and M. B. Müller (2015). Auction fever! How time pressure and social competition affect bidders' arousal and bids in retail auctions. *Journal of Retailing 91*(3), 468–485.

Adam, M. T., J. Krämer, and C. Weinhardt (2012). Excitement up! Price down! Measuring emotions in dutch auctions. *International Journal of Electronic Commerce 17*(2), 7–40.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics 84*(3), 488–500.

Alekseev, A., G. Charness, and U. Gneezy (2017). Experimental methods: When and why contextual instructions are important. *Journal of Economic Behavior & Organization 134*, 48–59.

Ambrus, A. and P. A. Pathak (2011). Cooperation over finite horizons: A theory and experiments. *Journal of Public Economics 95*(7-8), 500–512.

Arbak, E. and M.-C. Villeval (2013). Voluntary leadership: motivation and influence. *Social Choice and Welfare 40*(3), 635–662.

Ariely, D., A. Bracha, and S. Meier (2009). Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American Economic Review 99*(1), 544–55.

Astor, P. J., M. T. Adam, C. Jähnig, and S. Seifert (2013). The joy of winning and the frustration of losing: A psychophysiological analysis of emotions in first-price sealed-bid auctions. *Journal of Neuroscience, Psychology, and Economics 6*(1), 14–30.

Avery, C. and J. H. Kagel (1997). Second-price auctions with asymmetric payoffs: An experimental investigation. *Journal of Economics & Management Strategy 6*(3), 573–603.

Barrett, G. F. and S. G. Donald (2003). Consistent tests for stochastic dominance. *Econometrica 71*(1), 71–104.

Battaglini, M., S. Nunnari, and T. R. Palfrey (2016). The dynamic free rider problem: A laboratory study. *American Economic Journal: Microeconomics 8*(4), 268–308.

Bazerman, M. H. and W. F. Samuelson (1983). I won the auction but don't want the prize. *Journal of Conflict Resolution 27*(4), 618–634.

Blanco, M., D. Engelmann, and H. T. Normann (2011). A within-subject analysis of other-regarding preferences. *Games and Economic Behavior 72*(2), 321–338.

Cadigan, J., P. T. Wayland, P. Schmitt, and K. Swope (2011). An experimental dynamic public goods game with carryover. *Journal of Economic Behavior & Organization 80*(3), 523–531.

Camerer, C., S. Nunnari, and T. R. Palfrey (2016). Quantal response and nonequilibrium beliefs explain overbidding in maximum-value auctions. *Games and Economic Behavior 98*, 243–263.

Capen, E. C., R. V. Clapp, and W. M. Campbell (1971). Competitive bidding in high-risk situations. *Journal of Petroleum Technology 23*(6), 641–653.

Cappelen, A. W., B.-A. Reme, E. Ø. Sørensen, and B. Tungodden (2016). Leadership and Incentives. *Management Science 62*(7), 1944–1953.

Charness, G. and D. Levin (2009). The origin of the winner's curse: A laboratory study. *American Economic Journal: Microeconomics 1*(1), 207–236.

Cooper, D. J. and H. Fang (2008). Understanding overbidding in second price auctions: An experimental study. *The Economic Journal 118*(532), 1572–1595.

Costa-Gomes, M. A. and M. Shimoji (2015). A comment on "Can relaxation of beliefs rationalize the winner's curse?: An experimental study". *Econometrica 83*(1), 375–383.

Crawford, V. P. and N. Iriberri (2007). Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? *Econometrica 75*(6), 1721–1770.

Dannenberg, A. (2015). Leading by example versus leading by words in voluntary contribution experiments. *Social Choice and Welfare 44*(1), 71–85.

Detemple, J., M. Kosfeld, and M. Kröll (2019). Confusion does not explain conditional cooperation in public-goods experiments. *unpublished manuscript*.

Deversi, M., A. Ispano, and P. Schwardmann (2018). Spin doctors: A model and an experimental investigation of vague disclosure. *working paper*.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2009). Homo reciprocans: Survey evidence on behavioural outcomes. *The Economic Journal 119*(536), 592–612.

Eichenseer, M. (2019). Leading by example in public goods experiments: What do we know? *unpublished manuscript*.

Eichenseer, M. and J. Moser (2019a). Conditional cooperation: Type stability across games. *working paper*, Available at SSRN 3384821.

Eichenseer, M. and J. Moser (2019b). Leadership in dynamic public good provision: Endogenous growth and inequality. *working paper*, Available at SSRN 3269341.

Ellis, A. and M. Piccione (2017). Correlation misperception in choice. *American Economic Review 107*(4), 1264–92.

Enke, B. and F. Zimmermann (2017). Correlation neglect in belief formation. *Review of Economic Studies 86*(1), 313–332.

Esponda, I. and E. Vespa (2014). Hypothetical thinking and information extraction in the laboratory. *American Economic Journal: Microeconomics 6*(4), 180–202.

Esponda, I. and E. Vespa (2016). Contingent preferences and the sure-thing principle: Revisiting classic anomalies in the laboratory. *working paper*.

Evans, J. S. B. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement*, Volume 3. Psychology Press.

Eyster, E. and M. Rabin (2005). Cursed equilibrium. *Econometrica 73*(5), 1623–1672.

Eyster, E., M. Rabin, and G. Weizsäcker (2015). An experiment on social mislearning. *working paper*.

Eyster, E. and G. Weizsäcker (2016). Correlation neglect in portfolio choice: Lab evidence. *working paper*.

Falk, A., A. Becker, T. J. Dohmen, D. Huffman, and U. Sunde (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *working paper*.

Fallucchi, F., R. A. Luccasen, and T. L. Turocy (2018). Identifying discrete behavioural types: a re-analysis of public goods game contributions by hierarchical clustering. *Journal of the Economic Science Association*, 1–17.

Fehr, E. and S. Gächter (2000). Fairness and retaliation: The economics of reciprocity. *Journal of Economic Perspectives 14*(3), 159–181.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Fischbacher, U. and S. Gächter (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review 100*(1), 541–556.

Fischbacher, U., S. Gächter, and E. Fehr (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters 71*(3), 397–404.

Gächter, S., F. Mengel, E. Tsakas, and A. Vostroknutov (2017). Growth and inequality in public good provision. *Journal of Public Economics 150*, 1–13.

Gächter, S., D. Nosenzo, E. Renner, and M. Sefton (2012). Who makes a good leader? Cooperativeness, optimism, and leading-by-example. *Economic Inquiry 50*(4), 953–967.

Gächter, S. and E. Renner (2018). Leaders as role models and 'belief managers' in social dilemmas. *Journal of Economic Behavior & Organization 154*, 321–334.

Gleick, J. (1986). Prisoner's dilemma has unexpected applications. *The New York Times*.

Greiner, B. (2004). An online recruitment system for economic experiments. *Max Planck Institute of Economics, Strategic Interaction Group: Technical report*.

Gürerk, Ö., B. Irlenbusch, and B. Rockenbach (2009). Motivating teammates: The leader's choice between positive and negative incentives. *Journal of Economic Psychology 30*(4), 591–607.

Gürerk, Ö., T. Lauer, and M. Scheuermann (2018). Leadership with individual rewards and punishments. *Journal of Behavioral and Experimental Economics 74*, 57–69.

Güth, W., M. V. Levati, M. Sutter, and E. Van Der Heijden (2007). Leading by example with and without exclusion power in voluntary contribution experiments. *Journal of Public Economics 91*(5), 1023–1042.

Haigner, S. D. and F. Wakolbinger (2010). To lead or not to lead: Endogenous sequencing in public goods games. *Economics Letters 108*(1), 93–95.

Hermalin, B. E. (1998). Toward an economic theory of leadership: Leading by example. *American Economic Review*, 1188–1206.

Holt, C. A. and R. Sherman (1994). The loser's curse. *American Economic Review 84*(3), 642–652.

Hossain, T. and R. Okui (2018). Belief formation under signal correlation. *working paper*.

Ivanov, A., D. Levin, and M. Niederle (2010). Can relaxation of beliefs rationalize the winner's curse?: An experimental study. *Econometrica 78*(4), 1435–1452.

Jones, B. F. and B. A. Olken (2005). Do leaders matter? national leadership and growth since world war ii. *Quarterly Journal of Economics 120*(3), 835–864.

Kagel, J. H., R. M. Harstad, and D. Levin (1987). Information impact and allocation rules in auctions with affiliated private values: A laboratory study. *Econometrica 55*(6), 1275–1304.

Kagel, J. H. and D. Levin (1986). The winner's curse and public information in common value auctions. *American Economic Review 76*(5), 894–920.

Keser, C. and F. Van Winden (2000). Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics 102*(1), 23–39.

Klemperer, P. (1998). Auctions with almost common values: The wallet game and its applications. *European Economic Review 42*(3), 757–769.

Koch, C. and S. P. Penczynski (2018). The winner's curse: Conditional reasoning and belief formation. *Journal of Economic Theory 174*, 57–102.

Kocher, M. G., T. Cherry, S. Kroll, R. J. Netzer, and M. Sutter (2008). Conditional cooperation on three continents. *Economics Letters 101*(3), 175–178.

Kosfeld, M. (2019). The role of leaders in inducing and maintaining cooperation: The CC strategy. *The Leadership Quarterly*, forthcoming.

Levati, M. V., M. Sutter, and E. van der Heijden (2007). Leading by Example in a Public. *Journal of Conflict Resolution 51*(5), 793–818.

Levin, D., J. Peck, and A. Ivanov (2016). Separating bayesian updating from non-probabilistic reasoning: An experimental investigation. *American Economic Journal: Microeconomics 8*(2), 39–60.

Levy, G., I. Moreno de Barreda, and R. Razin (2018). Persuasion with correlation neglect: Media power via correlation of news content. *working paper*.

Levy, G. and R. Razin (2015a). Correlation neglect, voting behavior, and information aggregation. *American Economic Review 105*(4), 1634–1645.

Levy, G. and R. Razin (2015b). Does polarisation of opinions lead to polarisation of platforms? The case of correlation neglect. *Quarterly Journal of Political Science 10*(3), 321–355.

Li, S. (2017). Obviously strategy-proof mechanisms. *American Economic Review 107*(11), 3257–87.

Louis, P. (2013). The barrel of apples game: Contingent thinking, learning from observed actions, and strategic heterogeneity. *working paper*.

Madrian, B. C. and D. F. Shea (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *Quarterly Journal of Economics 116*(4), 1149–1187.

McCannon, B. C. (2018). Leadership and motivation for public goods contributions. *Scottish Journal of Political Economy 65*(1), 68–96.

Mellström, C. and M. Johannesson (2008). Crowding out in blood donation: was Titmuss right? *Journal of the European Economic Association 6*(4), 845–863.

Miettinen, T., M. Kosfeld, E. Fehr, and J. W. Weibull (2017). Revealed preferences in a sequential prisoners' dilemma: A horse-race between five utility functions. *working paper*.

Morgan, J., K. Steiglitz, and G. Reis (2003). The spite motive and equilibrium behavior in auctions. *Contributions in Economic Analysis & Policy 2*(1).

Moser, J. (2019). Hypothetical thinking and the winner's curse: An experimental investigation. *Theory and Decision 87*(1), 17–56.

Moser, J. and N. Wallmeier (2019). Correlation neglect in voting decisions: An experiment. *working paper*, Available at SSRN 3386981.

Moxnes, E. and E. Van der Heijden (2003). The effect of leadership in a public bad experiment. *Journal of Conflict Resolution 47*(6), 773–795.

Neugebauer, T., J. Perote, U. Schmidt, and M. Loos (2009). Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments. *Journal of Economic Psychology 30*(1), 52–60.

Ngangoué, K. and G. Weizsäcker (2015). Learning from unrealized versus realized prices. *working paper*.

Nickerson, R. (2015). *Conditional Reasoning: The Unruly Syntactics, Semantics, Thematics, and Pragmatics of "if"*. Oxford University Press.

Noussair, C. and C. Soo (2008). Voluntary contributions to a dynamic public good: Experimental evidence. *Economics Letters 98*(1), 71–77.

Ortoleva, P. and E. Snowberg (2015). Overconfidence in political behavior. *American Economic Review 105*(2), 504–535.

Pogrebna, G., D. H. Krantz, C. Schade, and C. Keser (2011). Words versus actions as a means to influence cooperation in social dilemma situations. *Theory and Decision 71*(4), 473–502.

Potters, J., M. Sefton, and L. Vesterlund (2005). After you—endogenous sequencing in voluntary contribution games. *Journal of Public Economics 89*(8), 1399–1419.

Rabin, M. (2002). A perspective on psychology and economics. *European economic review 46*(4-5), 657–685.

Rivas, M. F. and M. Sutter (2011). The benefits of voluntary leadership in experimental public goods games. *Economics Letters 112*(2), 176–178.

Rockenbach, B. and I. Wolff (2017). The effects of punishment in dynamic public-good games. *working paper*.

Sadrieh, A. and H. A. Verbon (2006). Inequality, cooperation, and growth: An experimental study. *European Economic Review 50*(5), 1197–1222.

Sahin, S. G., C. Eckel, and M. Komai (2015). An experimental study of leadership institutions in collective action games. *Journal of the Economic Science Association 1*(1), 100–113.

Schmidt, U. and P. C. Wichardt (2018). Inequity aversion, welfare measurement and the gini index. *working paper*.

Singmann, H., K. C. Klauer, and S. Beller (2016). Probabilistic conditional reasoning: Disentangling form and content with the dual-source model. *Cognitive Psychology 88*, 61–87.

Sutter, M. and F. Rivas (2014). Leadership, reward, and punishment in sequential public goods experiments. *Reward and punishment in social dilemmas*, 133–160.

Thaler, R. H. (1988). Anomalies: The winner's curse. *The Journal of Economic Perspectives 2*(1), 191–202.

Thöni, C. and S. Volk (2018). Conditional cooperation: Review and refinement. *Economics Letters 171*, 37–40.

Van den Bos, W., J. Li, T. Lau, E. Maskin, J. D. Cohen, P. R. Montague, and S. M. McClure (2008). The value of victory: Social origins of the winner's curse in common value auctions. *Judgment and Decision Making 3*(7), 483–492.

Volk, S., C. Thöni, and W. Ruigrok (2012). Temporal stability and psychological foundations of cooperation preferences. *Journal of Economic Behavior & Organization 81*(2), 664–676.

# A Figures & Tables (Section 2 - winner's curse)

## Decision screens


Figure A.1: Typical decision screen in stage I.

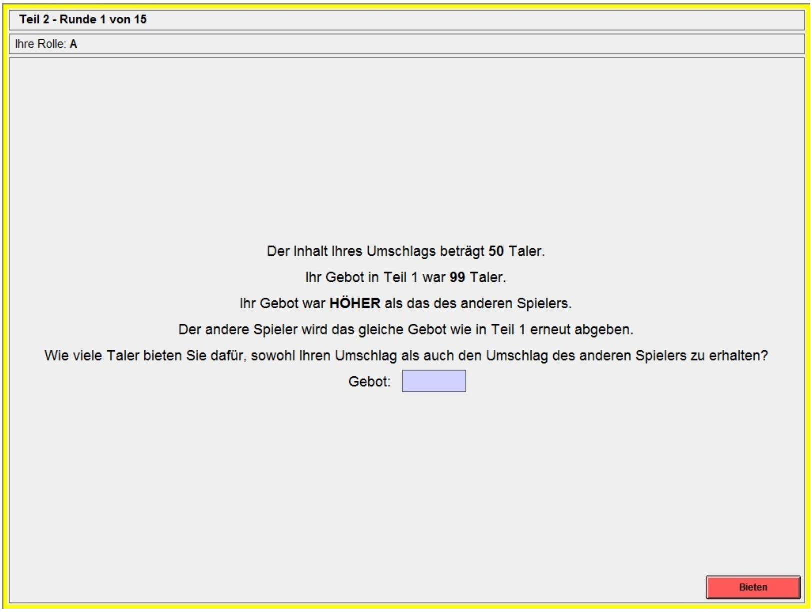



Figure A.2: Typical decision screen in stage II (treatment INF).

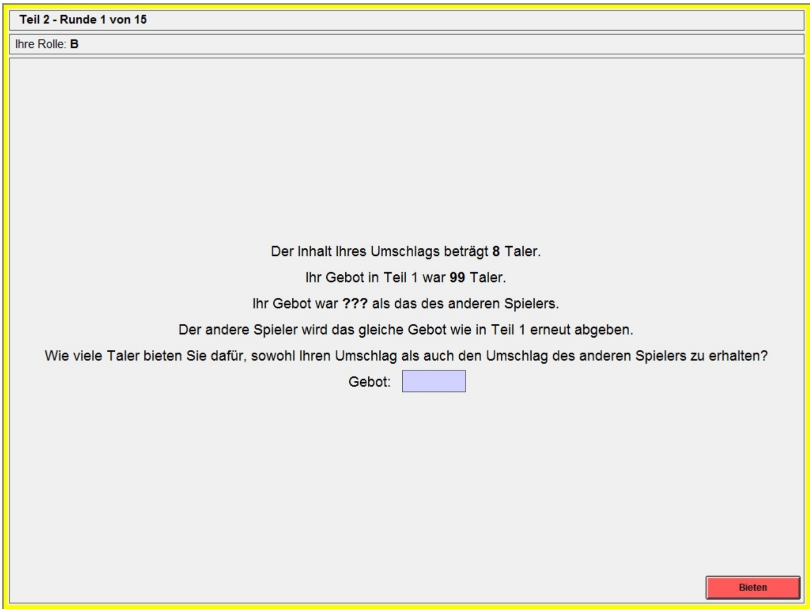



Figure A.3: Typical decision screen in stage II (treatment NOINF).

## Distribution of bids

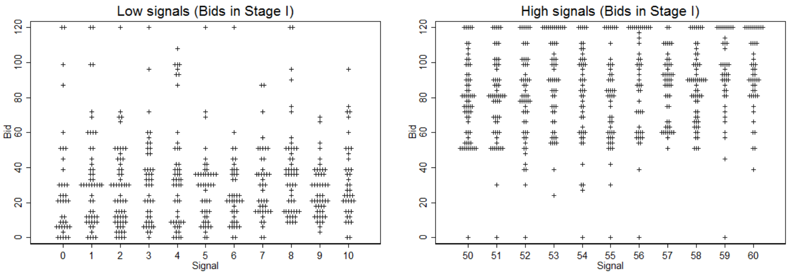



Figure A.4: Distribution of bids in stage I

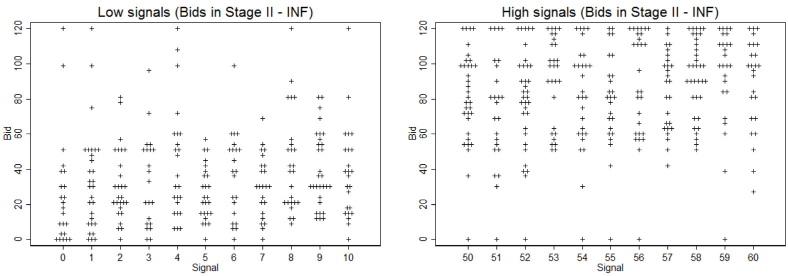



Figure A.5: Distribution of bids in stage II (INF)

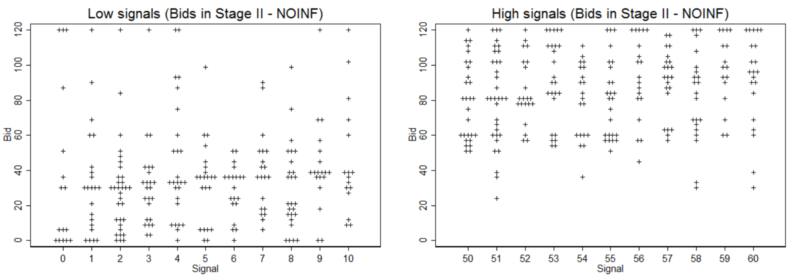



Figure A.6: Distribution of bids in stage II (NOINF)

## Illustration of Nash range

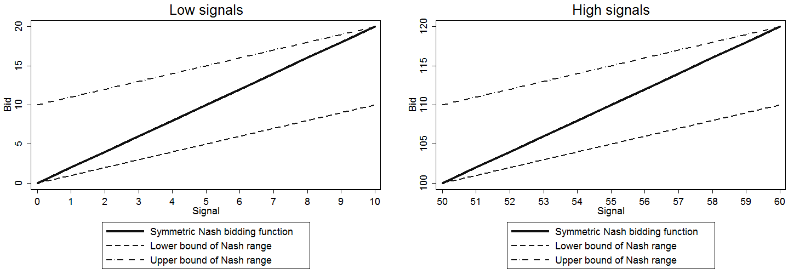



Figure A.7: Corridor of Nash range

## Robustness checks

| | Bid is in Nash range in stage I (YES [1] or NO [0]) | | | | | | |
|---|---|---|---|---|---|---|---|
| | $ls^H$ | $ls^L$ | $hs^H$ | $hs^L$ | $H$ | $L$ | *Overall* |
| INF | 0.002 | 0.081 | -0.023 | -0.053** | -0.010 | 0.039 | 0.021 |
| | (0.059) | (0.100) | (0.057) | (0.022) | (0.042) | (0.071) | (0.043) |
| Constant | 0.063 | 0.311*** | 0.153*** | 0.053** | 0.125*** | 0.229*** | 0.176*** |
| | (0.050) | (0.082) | (0.048) | (0.022) | (0.033) | (0.057) | (0.033) |
| Observations | 140 | 376 | 391 | 173 | 531 | 549 | 1080 |
| Subjects | 49 | 70 | 70 | 50 | 70 | 70 | 72 |
| $R^2$ | 0.000 | 0.007 | 0.001 | 0.030 | 0.000 | 0.002 | 0.001 |

Notes: Cluster-robust standard errors (on the subject-level) are in parentheses. Constellation $ls^H$ stands for low signals paired with the information HIGHER. Constellation $L$ stands for all signals paired with the information LOWER. The other constellations are defined analogously.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1: Regression Table - Bid in Nash range (Stage I)

| | Changes across the bidding ranges | | | | | | |
|---|---|---|---|---|---|---|---|
| | $ls^H$ | $ls^L$ | $hs^H$ | $hs^L$ | $H$ | $L$ | *Overall* |
| INF | 0.359*** | -0.222*** | -0.023 | 0.124 | 0.057 | -0.114* | -0.033 |
| | (0.096) | (0.077) | (0.039) | (0.102) | (0.036) | (0.068) | (0.039) |
| Constant | 0.051* | -0.018 | -0.000 | 0.092 | 0.016 | 0.017 | 0.016 |
| | (0.029) | (0.056) | (0.029) | (0.063) | (0.022) | (0.045) | (0.023) |
| Observations | 140 | 376 | 391 | 173 | 531 | 549 | 1080 |
| Subjects | 49 | 70 | 70 | 50 | 70 | 70 | 72 |
| $R^2$ | 0.193 | 0.054 | 0.002 | 0.026 | 0.008 | 0.014 | 0.002 |

Notes: Cluster-robust standard errors (on the subject-level) are in parentheses. Constellation $ls^H$ stands for low signals paired with the information HIGHER. Constellation $L$ stands for all signals paired with the information LOWER. The other constellations are defined analogously. The dependent variable indicates whether a bid was shifted from one range to the other across the stages (shifted from Nash to unsophisticated range ($-1$); not shifted from one range to the other (0); shifted from unsophisticated to Nash range (1)).

$^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table A.2: Regression Table - Changes across the bidding ranges

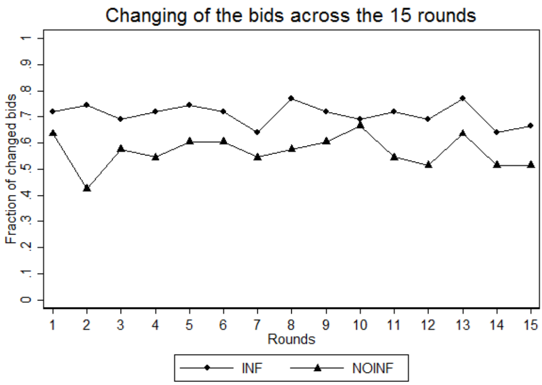



Figure A.8: Fraction of changed bids across the 15 rounds

# B Proofs (Section 2 - winner's curse)

*Proof of Proposition 2.* Bidding $b_1(x_1) = \alpha \cdot x_1$ and $b_2(x_2) = \frac{\alpha}{\alpha-1} \cdot x_2$ are equilibrium strategies for any $\alpha > 1$.

Consider two players, indexed by $i = 1, 2$. Suppose, without loss of generality, player 1 follows the bidding rule $b_1(x_1) = \alpha \cdot x_1$. Player 2 knows that the price he has to pay in the winning case is equal to $p = \alpha \cdot x_1$. So winning is beneficial for him as long as $x_1 + x_2 \geq p \Leftrightarrow x_2 + \frac{p}{\alpha} \geq p \Leftrightarrow x_2 \geq p \cdot \frac{\alpha-1}{\alpha} \Leftrightarrow p \leq \frac{\alpha}{\alpha-1} \cdot x_2$. Since we have a second-price auction, player 2 will bid exactly $b_2(x_2) = \frac{\alpha}{\alpha-1} \cdot x_2$.

Now, suppose player 2 follows the bidding rule $b_2(x_2) = \frac{\alpha}{\alpha-1} \cdot x_2$. Player 1 knows that the price he has to pay in the winning case is equal to $p = \frac{\alpha}{\alpha-1} \cdot x_2$. So winning is beneficial for him as long as $x_1 + x_2 \geq p \Leftrightarrow x_1 + p \cdot \frac{\alpha-1}{\alpha} \geq p \Leftrightarrow x_1 \geq \frac{p}{\alpha} \Leftrightarrow p \leq \alpha \cdot x_1$. Since we have a second-price auction, player 1 will bid exactly $b_1(x_1) = \alpha \cdot x_1$.

Hence, bidding $b_1(x_1) = \alpha \cdot x_1$ and $b_2(x_2) = \frac{\alpha}{\alpha-1} \cdot x_2$ are equilibrium strategies for any $\alpha > 1$. For $\alpha = 2$ we have the unique symmetric equilibrium. Note that this argumentation does not rely on the distribution of the signals $x_1$ and $x_2$. □

*Proof of Proposition 3.* Any bid $b_i(x_i)$ outside the interval $[x_i, x_i + 60]$ is weakly dominated. Consider two players, indexed by $i = 1, 2$.

(i) Bidding below $x_i$ is weakly dominated by bidding $x_i$. If player 1 would have lost the auction with $b_1(x_1) = x_1$, deviating to a bid below $x_1$ ($b'_1(x_1) < b_1(x_1)$) would not change the result because $b'_1(x_1) < b_1(x_1) \leq b_2(x_2)$. If player 1 would have won the auction with $b_1(x_1) = x_1$, he receives a payoff of at least 0, because the value of the object is at least $x_1$ and $b_2(x_2) \leq b_1(x_1) = x_1$. Deviating to a bid below $x_1$ can lead to losing an auction that generated a positive payoff. This is the case if player 1 bids $x_1 - \delta$ and player 2 bids $x_1 - \epsilon$ with $\delta > \epsilon \geq 0$. Thus, player 1 gives up an auction that guarantees a payoff of at least $x_1 + x_2 - x_1 + \epsilon = x_2 + \epsilon \geq 0$. If $\epsilon > \delta > 0$, player 1 still wins when he deviates, but the payoff does not change, because he still receives $x_1 + x_2 - x_1 + \epsilon = x_2 + \epsilon$ as before.

(ii) Bidding above $x_i + 60$ is weakly dominated by bidding $x_i + 60$. If player 1 would have won the auction with $b_1(x_1) = x_1 + 60$, he receives either a positive, negative or zero payoff. Deviating to a bid above $x_1 + 60$ (($b'_1(x_1) > b_1(x_1)$) would not change the result because $b'_1(x_1) > b_1(x_1) \geq b_2(x_2)$ and the price to pay remains $b_2(x_2)$. If player 1 would have lost the auction with $b_1(x_1) = x_1 + 60$, he receives a payoff of 0. Deviating to a bid above $x_1 + 60$ can lead to winning the auction, but the payoff is at best 0, because the value of the object is at most equal to $x_1 + 60$ and the price to pay remains $b_2(x_2)$ with $b_2(x_2) \geq b_1(x_1) = x_1 + 60$. So if player 1 deviates to $x_1 + 60 + \delta$ and player 2 bids $x_1 + 60 + \epsilon$ with $\delta > \epsilon \geq 0$, player 1 faces a loss of at least $\epsilon \geq 0$. □

*Proof of Proposition 4.* Suppose two players bid according to $b_1(x_1) = \alpha \cdot x_1$ and $b_2(x_2) = \frac{\alpha}{\alpha-1} \cdot x_2$. For any $\alpha \neq 2$ one player has a slope strictly greater than 2. For $\alpha > 2$ the proof is trivial. For $\alpha < 2$ we have the following equivalence: $\frac{\alpha}{\alpha-1} > 2 \Leftrightarrow \alpha > 2 \cdot (\alpha - 1) \Leftrightarrow \alpha > 2 \cdot \alpha - 2 \Leftrightarrow 2 > \alpha$.

Now suppose, without loss of generality, that player 1 uses the bidding function $b_1(x_1) = \alpha \cdot x_1$ with $\alpha = 2 + \epsilon$ for $\epsilon > 0$. For $x_1 = 60$ the bid is then given by $b_1(x_1) = \alpha \cdot 60 = 2 \cdot 60 + \epsilon \cdot 60 = 120 + \epsilon \cdot 60$. It follows that $120 + \epsilon \cdot 60 > x_1 + 60 = 120$ for all $\epsilon > 0$. By definition of Proposition 3, this bid is weakly dominated. So we have found at least one signal for which the resulting bid is weakly dominated for all $\epsilon > 0$.

Alternatively we can also show that for each $\alpha > 2$, we can find a $x_1 \in [50, 60]$ for which $b_1(x_1) = \alpha \cdot x_1$ is weakly dominated. We have to show that $\alpha \cdot x_1 > x_1 + 60$ holds for at least some $x_1$, for an arbitrary $\alpha > 2$. This inequation holds if $x_1 > \frac{60}{\alpha-1} := \gamma$. For $\alpha > 2$ we have that $\gamma < 60$. Hence, we can always find a $x_1 \in [50, 60]$ for which $x_1 > \gamma$ holds. □

*Proof of Proposition 5.* If a player bids according to $b_i(x_i) = x_i + 30$ in stage I and the other player does not use any weakly dominated bids, he can predict in stage II whether the signal of the other player is low or high when (i) he wins with a low signal or (ii) he loses with a high signal.

Consider two players, indexed by $i = 1, 2$. Suppose, without loss of generality, player 1 follows the bidding rule $b_1(x_1) = x_1 + 30$ and player 2 does not use any weakly dominated bids. This means for each $x_2$, player 2 may choose any bid in a range of $[x_2, x_2 + 60]$.

(i) Player 1 has a low signal ($x_1 \in \{0, \ldots, 10\}$) and wins: Player 1 wins with a bid of $b_1(x_1) = x_1 + 30$. This means his bid can be in a range of $[30, 40]$. If player 2 has a low signal, his bid can be in a range of $[0, 70]$. If player 2 has a high signal, his bid can be in a range of $[50, 120]$. So player 1 wins only if player 2 has a low signal.

(ii) Player 1 has a high signal ($x_1 \in \{50, \ldots, 60\}$) and loses: Player 1 loses with a bid of $b_1(x_1) = x_1 + 30$. This means his bid can be in a range of $[80, 90]$. If player 2 has a low signal, his bid can be in a range of $[0, 70]$. If player 2 has a high signal, his bid can be in a range of $[50, 120]$. So player 1 loses only if player 2 has a high signal. □

# C Instructions (Section 2 - winner's curse)

We would like to welcome you to this economic experiment! During the experiment you have the possibility to conduct a task that is explained in detail in the following instructions. In the experiment you can win a non-negligible amount of money. The amount of your payoff depends on your decisions, on the other participants' decisions and on chance. During the experiment it is forbidden to communicate with the other participants. Please read through the instructions at hand thoroughly. Should you have questions before or during the experiment, please raise your hand and an experimenter will come to your seat.

**General Structure**

The experiment consists of **two** parts. First, **part 1** will be explained. After part 1 ends, you will receive separate instructions for part 2. Your decisions in part 1 do not influence your payoff in part 2. During the whole experiment you can earn *Taler*. These will be converted into *Euros* after the experiment. The conversion rate is

**10 Taler = 1 EURO**

At the beginning of the experiment you are endowed with **50 Taler**. Your experimental credit at the end of the experiment consists of these 50 Taler plus your profits and minus your losses in part 1 and part 2. If you lose more than 50 Taler in the course of the experiment, your experimental credit drops down to 0 Taler. At the end of the experiment you receive your experimental credit in **EUR**. Anyway, independent of your decisions in the course of the experiment, you will receive **4 EUR** show-up fee at the end of the experiment. Your final payout will be calculated as follows:

**Final Payout = 4 EUR + experimental credit from part 1 and part 2 (in EUR)**

**Important remark**

All numerical examples that are used in the instructions, and later on in the control questions for exemplification, consist of arbitrary values and do not give a hint for optimal behavior in this experiment!

## Part 1

**Basic idea**

This experiment's underlying task is the following:

- Together with **one other player** you participate in an auction
- **You and the other player** receive each a randomly selected sealed **envelope that contains money**
- There are **red** and **blue** envelopes
- A **red** envelope contains a random integer amount between **0 and 10 Taler** (all values are equally likely)
- A **blue** envelope contains a random integer amount between **50 and 60 Taler** (all values are equally likely)
- Both colors are **equally likely**
- Thus, all together, there are 22 different amounts an envelope can contain and every amount is **equally likely**
- The colors and the amounts in both envelopes are **independent of one another** and it is also possible that both players receive the same amount

**The following combinations are possible:**

| | |
|---|---|
| **Player 1** *Envelope (0-10 Taler)* | **Player 2** *Envelope (0-10 Taler)* |
| **Player 1** *Envelope (0-10 Taler)* | **Player 2** *Envelope (50-60 Taler)* |
| **Player 1** *Envelope (50-60 Taler)* | **Player 2** *Envelope (0-10 Taler)* |
| **Player 1** *Envelope (50-60 Taler)* | **Player 2** *Envelope (50-60 Taler)* |

Both players are allowed to open their own envelope. This implies that every player learns his own amount, but not the other player's amount and color.
Then, both players participate in an auction, in which the highest bidder can win **both envelopes** and the money the envelopes contain.
Both players can submit a bid once only. The highest bidder wins **both envelopes** and pays the bid of the inferior bidder. The inferior bidder does not receive an envelope and does not have to pay anything - thus, he does neither make profit nor losses.

**The rules in detail**

**Bidding**

You and the other player can submit once only an *integer* bid between **0 and 120 Taler**. These bids are made *secretly*, i.e., the other player does not see what bid you have made and vice versa.

**Winning and Losing**

The winner and the payoff are determined as follows:
You win if:

1. Your bid is *higher* than the other player's bid
2. Your bid equals the other player's bid and your envelope contains *more* money

If your bids and the contents of the envelopes are equal, both players receive a payoff of 0. If your bid is lower than the other player's bid, you lose and receive a payoff of 0.

**Payoff**

If you have won the auction, your payoff is calculated as follows:
You receive the money of **both envelopes** and you have to pay the **other player's bid**. Thus, you do *not pay your own bid*, but the bid of the inferior bidder. This implies that in the winning case you must pay *at most* your own bid.
Is the amount of both envelopes higher than the bid you have to pay, you make profit. Is the amount of both envelopes lower than the bid you have to pay, you make a loss. If you have lost the auction, you receive a payoff of 0 - thus, you do neither make profit nor losses.

**Example**

Assume you would have 50 Taler in your envelope and the other player would have 10 Taler in his envelope (every player only knows his own amount). You bid 90 Taler and the other player bids 45 Taler (every player only knows his own bid). You win the auction, because you have submitted the higher bid. So you win both envelopes and pay the other player's bid. Your profit in this round would be $60-45 = 15$ Taler. The other player's profit would be 0 Taler.

**The course of the experiment**

**Trial phase**

At first you will bid in 5 *trial rounds* for the envelopes. During these 5 rounds you do not play against another human player, but against a computer. These rounds are **not relevant for your payoff** and their only purpose is to gain an understanding of the game and its general course. In every round you and the computer will receive a randomly selected amount between 0 and 10 or 50 and 60. This implies for you that you see in every round a randomly selected integer number between 0 and 10 or 50 and 60 on your screen. **This number symbolizes the content of an envelope** (you can find an example screenshot at the bottom of this page). You only learn your own amount, but not your computer opponent's amount and vice versa. Now, you can submit once (per round) any integer bid between 0 and 120 Taler. Overall, you participate in 5 auctions. After every round (i.e., after every bid) you receive an immediate feedback on your hypothetical profit or loss. However, you do not learn the opponent's bid or the amount the computer received. The computer is programmed to choose a bid that depends on his amount, i.e., the higher the computer's amount the higher the bid the computer chooses.

[SCREEN 1]

**Main phase**

After the trial phase you bid in 15 rounds for the envelopes and now you can receive an actual monetary payoff. In the main phase you do not compete with a computer but with a human player. Your opponent will be randomly drawn from this room. You do neither know who your opponent is nor does your opponent know this. The general course will be similar as in the trial phase. This means, in every round you and the other player will see a respective amount on your screens, which is a randomly selected number between 0 and 10 or between 50 and 60 that **symbolizes the content of an envelope**. In every round you and the other player can submit any integer bid between 0 and 120 Taler. But now, you do not receive an immediate feedback after each bid and you learn your final payoff only at the very end of the experiment (**i.e., only when part 2 is finished**). Out of the 15 rounds **3 randomly selected rounds are relevant for your payoff**. The other 12 rounds do not influence your payoff. Your profit or loss from this three randomly selected rounds is added to or subtracted from your initial endowment of 50 Taler, respectively.

**Example**

Assume round 1,2 and 3 are randomly selected for the payoff. In round 1 your payoff is 30 Taler, in round 2 your payoff is $-5$ Taler and in round 3 your payoff is 0 Taler. Your profit in part 1 would be 25 Taler. Your current experimental credit would be 50+25 = 75 Taler.

[SCREEN 2]

## Part 2

You will now bid for the envelopes once again. For this purpose you and your previous opponent will receive the **same amounts** as in part 1 again in the **same order**. In contrast to the previous part you do *not bid directly* against your opponent, but against **his decisions** that he made in part 1. This implies that your opponent does *not make new decisions* in part 2 and he will bid in every round exactly as in part 1. As you only compete with your opponent indirectly, your decisions in part 2 do not influence the payoff of your opponent in part 2.

In **part 2** you are randomly assigned to a role: either **A** or **B**. Both roles are equally likely. Your role is determined before the beginning of the 15 rounds by a virtual coin toss. Your respective role is constant for all 15 rounds and is displayed at the top of the screen.

If your role is **A**, you see for every amount additionally on the screen, whether your respective bid in part 1 was **HIGHER** or **LOWER** than the other player's bid. If the bids are equal, you will also see "**LOWER**" on your screen (thus, "HIGHER" means *strictly higher* and "LOWER" means *lower or equal*). If your role is **B**, you receive no further information in part 2 and instead of HIGHER or LOWER you only see "???" on your screen.

Independent of your role you are once again, in every round, allowed to submit a bid, which can be as in part 1 between 0 and 120 Taler. The rules for winning and losing are exactly as in part 1 and also your payoff is computed equally. As already in part 1, the same three randomly selected rounds determine your payoff (i.e., if round 1, 2 and 3 were selected in part 1, round 1, 2 and 3 also determine your payoff in part 2). Your profit or loss from part 2 is added to or subtracted from your current experimental credit. Also in part 2 you do not receive immediate feedback after every bid, but you learn your final payoff only at the very end of the experiment.

### Summarized

- Part 2 is a repetition of the main phase of part 1
- In part 2 you have the same opponent as already in part 1
- Now you do not compete with your opponent directly, but with his decisions he made in part 1
- Your opponent will bid in part 2 exactly as in part 1

- You are randomly assigned to either role A or B
- If your role is A, you additionally see whether your bid in part 1 was higher or lower than the other player's bid
- If your role is B, you do not receive additional information

[SCREEN 3]

[SCREEN 4]

**Example (role - A)**
Part 1 - 1st round: You have 50 Taler in your envelope and the other player has 10 Taler in his envelope (every player only knows his own amount). You bid 90 Taler and the other player bids 45 Taler (every player only knows his own bid).
Part 2 - 1st round: Now, you receive once again an envelope that contains 50 Taler and the other player again receives an envelope that contains 10 Taler (every player only knows his own amount). Now, you see on your screen that your bid in part 1 was **HIGHER** than the other player's bid (as 90 is larger than 45 - anyway, also in part 2 you do not know the other player's exact bid). Now, you can submit any bid between 0 and 120 Taler once again. The other player bids 45 Taler as in part 1.

**Example (role - B)**
Part 1 - 1st round: You have 50 Taler in your envelope and the other player has 10 Taler in his envelope (every player only knows his own amount). You bid 90 Taler and the other player bids 45 Taler (every player only knows his own bid).
Part 2 - 1st round: Now, you receive once again an envelope that contains 50 Taler and the other player again receives an envelope that contains 10 Taler (every player only knows his own amount). You do not receive further information about the other player's bid on your screen. Now, you can submit any bid between 0 and 120 Taler once again. The other player bids 45 Taler as in part 1.

# D Control questions (Section 2 - winner's curse)

The following control questions were asked before the subjects started with the actual bidding task. The subjects were only allowed to continue with the experiment after all control questions were answered correctly.

1. You have 50 Taler in your envelope and the other player has 10 Taler in his envelope. You bid 110 Taler and the other player bids 30 Taler. What would be your payoff? (*Correct answer: 30 Taler*)

2. You have 50 Taler in your envelope and the other player has 10 Taler in his envelope. You bid 50 Taler and the other player bids 30 Taler. What would be your payoff? (*Correct answer: 30 Taler*)

3. You have 5 Taler in your envelope and the other player has 50 Taler in his envelope. You bid 25 Taler and the other player bids 90 Taler. What would be your payoff? (*Correct answer: 0 Taler*)

4. You have 5 Taler in your envelope and the other player has 50 Taler in his envelope. You bid 100 Taler and the other player bids 90 Taler. What would be your payoff? (*Correct answer: −35 Taler*)

5. You have 10 Taler in your envelope and the other player has 5 Taler in his envelope. You bid 45 Taler and the other player bids 35 Taler. What would be your payoff? (*Correct answer: −20 Taler*)

6. You have 10 Taler in your envelope and the other player has 5 Taler in his envelope. You bid 60 Taler and the other player bids 35 Taler. What would be your payoff? (*Correct answer: −20 Taler*)

7. You have 55 Taler in your envelope and the other player has 50 Taler in his envelope. You bid 90 Taler and the other player bids 90 Taler. What would be your payoff? (*Correct answer: 15 Taler*)

8. You have 55 Taler in your envelope and the other player has 50 Taler in his envelope. You bid 80 Taler and the other player bids 90 Taler. What would be your payoff? (*Correct answer: 0 Taler*)

# E Proofs (Section 3 - correlation neglect)

*Proof of Proposition 7.* For a given state $G$, the following combinations of correct and false signals can possibly occur within a group of five members. This is illustrated by Table E.1.

| Signals | Probability | Cumulative probability |
|---|---|---|
| 5 correct; 0 false | $\frac{2}{3}^5 \approx 0.131$ | $\sim 0.131$ |
| 4 correct; 1 false | $5 \cdot (\frac{2}{3}^4 \cdot \frac{1}{3}) \approx 0.329$ | $\sim 0.461$ |
| 3 correct; 2 false | $10 \cdot (\frac{2}{3}^3 \cdot \frac{1}{3}^2) \approx 0.329$ | $\sim 0.790$ |
| 2 correct; 3 false | $10 \cdot (\frac{2}{3}^2 \cdot \frac{1}{3}^3) \approx 0.165$ | $\sim 0.955$ |
| 1 correct; 4 false | $5 \cdot (\frac{2}{3} \cdot \frac{1}{3}^4) \approx 0.041$ | $\sim 0.996$ |
| 0 correct; 5 false | $\frac{1}{3}^5 \approx 0.004$ | 1.000 |

Table E.1: Combinations of signals

Given that all five members vote according to the signal they receive, the group will match the state of the world in all cases where they receive at least 3 correct signals. Hence, this will happen with a probability of around 79%.

□

*Proof of Proposition 8.* (i) We first look at the case where all other group members vote according to their signal. Then the probability for being pivotal is given by $P(piv) = \binom{4}{2} \cdot (\frac{2}{3} \cdot \frac{2}{3} \cdot \frac{1}{3} \cdot \frac{1}{3}) = \frac{32}{81} \approx 0.395$.
Without loss of generality, we look at the probability for state $r$, conditional on the voter being pivotal.
We have

$$P(r|piv) = \frac{P(piv|r) \cdot P(r)}{P(piv)}.$$

Given that all other group members vote according to the signal they receive, the probability of being pivotal is not affected by one state or the other and we have

$$P(piv|r) = P(piv|s) = P(piv) = \frac{32}{81}.$$

This leads to

$$P(r|piv) = \frac{\frac{32}{81} \cdot \frac{1}{2}}{\frac{32}{81}} = \frac{1}{2}.$$

The proof works analogous for $G = s$.

(ii) Next we look at the case where all other group members vote according to their preference $p^W$. We denote the probability for having preference $p^R$ as $q$ and the

probability for having preference $p^S$ as $(1-q)$ with $q \in [0,1]$. Then the probability for being pivotal is given by $P(piv) = \binom{4}{2} \cdot (q \cdot q \cdot (1-q) \cdot (1-q)) = 8 \cdot (q^2 - 2q^3 + q^4) = z$. For $z$ it holds that $z \in [0, \frac{1}{2}]$.
Without loss of generality, we look at the probability for state $r$, conditional on the voter being pivotal.
We have

$$P(r|piv) = \frac{P(piv|r) \cdot P(r)}{P(piv)}.$$

Since state $G$ and preference $p^W$ are independent of one another, the probability for being pivotal is not affected by one state or the other and we have

$$P(piv|r) = P(piv|s) = P(piv) = z.$$

This leads to

$$P(r|piv) = \frac{z \cdot \frac{1}{2}}{z} = \frac{1}{2}.$$

The proof works analogous for $G = s$.

□

*Proof (Treatment TwoInd).* When receiving two independent signals, $\sigma_G^1$ and $\sigma_G^2$, which provide the same information, we have $P(r|\sigma_r^1, \sigma_r^2) = P(s|\sigma_s^1, \sigma_s^2) = 0.8$.

Without loss of generality, we look at the case where a subjects receives $(\sigma_r^1, \sigma_r^2)$. Then we have

$$P(r|\sigma_r^1, \sigma_r^2) = \frac{P(\sigma_r^1, \sigma_r^2|r) \cdot P(r)}{P(\sigma_r^1, \sigma_r^2)}.$$

Since both signals are independent and both states of the world are equally likely, we have

$$P(\sigma_r^1, \sigma_r^2|r) = \frac{2}{3} \cdot \frac{2}{3} = \frac{4}{9},$$

$$P(\sigma_r^1, \sigma_r^2|s) = \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9},$$

and

$$P(r) = \frac{1}{2}.$$

This yields

$$P(\sigma_r^1, \sigma_r^2) = \frac{1}{2} \cdot \frac{4}{9} + \frac{1}{2} \cdot \frac{1}{9} = \frac{5}{18}.$$

Plugging this in leads to

$$P(r|\sigma_r^1, \sigma_r^2) = \frac{\frac{4}{9} \cdot \frac{1}{2}}{\frac{5}{18}} = \frac{4}{5} = 0.8.$$

The proof works analogous for $G = s$.

□

# F Instructions (Section 3 - correlation neglect)

Welcome to this online experiment!
It will take approximately **10 minutes**. In addition to your completion fee, **you can earn money depending on your decisions in this experiment**. All possible earnings reported on the following pages are **additional** to your completion fee of 1.00 Pound.

Please read the following consent form before continuing:
I consent to participate in this research study. I am free to withdraw at any time without giving a reason (knowing that any payments only become effective if I complete the study).
I understand that all data will be kept confidential by the researchers. All choices are made in private and anonymously. Individual names and other personally identifiable information are not available to the researchers and will not be asked at any time. No personally identifiable information will be stored with or linked to data from the study.
I consent to the publication of study results as long as the information is anonymous so that no identification of participants can be made.
By clicking the 'Next' button you declare that you have read and understand the explanations and you voluntarily consent to participate in this study.

This experiment will consist of **two parts**.

**IMPORTANT: Only one of the parts will be payoff-relevant for you!**

Hence, your overall payoff in this study will be either:

**(1) Overall payoff** = completion fee + payoff in part 1
or
**(2) Overall payoff** = completion fee + payoff in part 2

Both options, (1) and (2), are equally likely.

We will inform you about your payoff-relevant part at the end of the experiment. Your overall payoff will be calculated when all participants have completed this experiment.

**Part 1**

Basic setup:

You are part of a group consisting of 5 members in total. The remaining 4 members are also participants in this experiment. The whole setup is completely anonymous. This means you neither know the other participants or their decisions, nor do they know you or your decisions.

**The task of your group** is to decide on one of two options, PURPLE or YELLOW, by a **majority vote** (i.e., if at least 3 members of the group vote for the same option, this option will be implemented). The implemented option will determine your payoff.

You will be paid according to the implemented option in the following way:

a) If the group implements option PURPLE, every group member earns a **payoff** that is with **90% chance 0.25 Pounds and with 10% chance 7.75 Pounds** (for each member of the group this will be drawn individually).

b) If the group implements option YELLOW, every group member earns a **payoff** of **1.00 Pound for sure**.

**Part 2**

Basic setup:

As before, **the task of your group** is to decide on one of two options, PURPLE or YELLOW, by a **majority vote**. The implemented option will determine your payoff.

In addition, your group also has a **secret color** (purple or yellow, **and each color is equally likely**).

The **secret color** of your group has an additional effect on your payoff.

In advance, **you do not know whether your group is purple or yellow**.

**Your payoff in part 2 will consist of two components:**

**Component 1:** The implemented option affects the payoff itself (**this is the same**

**as in part 1**).

**Component 2:** If your group **implements the color of your group** (e.g., group is purple and the group votes PURPLE), every group member receives a **bonus of 0.25 Pounds (+0.25)**. If your group **implements the other color** (e.g., group is yellow and the the group votes PURPLE), every group member pays a **penalty of 0.25 Pounds (−0.25)**.

**Observers:**
In order to make an informed decision, there are **three OBSERVERS who know the secret color of your group**.

**OBSERVER 1** (*tells always the truth*): This OBSERVER will tell you the correct color of your group.

**OBSERVER 2** (*tells always the truth*): This OBSERVER will tell you the correct color of your group.

**OBSERVER 3** (*lies always*): This OBSERVER will tell you the wrong color of your group.

You will receive a message by **a randomly drawn OBSERVER** in the form of a short **Note**.

Hence, your **Note** will be **correct in two out of three times**.

Each member of the group receives an individual and randomly drawn **Note** from one of the three OBSERVERS. **Each of the three OBSERVERS is equally likely for each group member**.

*Add-on in treatment TwoCorr*
Furthermore, you also receive an **additional message**. The additional message **is from the same OBSERVER who gave you the first Note**.

*Add-on in treatment TwoInd*
Furthermore, you also receive an **additional message**. The additional message **is randomly drawn from any OBSERVER (including the one who gave you the first Note – all with equal likelihood)**.

# G Voting screens (Section 3 - correlation neglect)

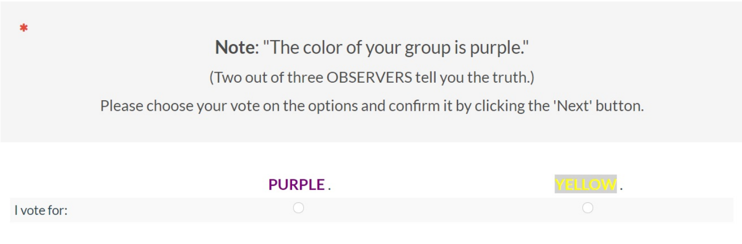



Figure G.1: Typical voting screen in treatment *OneSignal*

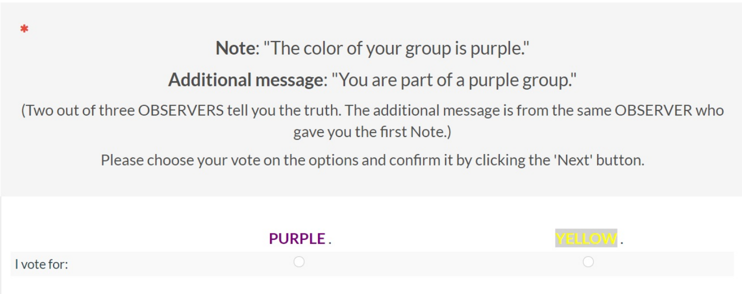



Figure G.2: Typical voting screen in treatment *TwoCorr*

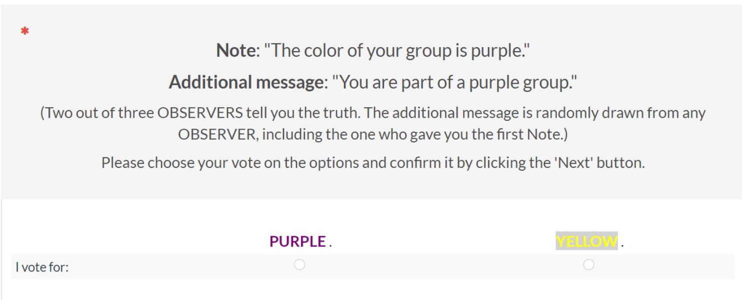



Figure G.3: Typical voting screen in treatment *TwoInd*

# H Balancing table (Section 4 - leadership)

| **Type** | **Treatment** | |
|---|---|---|
| | *NOLEAD* | *LEAD* |
| *Selfish* | 14 | 8 |
| *CC* | 26 | 31 |
| *Altruist* | 4 | 8 |
| *Mismatcher* | 0 | 1 |
| ***Total*** | 44 | 48 |
| | Pearson $\chi^2(3) = 4.2424$; $p = 0.236$ | |

Table H.1: Number of different cooperation types across treatments

# I Translated instructions (Section 4 - leadership)

Welcome to this economic experiment! During the experiment you have the possibility to conduct a task that is explained in detail in the following instructions. In the experiment you can win a non-negligible amount of money. The amount of your payoff depends on your decisions and on the other participants' decisions. During the experiment it is forbidden to communicate with the other participants. Please read through the instructions at hand thoroughly. Should you have questions before or during the experiment, please raise your hand and an experimenter will come to your seat.

**General structure**

The experiment consists of two parts, part 1 and part 2. These two parts are independent from each other - i.e., your decisions in part 1 have no influence on the outcome of part 2 and vice versa. First, you will get instructions for part 1. Once this part is completed, you will get instructions for part 2. In both, part 1 and part 2, you can earn money. Your final payoff consists of the sum of your payoffs from part 1 and part 2. In addition, you will receive a show-up fee of 4 EUR at the end of the experiment. In any case, you will receive your show-up fee, regardless of your decisions in the experiment. Your final payoff will therefore be made up as follows:

**Total payoff = 4 EUR + payoff in part 1 + payoff in part 2**.

Therefore, your total payoff will be at least 4 EUR.

## Part 1

In part 1 you will be grouped with a randomly selected experimental participant from this room, who is your peer for this part. One of you randomly gets the role "Player 1", while the other one gets the role "Player 2". Each of you will now receive **1 EUR** from the experimenter. You can now decide whether you want to KEEP this Euro or SEND it to your peer. When the Euro is sent, it doubles for your peer. Your peer is facing same decision as you. This decision situation takes place sequentially - i.e., the players make their decisions one after the other. This is also shown in the graph below. First, Player 1 (BLUE) makes a decision. Player 2 (RED) can observe this decision and then makes a decision as well. In the experiment you make a decision for both, the case that you are Player 1 and for the case that you are Player 2. For this purpose, please follow the instructions on the screen. At the end of the experiment a random generator will decide whether you or your peer will be given the role of “Player 1”. The other player automatically gets the role of “Player 2”. By combining your decisions, your payoff will finally be formed, as shown in the graphic below. Here, the payoff for player 1 is shown in BLUE and the payoff for player 2 is shown in RED.

[GRAPH]

**Example**
Player 1 chooses SEND.
Player 2 chooses SEND if Player 1 chooses SEND and KEEP if Player 1 chooses KEEP.
In this scenario, both players receive **2 EUR each**.

**Example**
Player 1 chooses KEEP.
Player 2 chooses SEND if Player 1 chooses SEND and SEND if Player 1 chooses KEEP.
In this scenario, player 1 receives **3 EUR** and player 2 receives **0 EUR**.

## Part 2

Part 2 consists of **4 periods** and each period consists of **7 rounds**. This is illustrated in the following table.

| **Period** | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| **Round** | 1 2 3 4 5 6 7 | 1 2 3 4 5 6 7 | 1 2 3 4 5 6 7 | 1 2 3 4 5 6 7 |

In part 2 you can earn Taler. The conversion is as follows:

**10 Taler = 1 Euro**

At the end of the experiment you will receive your endowment of Taler paid out in EUR. A **random period is selected** for your payment in part 2. This means that only **one of the four periods is relevant for your payoff**. However, you do not know beforehand which of the periods is relevant, but you get this information only at the end of the experiment. At the beginning of each period, groups of **4 participants** are formed. These groups are fixed for the complete period - i.e., they will not be mixed within a period. Each group member is called either **A**, **B**, **C**, or **D**. For each new period you are **randomly assigned to a new group**. This also means that your label (**A**, **B**, **C**, or **D**) can change.

**The task**
At the beginning of each period you are endowed with **20 Taler**. In each round, you can invest all or a part of your endowment into a **group project** (individual contribution). The part of your endowment that you do NOT to invest, goes into your **private pot**, which you keep for yourself. The investments of **all 4 group members** go into a **group pot**. The amount within the group pot will then be **increased by 50%** and equally split among all group members, regardless of how much each individual has contributed. The new endowment of a participant at the end of a round is then calculated as follows:

**new endowment = old endowment − individual contribution + 1.5*(group pot/4) = private pot + 1.5*(group pot/4)**

After each round you will be informed about your new endowment. **Additionally, you can also see** how much the other members in your group have contributed and their current endowment (see Screenshot 1). All amounts are always rounded to the next integer. For the subsequent rounds, each player can choose any individual contribution to the group pot which can be at most as high as the current equipment of Taler of this player. In contrast to round 1, the equipment of Taler within the rounds 2 - 7 can be smaller or larger than 20.

[SCREENSHOT 1]

***** ONLY in treatment NOLEAD *****

**The players**
In each round all group members (**A**, **B**, **C** and **D**) choose their individual contributions simultaneously. (A typical decision screen is shown in Screenshot 2). This means that all group members only learn at the end of a round which group member contributed how much to the group pot.

**Your current endowment is displayed at the top of the screen. This is the maximal amount you can choose.**

[SCREENSHOT 2]

***** ONLY in treatment LEAD *****

**The starting player**
At the beginning of each period, a **starting player** is randomly determined. This player remains the starting player for the entire period - i.e., for the entire 7 rounds.

All group members will be informed about who the starting player is (**A**, **B**, **C**, or **D**). Each round now consists of the following two sections:

1. The selected starting player can now **first** decide on his individual contribution (see Screenshot 2). If you are NOT the starting player, simply click "OK" to continue while the starting player chooses his individual contribution (see Screenshot 3).

2. The other players will be informed about the contribution of the starting player and can then choose their own individual contributions (see Screenshot 4).

**Your current endowment is displayed at the top of the screen. This is the maximal amount you can choose.**

[SCREENSHOT 2]
[SCREENSHOT 3]
[SCREENSHOT 4]

***** For BOTH treatments *****

**Procedure during the experiment**

(i) At the beginning you will be asked to answer a number of **control questions**. You cannot earn money by answering these questions. The purpose of these questions is to make sure that you understand the game and the rules correctly.

(ii) You will then be given the opportunity to test on a example calculator what endowment you would have if you and the other members in your group would select a specific contribution (a maximum of 20 Taler, as in round 1). You can choose your contribution with a slider next to the button "Your contribution", and the contributions of the other group members by using the sliders below. With this example calculator, you can test different combinations. The purpose of the example calculator is to make yourself familiar with the basic principle of the task. Your entries here have no influence on the further course of the experiment. Please note that you can use this example calculator for a maximum of 180 seconds (= 3 minutes). Then the experiment will automatically continue (you cannot actively continue by clicking).

(iii) After the testing phase with the example calculator, the actual task begins and you start with the first period as described above. Each period is exactly the same, but in each period new groups are formed randomly.

# J Instructions (Section 5 - cooperation types)

***Basics***
This online experiment will consist of **two parts**.
First, part 1 will be explained. After part 1 ends, you will receive instructions for part 2. Your decisions in part 1 do not influence your payoff in part 2 and vice versa.

In this study you will earn POINTS. **Each POINT is worth 0.05 Dollar (20 POINTS = 1 DOLLAR)**. At the end of the study you receive your amount of POINTS cashed out in Dollar.

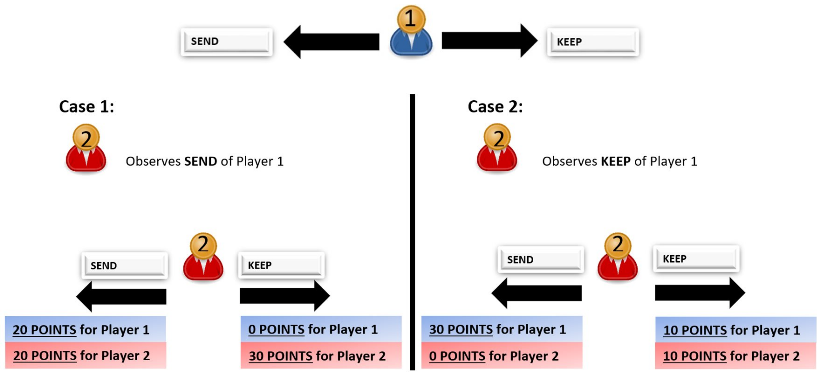
In this study, you must answer control questions to ensure that you have understood the task correctly (there are five control questions in total). Only if you answer them correctly you can complete this survey. **The control questions require some small calculations**. If you give a wrong answer to a control question, you can try **multiple times** until you find the correct solution.

***Treatment SPD***
You will be matched with one other random MTurker who also participates in this study. One of you has the role of "Player 1" and the other one has the role of "Player 2". Each of you is endowed with **10 POINTS**. You have to decide whether you want to **KEEP** your 10 POINTS or whether you want to **SEND** your 10 POINTS. If the POINTS are sent, they double for the other player. The other MTurker has to make the same decision.

This game is played sequentially – i.e., the players make their decisions subsequently (this is illustrated by the graph below).

First Player 1 (BLUE) makes a decision. Player 2 (RED) observes this decision and makes a decision as well. In this study, you make a decision for both roles, Player 1 and Player 2 (follow for this purpose simply the instructions on the screens). At the end of the study, a random device determines the role of you and the other MTurker. There are two possibilities: you are Player 1 and the other MTurker is Player 2 or you are Player 2 and the other MTurker is Player 1. The combination of the decisions of you and the other MTurker determines your payoff in this game, as shown in the graph below.

**Conversion rate: 20 POINTS = 1 DOLLAR.**




Before starting with the actual decisions, you are asked to answer two short control questions to make sure that you have understood all rules of the game correctly.

***Treatment FGF***

You will now be in a group of 4 MTurkers. Each MTurker must decide on the division of 20 POINTS. You can put these 20 POINTS in a **private account** or you can invest them fully or partially into a **project**. Any POINT that you do NOT invest into the **project**, will automatically be transferred to your **private account**.

<u>Your income from the private account:</u>

For each POINT you put in your private account you will earn exactly one POINT. Nobody except you earns something from your private account.

<u>Your income from the project:</u>

The amount of POINTS contributed to the project by ALL group members, will be increased by 60% and then equally split among all group members. This means, each group member will receive the same income from the project. Consequently, for each POINT invested in the project each group member (including yourself) receives $1.6/4 = 0.4$ POINTS.

Hence, for each group member the income from the project will be determined as follows:

Income from the project = sum of contributions to the project x 0.4.

For example, if the sum of all contributions to the project is 70 POINTS, then you and all group members will get a payoff of 70 x 0.4 = 28 POINTS each from the project. If the sum of contributions is 15 POINTS, then you and all group members will get a payoff of 15 x 0.4 = 6 POINTS each from the project.

Your total income:

Your total income is the sum of your income from the **private account** and the **project**:

**Income from the private account** (= **20** - contribution to the project)
+
**Income from the project** (= 0.4 x Sum of contributions of all four players to the project)
= **TOTAL INCOME**.

**Conversion rate: 20 POINTS = 1 DOLLAR.**

**Your decisions:**

In this part of the study, each participant has to make two types of decisions. In the following we call them "**unconditional contribution**" and "**conditional table**":

With the "**unconditional contribution**" to the project you have to decide how many of your 20 POINTS you want to invest into the project. You do not know how much the other players will invest.

Your second task is to fill a "**contribution table**". For each possible average contribution of the other group members (rounded to the nearest integer), you must specify how many POINTS you want to contribute to the project. Thus, you can condition your contribution on the contribution of the other group members.

In each group a random mechanism will select **one group member**. For the randomly selected group member only the **contribution table** will be the payoff-relevant decision. For the other three group members only the **unconditional contribution** will be the payoff-relevant decision. When you make your **unconditional contribution** and when you fill out the **contribution table**, you do not know whether you will be selected by the random mechanism. Hence, you have to think carefully about both types of decisions because both can be relevant to you. The combination of the decisions of you and the other group members determines your payoff in this game.