## 

**Using social preferences and risk-attitude to explain and predict decision-making in game theory**

**Jeroen van Buren**

u490632 / 2017269

Supervisor: P. Spronck

Inhoudsopgave

[Preface 4](#_Toc521589232)

[Abstract 5](#_Toc521589233)

[2 Introduction 6](#_Toc521589234)

[3 Related work 8](#_Toc521589235)

[3.1 Game theory & behavioural game theory 8](#_Toc521589236)

[3.2 Social preferences 8](#_Toc521589237)

[3.2.1 Ultimatum game and negative reciprocity 8](#_Toc521589238)

[3.2.2 Dictator game and altruism 9](#_Toc521589239)

[3.2.3 Trust game, trust, and positive reciprocity 10](#_Toc521589240)

[3.3 Risk-attitude 10](#_Toc521589241)

[3.4 Opportunities 10](#_Toc521589243)

[3.5 Game scenarios 11](#_Toc521589244)

[3.5.1 Public goods 11](#_Toc521589245)

[3.5.2 Ultimatum game 12](#_Toc521589246)

[3.5.3 Prisoner’s dilemma 12](#_Toc521589247)

[3.5.4 The chicken game 12](#_Toc521589248)

[4 Methods 13](#_Toc521589249)

[4.1 Questionnaire 13](#_Toc521589250)

[4.1.1 Designing the questionnaire 13](#_Toc521589251)

[4.1.2 Piloting the questionnaire 13](#_Toc521589252)

[4.1.3 Amazon Mechanical Turk 14](#_Toc521589253)

[4.2 Data and pre-processing 14](#_Toc521589254)

[4.2.1 Time 15](#_Toc521589255)

[4.2.2 Multi-indicator concepts 15](#_Toc521589256)

[4.3 Exploratory data analysis 16](#_Toc521589257)

[4.3.1 Public goods 16](#_Toc521589258)

[4.3.2 Ultimatum game: offer 17](#_Toc521589260)

[4.3.3 Ultimatum game: response 17](#_Toc521589261)

[4.3.4 Prisoner’s dilemma 18](#_Toc521589262)

[4.3.5 Chicken game 18](#_Toc521589263)

[4.3.6 Preferences and risk-attitude 18](#_Toc521589264)

[4.4 Experimental procedure 19](#_Toc521589265)

[4.4.1 Algorithm types and evaluation criteria 19](#_Toc521589266)

[4.4.2 Cross-validation 19](#_Toc521589267)

[4.4.3 Bivariate analysis 20](#_Toc521589268)

[4.4.4 Model building and parameters 20](#_Toc521589269)

[5 Results 21](#_Toc521589270)

[5.1 Bivariate analysis 21](#_Toc521589271)

[5.2 Modelling 23](#_Toc521589272)

[5.2.1 Public goods 23](#_Toc521589273)

[5.2.2 Ultimatum game: offer 25](#_Toc521589274)

[5.2.3 Ultimatum game: response 26](#_Toc521589275)

[5.2.4 Prisoner’s dilemma 27](#_Toc521589276)

[5.2.5 Chicken game 29](#_Toc521589277)

[6 Discussion 30](#_Toc521589278)

[6.1 Discussing and answering the research questions 30](#_Toc521589279)

[6.1.1 Deviations 30](#_Toc521589280)

[6.1.2 Predicting and explaining 30](#_Toc521589281)

[6.1.3 General findings and predictive limitations 31](#_Toc521589282)

[6.2 Limitations 31](#_Toc521589283)

[6.3 Future research 32](#_Toc521589284)

[Conclusion 33](#_Toc521589285)

[Bibliography 34](#_Toc521589286)

[Appendix A: Tables and figures 36](#_Toc521589287)

[Appendix B: Questionnaire 39](#_Toc521589288)

[Information sheet 39](#_Toc521589289)

[Informed consent form 40](#_Toc521589290)

[Structure and explanation 41](#_Toc521589291)

[Game scenarios 42](#_Toc521589292)

[Prisoner’s dilemma 43](#_Toc521589293)

[Chicken game 44](#_Toc521589294)

[Ultimatum game: offer 45](#_Toc521589295)

[Ultimatum game: response 46](#_Toc521589296)

[Public goods 47](#_Toc521589297)

[Investment game 48](#_Toc521589298)

[The preference survey module 49](#_Toc521589299)

[Demographics 56](#_Toc521589300)

[Debriefing 57](#_Toc521589301)

[Additional questions 57](#_Toc521589302)

[Feedback 57](#_Toc521589303)

[Appendix C: Pilot questions 58](#_Toc521589304)

[Appendix D: Packages and “thesisFunctions.R” 60](#_Toc521589305)

[Packages 60](#_Toc521589306)

[thesisFunctions.R file 60](#_Toc521589307)

# Preface

This thesis is written for a research graduation project for the Master information and communication science: data science for business and governance track Tilburg University.

In game theory, a field that assumes rationality, theory often does not correctly predict human decision-making. Since neuroscience, psychology, human decision-making, and rationality are of great interest of mine and being familiar with the fact that human decisions often violate rational choice theory, a combination of these two fields seemed an interesting research topic. It was not until later that behavioural game theory came to my notice, a research area that combines these fields and became the focus of this research.

I would like to thank my supervisor for helping me through the thought process of formulating and narrowing down my research question, and for his availability and supervision. Furthermore, I would like to thank the respondents who filled in the questionnaire and provided the data needed for the research. I would also like to thank all researchers that contribute to the field of game theory and behavioural game theory, neuroscience and psychology, computer science and statistics. This thesis has a multidisciplinary nature and builds on knowledge of all these different domains.

# Abstract

The focus of this research is the explanatory and predictive value of social preferences and risk-attitude for human decision-making in game theory. It focuses on a subset of games: the public goods game, the ultimatum game, the prisoner’s dilemma, and the chicken game. It tries to predict and explain decisions in these scenarios using statistical- and machine learning techniques. Earlier theoretical models do not incorporate heterogeneity of the preferences, and merely fit and post-hoc explain the findings. The models in this research incorporate individuality and apply cross-validation to avoid overfitting. Online questionnaires are the primary source of data for this research and the sample size of the dataset is 204. The models and analysis show that the associated preferences and risk-attitude are dissimilar for all scenarios. More precisely, altruism relates with the public goods game. Positive reciprocity with the offer, and negative reciprocity with the minimum acceptable amount in the ultimatum game. Prisoner’s dilemma relates with altruism, trust, and possibly with negative reciprocity. The chicken game lacked any valuable predictors for the features investigated. Overall, the social preferences provide insights and understanding of the decision-making in the game scenarios. The social preferences have moderate to weak predictive value depending on the game. However, the chicken game and risk-attitude have no predictive or explanatory value. Furthermore, the best algorithm types and models are dissimilar, showing that a flexible modelling method is needed for making predictions. Future research may try to apply this method on cleaner datasets, test different game scenarios or design new models that incorporate flexible, heterogeneous social preferences models.

# Introduction

The goal of this study is to explain and predict decision-making in game theory. This research takes a psychological grounded perspective of game theory, called behavioural game theory. Whereas game theory is a mathematical- and analytical study that supposes self-regarding rational decision makers, behavioural game theory tries to extend game theory with concepts such as cognitive limitations, emotions, and social preferences. This is needed because humans have a limited working memory and their behaviour is complex and diverse. Accordingly, their decisions frequently violate game theory; experimental research provides support for the fact that game-theoretical solutions often deviate from human decision-making in game scenarios (Camerer, 2003; Gintis, 2005). This is especially relevant in games wherein the best strategy relies on the opponent’s decision, as is in the ultimatum game and the chicken game. For example, in an ultimatum game scenario where an offeror randomly receives ten dollars to distribute. A game theoretical player that makes an offer of one dollar probably faces rejection by the responder, since this offer may be considered unfair. If so, the offeror should have offered more money to avoid rejection (Guth, et al., 1983; Güth & Kocher, 2014). In these scenarios correct prediction of the opponent’s behaviour is essential to maximise pay-off. However, not only predictions are of importance, but also the understanding of what exactly makes people reject such an offer.

Behavioural game theory uses experimental evidence to find the deviations and propose relevant concepts. It also tries to build models that aim to predict and explain human decisions better than game theory does. A specific family of models from behavioural game theory are social preference models. The social preference models consider the role of social preferences in decision-making and uses these to explain and predict behaviour in single-shot games (Rabin, 1993; Camerer & Ho, 2015). Beside these social preferences, risk preferences also influence decision-making in situations of uncertainty and therefore may be important for understanding and predicting behaviour (Falk, et al., 2016; Dohmen, et al., 2011).

The concepts of social preferences and risk-attitude form the theoretical framework of this research. Since the literature describes a lot of different underlying motives and concepts, the goal of this research is to find out whether, which and if proper, how much explanatory and predictive value these social preferences and risk-attitudes have in game theory. More specifically, the research investigates the following game scenarios: the public goods game, the ultimatum game, prisoner’s dilemma, and the chicken game. It uses statistical learning techniques to answer these questions. Earlier social preference models take a general approach, solemnly fit the experimental data, and explain the findings afterwards. In contrast, statistical learning techniques allow for prediction on a separate test data, thereby avoid fitting models that are too complex and contain irrelevant concepts. These techniques also incorporate heterogeneity. All in all, the use of statistical learning techniques targets the shortcomings of earlier models. The use of these statistical learning techniques is uncommon in behavioural game theory (Wright & Leyton-Brown, 2017). It therefore provides an opportunity to answer the research questions in a novel way.

The novel applications of the techniques, the methodology, and the potential understanding of the decision-making process are not only of practical relevance to an existing and active research field, but also to the field of decision-making in general. In other words, these methods and concepts are applicable in different game scenarios as well as to decision-making scenarios in the social sciences, business, economics, politics, and psychology. It can help to understand and reduce uncertainty in these situations.

Altogether this leads to the following research questions:

1. *Do the findings in the game scenarios deviate from the game theoretical solution?*
2. *Which and how do social preferences and risk-attitude relate with the decisions in the game scenarios?*
3. *How much explanatory and predictive value do social preferences and risk-attitude have?*
4. *Are the concepts and models generalizable across the games?*

Regarding the research questions, this thesis answers these using a cross-sectional research design, wherein a self-completion questionnaire serves as the primary data source. The questionnaire contains the game scenarios and measures the preferences and risk-attitude of the respondents. Exploratory data analysis answers the first question and concludes whether results correspond with the game theoretical solutions. Bivariate analysis answers the second and third questions. Statistical learning techniques accommodate in the understanding of the relations between the concepts and the decisions in the game scenarios and answer the second and third questions more thoroughly. The use of cross-validation supports the measurement of predictive performance of the models and concepts and answers the third question. Summarizing and comparing the findings answers the third and fourth question.

The thesis starts with an extensive description of related work that provides a theoretical framework. It focuses on the underlying social preferences and risk-attitude that explain deviations from rational decision-making. These preferences serve as the predictors in the models. Next, the section methods address the general implementation, the process of designing and piloting the questionnaire, and the use of Amazon Mechanical Turk to collect the data. Additionally, it addresses the data- and pre-processing steps. Thereafter the section continues with an exploratory data analysis of all game scenarios, and the preferences and risk-attitude. The section ends with the experimental procedure and contains a description of the bivariate analysis, evaluation criteria, and the statistical learning models. The results section elaborates on the models and findings of all game scenarios. The discussion section evaluates and combines all findings from the previous sections and discusses limitations. The last section is the conclusion of the thesis and concisely answers the research questions. The appendix encompasses the tables and figures, the questionnaire, the piloting questions, and information about the used packages and functions.

# Related work

## Game theory & behavioural game theory

Game theory uses simplified assumptions to find game theoretical solutions. These assumptions regard players as exclusively self-interested, having full rationality and common knowledge (Colman, 2003; Gintis, 2005). Analysis of traditional game theoretical solutions such as Nash equilibrium and backward induction require these strong assumptions to make definite predictions about the opponent’s behaviour. Experimental studies show that these assumptions do not always hold and that experimental subjects often play dominated or non-equilibrium strategies (Camerer, 2003). Even though players might show a degree of rationality and self-interest, it is not enough to justify analysis based exclusively on traditional game theoretical analysis (Crawford, 1997). Consequently, the use of traditional game theoretical analysis leads to inaccurate predictions of actual behaviour. In addition, the presence of a player that plays non-game theoretically can also affect the optimal strategy.

To deal with these shortcomings of traditional game theory, behavioural game theory emerged. The main concern of behavioural game theory is careful observation of how people actually play in game scenarios. Behavioural game theory uses experimental studies to examine deviations and to understand human behaviour and this provides the theoretical framework for future studies (Crawford, 2002). This theoretical framework is essential to explain and predict behaviour correctly and is discussed in more detail below.

## Social preferences

Experimental studies show that players are influenced by various kinds of social preferences that contradict with the simplified assumptions of traditional game theory. In principal, game theory allows for the pay-offs to be determined by rational preferences other than monetary or pure self-interest and may include social preferences (Colman, 2003; Gintis, 2005; Levitt & List, 2007). Most social preferences models tried to transform pay-offs by accounting for these preferences based on the insights and concepts of experimental research. However, the utilities are hard to determine with certainty and precision, and often still fail to capture the actual pay-offs. This is even harder for the individual player. Therefore, the models do not provide accurate, individual predictions (Camerer & Ho, 2015).

Nevertheless, the models and experimental research help to explain behaviour and improve general predictions in games where traditional game theory fails (Camerer, 2003). It provides the needed theoretical framework and concepts for this research. The experimental game scenarios and the corresponding relevant social preferences (such as: positive- and negative reciprocity, altruism and trust) that may provide insights for decision-making in game theory are described in more detail below (Camerer & Ho, 2015; Falk, et al., 2016).

### Ultimatum game and negative reciprocity

The ultimatum game captures the phenomenon of negative reciprocity. In an ultimatum game the offeror makes an offer to divide a sum of money. The responder can, in turn, either accept or reject the offer. In case that the responder rejects the offer, both get nothing; if the responder accepts the offer, the proposed offer is paid out. The game-theoretical solution to this game is for the responder to accept any offer that is larger than zero. However, experimental research has shown that players often reject bargains that seem unfair to them, even if that leads to no monetary pay-off. This indicates that subjects do not perceive their pay-offs as purely monetary and care about fairness (Guth, et al., 1983). This game scenario captures the phenomena of negative reciprocity: people are willing to punish others if treated unfairly, even at a cost for themselves (Güth & Kocher, 2014; Camerer & Ho, 2015). Which in this scenario means rejection of an unfair offer.

As mentioned above, a responder should accept the smallest, positive amount offered, since receiving any amount greater than zero dollars is better than receiving nothing. A traditional offeror knows this and thus applies backward induction to determine the best strategy, which is to offer the smallest, positive amount. According to traditional game theory the traditional responder accepts this and the offeror then earns the maximum pay-off. However, in practice, offering an amount that is more evenly distributed and that is perceived as fairer by the responder has a greater chance of being accepted.

It is noteworthy that, in practice, a traditional offeror receives a significantly worse pay-off than the offeror that would incorporate the social preference (i.e., negative reciprocity). Accordingly, an offeror must make an educated guess based on their belief about the responder’s care of fairness and what their minimum acceptable amount is. This approach integrates the responder’s negative reciprocity in the decision-making of the offeror (Güth & Kocher, 2014).

However, the offeror could also have social preferences and prefer to make a generous offer, to appear, feel, or be fair. Such a decision relates to positive reciprocity and altruism (Camerer, 1997). Risk-attitude possibly plays a role in determining the offer as well, since offerors might balance the expected minimum acceptable amount with the risk they are willing to take by offering a lower amount.

An interesting point is that for a responder it is beneficial to have a reputation of having negative reciprocity. Because if the offeror knows that the responder accepts an offer of one dollar, there is no reason to offer more, except for generosity, as mentioned above. In this game the possible rejection of unfair behaviour causes offers to rise, and ‘co-operation’ is achieved in the ultimatum game, this behaviour is often the effect when reciprocity is into play (Axelrod, 1984). Even though the idea is interesting, it is not pursued any further in this thesis.[[1]](#footnote-1)

### Dictator game and altruism

The dictator game captures the social preference altruism. In a dictator game a dictator divides a sum of money, but unlike in the ultimatum game, the responder turns into a receiver and *cannot* accept or reject the offer. A traditional dictator would offer nothing since this maximises the dictator’s pay-off. In contrast, empirical results have shown that some people offer more than zero dollars and decide to share money with the responder even at cost for themselves (Forsythe, et al., 1994; Engel, 2011). The dictator game supposes to capture altruism, the care and concern for the happiness and well-being of other people (Camerer & Ho, 2015). This preference might, as noted before, influence the offer in an ultimatum game, but also the investment in the public goods game, and co-operation in the prisoner’s dilemma and the chicken game.

However, this altruism might be impure altruism and not genuine concern for the welfare of others but emerge from manners and the norms of generosity (Camerer & Ho, 2015). Similarly, it might be self-serving altruism. Whatever the exact nature is, the general concept of altruism still plays a key role in explaining the findings. Still, there is more criticism on the role of the altruism in the dictator game; for instance, the research by (Bardsley, 2007) shows that altruistic behaviour can also work in the opposite direction. When a dictator can take, instead of give money, they are just as willing to take money from the other player. This provides support for the argument that altruistic behaviour in the dictator game is merely an experimental artefact. This is not necessarily a rejection of the role of altruism in game theory and decision-making but does indicate that context-specific social norms play a role (Bardsley, 2007).

### Trust game, trust, and positive reciprocity

The trust game covers the phenomenon of trust and positive reciprocity. This game is like the dictator game but includes an added first stage. In this stage the proposer has the choice to keep a fixed amount of money or transfer a variable amount to the responder. When money is transferred the total amount is increased in value by a factor of more than two and is then in the hands of the responder, from now on called the dictator. The dictator can unilaterally decide how much money to keep and how much to return to the original offeror. Both players can receive a higher monetary pay-off if both co-operate and transfer the money back and forth.

The game-theoretical solution for the dictator is to maximise pay-off by returning nothing. Consequentially, a traditional offeror with common knowledge of this applies backward induction and does *not* transfer any money. Despite this, experiments have shown that in general, some offerors transfer money to the dictator and that some dictators return the transferred money (Berg, et al., 1995; Johnson & Mislin, 2011). Dictators tend to return the transferred money with a slightly positive surplus (Camerer, 1997).

The offeror shows trust in the dictator and the dictator returns this trust by showing positive reciprocity (Berg, et al., 1995). The concept of trust is defined as the reliance on, and the assumption that others have good intentions and integrity. It might influence decision-making in the public goods game, prisoner’s dilemma, and the chicken game.

The concept of positive reciprocity is defined as the willingness to treat others well if they are treated fairly even at a cost for themselves. It possibly relates to the offer in an ultimatum game, the investment in the public goods game, and co-operation in prisoner’s dilemma and the chicken game. A criticism on the preference positive reciprocity is that it is quite like altruism and that this experimental set-up allows for the return of money to be partially determined by altruism, as these preferences are hard to separate in this game. However, in general the amount returned by the dictator is higher in the trust game than in the dictator game (Berg, et al., 1995; Forsythe, et al., 1994). In any case, the trust game shows that extending traditional game theory with the social preferences trust and positive reciprocity seem to better explain actual behaviour.

## Risk-attitude

Risk-attitude influences decision-making in situations of uncertainty. Moreover, risk-attitude relates to risky economical behaviour and could influence decision-making in game theory (Dohmen, et al., 2011; Houser, et al., 2010). Because this research extends game theory with social preferences uncertainty increases and risk-attitude might affect decision-making in game scenarios. As noted before, risk may influence in the height of the offer in the ultimatum game or whether players in the chicken game pick the riskier- or the safer option.

## Opportunities

The empirical findings from above show theoretical and empirical support for the role of social preferences and risk-attitude in the context of game theory. Earlier models incorporated different social preferences in the models; however, risk preferences are not integrated in these models (Rabin, 1993; Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999). In addition, these models try to transform pay-offs and are homogeneous models, even though these preferences and risk-attitude are in fact heterogeneous (Dohmen, et al., 2011; Falk, et al., 2016). Subsequently, a method that does not consider these differences is not sufficiently flexible and too general to make correct individual predictions. Moreover, as these models train and fit on all available data and therefore do not prevent overfitting, they may fit relations that do not have actual predictive value (Wright & Leyton-Brown, 2017). There are more factors that influence decision-making that are not integrated in these models, such as framing effect and focal points (Camerer, 1997).

These shortcomings of earlier models limit the predictive and explanatory value of these models and provide opportunities for statistical learning models to answer the main research question whether, which and how much these social- and risk preferences influence decision-making in game theory. To answer this question this research takes the same basic approach as in all behavioural game theory research: it uses a theoretical framework, collects empirical data, and tests and builds models on this data.

The theoretical framework provides knowledge about what predictors may influence decision-making. In this research, the social preferences and risk-attitude serve as predictors for the models and include: negative- and positive reciprocity, altruism, trust and risk-attitude.

A questionnaire was used to collect the empirical data and to measure the predictors and the game scenarios. The part of the questionnaire that compromises the social preferences and risk-attitudes is specifically designed for employing these preferences to allow for improved, individual prediction of behaviour (Falk, et al., 2016; Falk, et al., 2018). The questionnaire is explained in more detail in the next chapter. The availability of a validated questionnaire facilitates the research of relations between social- and risk preferences and the decisions made by heterogeneous players in a more direct, generalizable, and quantitative way.

The use of statistical learning techniques for modelling has multiple advantages. Firstly, it gives a sign of the predictive performance of the model and avoids overfitting. The combination of cross-validation and variable selection models allow that only the predictors that have predictive value are included in the model. It also provides understanding in which concepts relate with decision-making in these scenarios. Furthermore, these methods facilitate individual predictions since they use empirical findings to learn more general relations between the preferences and the dependent variable and predict the dependent variable given these individual preferences. Lastly, these models can consider focal points and framing effects from the empirical findings and have no theoretical limitations regarding this.

## Game scenarios[[2]](#footnote-2)

In this research the following games are studied to examine the predictive and explanatory performance of the social- and risk preferences. These scenarios were chosen because they improve external validity, they rely on different theoretical axioms, and are simple one-shot games.

### Public goods

The first game scenario is the public goods game. It is a simplification of a type of cooperative system. A player has the choice to contribute to a public pot. Each dollar contributed to the pot doubles in amount. The uncontributed amount is kept by the player. After all players decide, each player receives an equal share of the public pot *regardless* of their individual contribution. The game theoretical solution and dominating strategy is to invest zero dollars, since a player shares evenly in the pot no matter what amount that player invests in the pot. However, this is an inefficient solution; if everybody cooperates *all* respondents get more money, but this is not a stable solution since the dominating strategy remains unchanged.

In the game scenario in the questionnaire, respondents had to decide how much dollars to contribute to the public. This investment must range from zero to ten, in whole dollars. The respondents were told that the game was played against ninety-nine other respondents, which is irrelevant from a traditional game theoretical perspective.

### Ultimatum game

The ultimatum game was already discussed in section 2.2.1. In the game scenario used in the research the offeror can offer from zero up to ten dollars, in whole dollars. The responder can pick their minimum acceptable amount in the same range. In this scenario the game theoretical solution is to offer and accept the minimum, positive amount of one dollar.

### Prisoner’s dilemma

In the prisoner’s dilemma both players must choose whether to cooperate or defect. The used pay-offs in the questionnaire are found in Table 1. The dominating strategy is to always defect for both players and therefore the Nash equilibrium is defect-defect. It is an inefficient equilibrium, since in this equilibrium both players would have rather co-operated together or preferred that the other player would do so rather than defect together. When both players co-operate this maximises total pay-offs.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | Player two | |
| Cooperate | Defect |
| Player  one | Cooperate | 12, 12 | 0, 20 |
| Defect | 20, 0 | 6, 6 |

Table 1: Pay-off table of the prisoner’s dilemma. The underlined values correspond with the optimal strategy.

### The chicken game

In the chicken’s game both players must choose chicken or dare. The best strategy in the chicken game is dependent on the other player’s decision and is the opposite of it. The chicken game has two Nash equilibria: cooperate-dare and dare-cooperate and therefore has a mixed Nash equilibrium. The mixed Nash equilibrium for the scenario in the questionnaire is calculated and corresponds with randomly choosing chicken in 30 percent and dare in 70 percent of the times.[[3]](#footnote-3) The pay-offs are illustrated in Table 2.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | Player two | |
| Chicken | Dare |
| Player  one | Chicken | 10, 10 | 3, 17 |
| Dare | 17, 3 | 0, 0 |

Table 2: Pay-off table of the chicken game. The underlined values correspond with the optimal strategy.

# Methods

Statistical learning techniques need relatively large samples; in general models perform better on more data. However, linear models are more robust under smaller sample sizes and do not suffer from the curse of dimensionality (James, et al., 2013). These robust models were preferred in this research. In any case, a fast and economically feasible method was needed to collect this relatively large dataset. An online self-completion questionnaire was conducted for this purpose. The questionnaire measured the social preferences, risk-attitude, and responses on the game scenarios. For the collection of the data and design of the online questionnaire, the platform Qualtrics was used.[[4]](#footnote-4) Amazon Mechanical Turk was used to gather respondents.[[5]](#footnote-5) The programming language R was used under the version 3.4.3. The used packages and especially created functions for this thesis are described in Appendix D: Packages and “thesisFunctions.R”.[[6]](#footnote-6) All raw data, the questionnaire, the code, and analysis are available online to assure reproducibility; to enhance this an R markdown document with a complete and extensive description of all analysis steps and code is also available online.[[7]](#footnote-7)

## Questionnaire

### Designing the questionnaire

The full questionnaire used for the research is given in Appendix B. The first part of the questionnaire is the game scenario section. The second part consists of the social preferences and risk-attitude. This part makes use of an extensively validated survey that improves measurement validity, reliability and generalizability: *The Preference Survey Module: A Validated Instrument for Measuring Risk, Time and Social Preferences* designed by (Falk, et al., 2016) and used in (Falk, et al., 2018).[[8]](#footnote-8) The preferences are all multi-indicator concepts, except for trust. Multi-indicator concepts are not caught by one measure, but by the combination of measures that captures the underlying concept. Furthermore, the module is designed to measure these preferences in a time and cost-efficient way outside of the lab and with minimal loss of explanatory value. Additionally, the use of this survey module improves the generalizability and reproducibility of this research (Falk, et al., 2016). The last part of the questionnaire embodies demographic questions about the responder’s gender, age, and nationality.

### Piloting the questionnaire

The self-completion questionnaire was carefully piloted before final implementation. The questions and instructions were piloted on the ease and clarity, common pitfalls, and wording issues which was especially important for the game scenario section (Bryman & Bell, 2011). In addition, the questionnaire was piloted on the time to complete and the order and flow of the questionnaire. The questionnaire was piloted on five respondents, every respondent received a different version since it was designed, tested, and re-designed again after each interviewee with the help of qualitative research methods. In Appendix C: Pilot questions, the pilot questions are described. The pilot followed two stages.

In the first pilot stage three respondents were interviewed and took around two hours per interview. The set-up was a combination of an *offline* self-completion-questionnaire and a semi-structured interview. The relevant pilot questions were asked after each game scenario, the preferences survey module, the demographics, and at the end of the questionnaire. This interrupted the flow of the survey but improved the recollection of the questions. After each respondent, the answers were evaluated and the questionnaire was adjusted and re-designed. For instance, the wording, the order, and description of the games were changed. Because of time constraints multiple games such as the trust game and the beauty contest were excluded.

After the first stage the online questionnaire was developed via Qualtrics. Attention check questions were added because these improve the quality of the responses by avoiding robots and inattentive respondents on Amazon Mechanical Turk. The questionnaire was conducted, uninterrupted, twice online to sustain the flow of the questionnaire. The focus was on how much time the respondents spend on the questionnaire. After the respondents completed the full questionnaire, the interview was conducted. Again, the answers were evaluated and re-designed after each interview. For example, some wording was changed. This resulted in the definitive version of the questionnaire (see Appendix B: Questionnaire).[[9]](#footnote-9)

### Amazon Mechanical Turk

The final online questionnaire was conducted from 29th of May to 31st of May of 2018 via Amazon Mechanical Turk. Amazon Mechanical Turk was chosen because it offers many respondents at a relative low-cost and in a brief time span. The quality of responses are better than data gathered via new media platforms since, beside the attention check questions, qualifications ascertain this.[[10]](#footnote-10) As an example, since good comprehension of English was necessary for the understanding of the games, the location of the respondents was set to the United States.

## Data and pre-processing

Respondents were directed via Amazon to Qualtrics and filled in the questionnaire on this platform. To assure that the input from Amazon matched with Qualtrics, a validation code was generated by Qualtrics that respondents had to fill in on Amazon at the end of the questionnaire. The result was an Amazon dataset with 209 submitted validation codes and a Qualtrics dataset of 317 finished and unfinished entries. The datasets were prepared and matched using the validation code of Amazon and Qualtrics. All invalid workers and unfinished entries were filtered based on the code.

After filtering the dataset, the sample included 204 observations and 61 variables. Five workers entered an invalid code, the rest were unfinished surveys due to multiple reasons. The variables were imported as numeric variables and were pre-coded via Qualtrics. For the second variable of risk the original discrete values ranged from five to 315 dollars and were re-coded to a scale from zero to 32. The second variable of positive reciprocity was as a discrete variable from zero to 30 dollars with steps of five and was re-coded from zero to six.

After inspection of the columns it was observed that there were a lot of metadata- and irrelevant variables. Some variables had unnecessary complicated names inherited from Qualtrics[[11]](#footnote-11). Others needed recoding because of the staircase questions used to measure the second variable of risk.

The irrelevant metadata variables, the validation code, the attention check questions and the variables gender, nationality and age were removed. Gender, nationality, and age were removed since it was decided to focus purely on the social preferences and risk-attitude and not on demographics. The unnecessary complicated names were renamed. The filter questions of risk were used to calculate the risk score of the second risk variable and were removed afterwards. Since the prisoner’s dilemma and the chicken game were imported as as numeric variables, the variables were transformed to factor variables and the levels of the variables were renamed to their corresponding categorical classes. Finally, all the remaining variables were re-ordered in a logical order. After all pre-processing, the final dataset has 15 variables: five game variables, eight preferences variables and two risk variables.[[12]](#footnote-12)

### Time

The time variable was inspected in Figure 1 via descriptive statistics, a boxplot, and a histogram, in order to find out whether some respondents had extremely fast responses that indicate that they provided random answers. The approximated time for Dutch readers based on the second pilot round is about 10 to 12 minutes (600 to 720 seconds). The median time to finish the questionnaire is 433 seconds, the minimum is 59 seconds and the first quantile is 293 seconds. A valid explanation for this may be that native English speakers read faster and that respondents on Amazon Mechanical Turk are experienced survey takers and therefore able to finish faster. In any case, some highly unlikely times are observed. This might coincide with under zealous effort of respondents and cause extra noise and respondent error in the data. However, to avoid meddling with the data too much they are kept in for further analysis.



Figure 1: Boxplot and histogram of the time variable, with the mean and median line shown in red and blue, respectively.

### Multi-indicator concepts

Whenever possible it was preferred to combine the multi-indicator concepts to provide a more internal reliable and full measure of the concept, but also to avoid collinearity in-between the concepts. This is important because when collinearity is present in linear regression, it can be problematic to figure out how each variable is independently associated with the target variable (James, et al., 2013). The correlations of the preferences show that collinearity might be present.[[13]](#footnote-13) The variables negative reciprocity have remarkably high in-between correlations; the same is true but to a lesser extent for altruism, risk, and positive reciprocity. High correlations are also present between preferences.

The Cronbach’s Alpha was then used to assess internal reliability and to discover whether the underlying concepts could be joined to together. Once for the unscaled input and once after scaling the input to z-scores. The alpha scores showed that only the variables of negative reciprocity were internally reliable, so these were joined together. The remaining negative reciprocity variables were removed and all variables were re-ordered.[[14]](#footnote-14) The other multi-indicator variables, positive reciprocity, altruism, and risk all lacked a sufficient Cronbach’s alpha score even after scaling the variables. It was decided to keep these multi-indicator concepts apart. After the joining the negative reciprocity variables, the correlation plot is inspected in Figure 2.



Figure 2: Spearman’s rang correlation plot of all the predictors: preferences and risk-attitude variables. All correlations are coloured gradually from -1 to +1, negative relations are coloured reddish and positive relationship are coloured blueish.

The figure shows that after joining the variables correlations are still present, but the highest in-between correlations of negative reciprocity are solved. Nonetheless, collinearity might still be an issue.

## Exploratory data analysis

This section has two parts: the game scenarios, and the preferences and risk variables.

### Public goods

The dominating strategy in the public goods game it to invest zero dollars and was described in detail in section 2.5.1. The amount invested in the public goods game is examined in Figure 3 below. Overall, only 40 of the 204 (19,6%) respondents play the dominating strategy of zero dollars. Presumably, altruism, trust and positive reciprocity plays a role in the public goods game. These relations are expected to be positive and might explain higher investments. The mean investment is 4.53, the median and the mode is five dollars, with 59 respondents (28.9%) that invest this amount. Multiple peaks are visible, players that offer all, half, or nothing. Furthermore, the data show a higher tendency in the lower numbers as is seen in the interquartile range of the boxplot and lack a normal distribution caused by the multiple peaks.



Figure 3: Boxplot and bar plot of the investment in the public goods game.

### Ultimatum game: offer

The game theoretical solution for the offer in the ultimatum game is one dollar and was described in detail in section 2.2.1 and 2.5.2. In the empirical findings, shown in Figure 4, only ten respondents (4.9%) offer one dollar. Moreover, the mode is at five dollars with 114 respondents (55.9%) that offer an even distribution. These findings correspond with earlier findings and indicate that offerors do not tend to assume that the responders play in a traditional game theoretical manner and offer more to avoid rejection (Güth & Kocher, 2014). The assumption that players have about another player’s negative reciprocity is not measured and therefore cannot be used to predict the offer. Nonetheless, preferences of the offeror are measured. The height of the offer may positively relate to altruism, positive reciprocity, and negatively with the risk score of the offeror.



Figure 4: Boxplot and bar plot of the offer in the ultimatum game.

Beside the information from above, the mean of the offer is 4.32, the median five dollars and the standard deviation is 1.72. The data is centred to the left and the boxplot in Figure 4 shows that the inter quartile range from three to five dollars. Some outliers are present, namely seven offers of 10 dollars, but all offers greater than five dollars are rare, as are offers of zero dollars.

### Ultimatum game: response

The game theoretical solution for what the responder should minimally accept is one dollar, as was discussed in section 2.2.1 and 2.5.2 . Only 37 respondents (18.1%) accept one dollar and is shown in Figure 5. The mode is five dollars with 69 respondents (33.8%) that demand an even split. As previously mentioned, theory states that negative reciprocity relates with rejections in the ultimatum game and therefore positively with a higher minimum acceptable amount (Camerer & Ho, 2015). Results show that the mean is 3.76, the median 4 dollars and the standard deviation 1.81. Noteworthy is that the data is more centred to the lower values than the offer in the ultimatum game. Again, some outliers are present, in total fourteen respondents demand more than 5 dollars.



Figure 5: Boxplot and bar plot of the minimum acceptable amount in the ultimatum game.

### Prisoner’s dilemma

As discussed in section 2.5.3 the dominating strategy in the prisoner’s dilemma is to defect. Only 66 of all 204 respondents are defectors (32.4%); most respondents choose to co-operate. These findings are in stark contrast with the game theoretical solution and shown in Table 3. Trust, altruism, and positive reciprocity might influence the decision-making in prisoner’s dilemma. This relation is presumably negative; this means that the chance of defecting decreases with higher scores for these preferences.

|  |  |  |
| --- | --- | --- |
|  | Cooperate | Defect |
| Frequency | 138 | 66 |
| Percentage | 67.6% | 32.4% |

Table 3: A table of the frequency and proportion distribution of the prisoner’s dilemma.[[15]](#footnote-15)

### Chicken game

As discussed in section 2.5.4 the mixed Nash equilibrium is to randomly pick co-operate 30 percent, and defect 70 percent of the times. In contrast, of all the respondents 56 chose dare (27.5%), this percentage is a lot lower than the 70 percent of the mixed Nash equilibrium. The results are shown in Table 4. Again, trust, altruism and positive reciprocity might have a negative relationship with the decision in the chicken game. Additionally, risk may have a positive relationship, this means that the chance of picking dare increases with higher risk scores.

|  |  |  |
| --- | --- | --- |
|  | Chicken | Dare |
| Frequency | 148 | 56 |
| Percentage | 72.5% | 27.5% |

Table 4: A table of the frequency and proportion distribution of the chicken game.[[16]](#footnote-16)

### Preferences and risk-attitude[[17]](#footnote-17)

All variables, except for the second variable of positive reciprocity, altruism, and risk are ordinal values but measured as a Likert interval from zero to ten. The second variable of altruism and risk were measured on a ratio scale. The histograms are plotted in Figure 6.



Figure 6: Multiple histograms of all preferences and risk-attitude variables[[18]](#footnote-18).

Negative reciprocity, and the second variable of risk and altruism are right-skewed and have a tendency for lower scores. The opposite is true for the first variable of positive reciprocity and altruism. It also is noteworthy that in first variables of altruism, positive reciprocity, and negative reciprocity the mode corresponds with the most socially desirable option and potentially indicate social desirability bias. None of the variables, except for the second variable of positive reciprocity and trust suggest a normal distribution; this is because of the skewness or the lack of a platykurtic character of the variables. The second variable of altruism has the characteristics of an exponential distribution and has a few large outliers.

## Experimental procedure

### Algorithm types and evaluation criteria

For the public goods game and the ultimatum game a regression model was used, even though a classification model can be used as well. However, since the goal of this research is not only prediction, but also explanation, a regression model is preferred since accuracy is less important than explaining variance. The means squared error was used as the evaluation criterium for the regression models.

The prisoner’s dilemma and chicken game are categorical variables with two options. Therefore, classification models are the only suitable choice. The evaluation criterium for the classification model was accuracy even though the groups were imbalanced but since no class was preferred over the other, accuracy was the best choice.

### Cross-validation

Cross-validation assesses the evaluation criteria and allow for parameter tuning. The data was split in a test and train set, setting apart a test set to measure final performance and to avoid data snooping. The training set was used for cross-validation. The training set observations were randomly assigned a fold id from one to five, using a five-fold cross-validation. This number of folds was chosen to get a large enough validation set and to avoid collinearity of the validation scores (James, et al., 2013). All folds, including the separate test set contained 34 observations. 170 observations were left for training. Cross-validation was preferred over the hold-out method since the data was already scarce and by using cross-validation, more data was used for training. Additionally, cross-validation served as a more reliable measure of the test fold performance and generated standard deviations of these performances. In all methods, except for the boosting method, the same fold id’s were used to compare the results; this is not possible in the boosting method.

### Bivariate analysis

All five-game scenarios and the relations with the preferences were examined via a bivariate analysis on the training data. Most of the preferences show no characteristics of normal distribution, and many of them are ordinal or interval variables. Therefore, as the public goods and ultimatum games are numeric variables, a non-parametric Spearman’s rang correlation test was used. For the prisoner’s dilemma and the chicken game the relations were examined via boxplots that were split by the groups.

### Model building and parameters

Multiple models were constructed for all game scenario’s and are described in more detail below. Since the goal of the research is prediction and understanding of the role of social preferences and risk-attitude, insightful statistical models were preferred that perform well under a smaller sample size.

At first, a base model generated to have a baseline to compare performance against. The regression model used the mean as a baseline. The classification models used the majority class.

Then a variable selection algorithm, called best subset selection was conducted. It used the linear regression model and the R2 score to determine the best variables for the regression models. The logistic regression and AIC statistics were used in case of a classification model. Cross-validation was performed to select the optimal number of variables and to assess performance. The optimal number of variables was afterwards supplied as a parameter of the model and trained on all training data to find out which variables have predictive value.

Another variable selection model was conducted: the lasso model. Lasso shrinks the coefficients of the linear regression or the logistic regression model and uses cross-validation to find the best lambda. The lambda and the shrunk non-negative coefficients were inspected for all training data to see which variables have predictive value.

A benefit of these methods is that they avoid overfitting too complex models. Moreover, these models can examine which variables have predictive and explanatory value and compare whether the predictors are the same for both methods.

Whenever there was an indication of non-linearity, the best subset selection model was extended using non-linear terms. Cross-validation was applied to assess performance of these polynomial extensions. In case of the response in the ultimatum game, a smoothing spline was conducted since there was evidence for non-linearity but polynomial extensions had no effect. Cross-validation was then applied to find the best degrees of freedom for the smoothing spline.

The final model was a non-parametric method, namely a gradient boosting machine, and was chosen to assess performance of a non-parametric method on this data. The tuned parameters were the number of trees, the number of observations in a node, and the interaction depth. The shrinkage factor was fixed at 0.001.

# Results

In the previous section it was found that most respondents do not play in the traditional game-theoretical way. This section focuses on the main goal of this research is to see whether, which and how much social preferences and risk-attitude predict and explain in game theory.

## Bivariate analysis

First, a bivariate correlation plot is examined in Figure 7. It shows the presence of moderate, positive correlations between the investment amount in the public goods game and altruism, and trust. A weaker positive correlation is present between positive reciprocity. Secondly, a moderate, positive correlation is present between the offer in the ultimatum game and the second variable of positive reciprocity. A weaker correlation is present for the first variable of positive reciprocity. Lastly, the minimum acceptable amount in the ultimatum game has a moderate, positive correlation with negative reciprocity.



Figure 7: Spearman’s rang correlation plot of the preferences and the dependent regression variables on the trainings data.

In the bivariate analysis of the prisoner’s dilemma and the chicken game the differences in distribution for both groups are examined via boxplots. The boxplots of the prisoner’s dilemma are found in Figure 8. The boxplots show large differences in the distributions between the co-operate and defect group for: positive reciprocity, altruism, and trust. These are all negative relationships, where there is a lower distribution for the defect group.



Figure 8: Boxplots of all preferences of the prisoner’s dilemma. Split by cooperate and defect, where cooperate corresponds with light blue and defect with red.

The analysis of the chicken game in Figure 9 shows the same differences, but the differences are smaller and more distributed toward each other than in the prisoner’s dilemma. In contrast to the prisoner’s dilemma, there is no difference in distribution of the first positive reciprocity variable. Regarding the possible influence of risk, the box plots show no differences in the distribution of the groups for the risk variables.



Figure 9: Boxplots of all preferences of the chicken game. Split by chicken and dare, where chicken corresponds with light blue and dare with red.

These bivariate relationships may have confounding variables, especially prevalent due to the high correlations between the preferences. This, however, is subdued in the multivariate models. Another issue is that these correlations are not extraordinarily strong, but in terms of noisy psychological concepts these correlations do indicate the nature of the relationship. They may, however, lack strong enough correlations for accurate predictions.

## Modelling

### Public goods

The plot below shows, for all different models, the predictive performance of the investment in the public goods game. These are examined in detail below and shown in Figure 10.



Figure 10: Plot of mean squared error cross-validation scores of the investment in the public goods game for all different models

First the base model was conducted to set a baseline to compare performance against. Thereafter, the best subset selection model was performed. Figure 11 illustrates that the first variable of altruism explains nine percent, together with the second variable of altruism explains more than 12 percent of all the variance. Further variables only provide trivial improvement. Cross-validation confirms these findings and the best model includes two variables: the first- and the second variable of altruism. The predictive performance increases with 11 percent compared to the base model.



Figure 11: R2 scores vs. the number and best variables plots for the public goods game.

The performance of the lasso model is worse than the best subset selection model, but the standard deviation shows that the model is more stable. Figure 12 illustrates that this model includes the same variables as in the previous model. By way of contrast, it also includes trust and it barely includes the second variable of positive reciprocity. The most important preference is altruism followed followed by trust and positive reciprocity. All the relationships are positive. Moreover, the predictive performance of the lasso model is not strong since the lambda plus one standard error still corresponds with the base model.



Figure 12: Scaled coefficients and the shrinkage factor plot of the lasso model of the public goods game. The red dotted line corresponds with the best lambda parameter and the light blue line with the best lambda plus one standard error.

For non-linear extensions, the found model via the best subset selection model was studied. There was evidence for non-linearity of the first variable of altruism. Cross-validation shows that the best performing model has a polynomial to the third degree of altruism and also includes the second variable of altruism. It performs slightly better compared to the best subset selection model and results are more stable too. The bivariate plot in Figure 23 in the Appendix A: Tables and figures shows a better fit as well.

The boosting method performance worse than the other models, nonetheless the relative importance table confirms the findings that the first variable of altruism is the most important variable, followed by the second variable of altruism in the public goods game, explaining 42 and 20 percent respectively, trust has a relative importance of ten percent.

The results are summarized in Table 5, and the best performing model is the model that includes a polynomial to the third degree of the first variable of altruism and includes the second variable of altruism. The polynomial model improves with 13.3 percent compared to the base model and has a mean squared error of 9.86 and a root mean squared error of 3.14. For the final prediction on the test set the model is trained on all trainings data. The base model has a mean squared error of 15.15 on the test set. The polynomial model performs better and has a mean squared error of 12.43 and performs 18 percent better than the base model.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Base model | Best subset | Lasso | Polynomial | Boosting |
| MSE | 11.37 | 10.13 | 10.28 | 9.86 | 10.42 |
| Percental improvement\* | 0% | 10.9% | 9.6% | 13.3% | 8.4% |
| RMSE | 3.37 | 3.18 | 3.21 | 3.14 | 3.22 |
| Standard deviation | 2.57 | 2.36 | 1.15 | 2.25 | NA |

Table 5: Summarized results of the performance in the public goods game.

\*compared to base model when measured in mean squared error.

The polynomial model was then trained on all available data, this model explains more than 16% of all the variance. The first variable of altruism is extremely significant (<0.001) and the second variable of altruism is very strongly significant (0.009). The second-degree polynomial is also significant (0.04) and the third degree is weakly significant (0.09). All coefficients are positive, indicating positive relationships. The output that contains the coefficients and the p-values of the final model can be found in Figure 24 of the Appendix A: Tables and figures. Furthermore, diagnostics show no oddities and no presence of multicollinearity.

### Ultimatum game: offer

The plot below shows, for all different models, the predictive performance of the offer in the ultimatum game. These are examined in detail below and shown in Figure 13.



Figure 13: Plot of mean squared error cross-validation scores of the offers in the ultimatum game for all different models

After conducting the base line model, the best subset selection model was performed. Figure 14 shows that the second variable of positive reciprocity explains more than six percent of all the variance and that adding an extra variable is inefficient. Cross-validation confirms these findings and the best performing model has one variable, the second variable of positive reciprocity. The performance increases with 5 percent compared to the base model, the corresponding coefficient is 0.26 and the variable is strongly significant (<0.001). There was no evidence for a non-linear relationship.



Figure 14: R2 scores vs. the number and best variables plots for the offer in the ultimatum game.

The performance of the lasso model is slightly worse but is more stable than the best subset selection method. Both models include the same variable, the second variable of positive reciprocity. The predictive performance is again weak, since the one standard error rule includes the base model.

The final model, the gradient boosting machine is the best performing model. The optimal settings are 710 trees, an interaction depth of 2, and a minimum number of observations of 20. As in all earlier models, the most important variable is again the second variable of positive reciprocity and is followed in importance by the second variable of altruism.

To summarize, Table 6 shows all results. The best model, boosting, is more complex but performs 8.5 percent better than the base model. The mean squared error of the boosting method is 3.03, and the root mean squared error is 1.74. On the final test set the base model has a mean squared error of 1.56. The boosting method performs 3 percent better and has a mean squared error of 1.52.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Base | Best subset | Lasso | Boosting |
| MSE | 3.30 | 3.13 | 3.17 | 3.02 |
| Percental improvement | 0% | 5.2% | 3.9% | 8.5% |
| RMSE | 1.82 | 1.77 | 1.78 | 1.74 |
| Standard deviation | 1.71 | 1.80 | 0.79 | NA |

Table 6: Summarized results of the performance for the offer in the ultimatum game.

The boosting model was then trained on all available data and shows that the relative influence of the second variable of positive reciprocity is 61 percent and 20 percent for the second variable of altruism. The bivariate plots, displayed in Figure 25 in Appendix A: Tables and figures, shows that positive reciprocity and altruism have a positive relationship with the offer in the ultimatum game.

### Ultimatum game: response

The plot below shows, for all the different models, the predictive performance of the minimum acceptable amount in the ultimatum game. These are examined in detail below and is shown in Figure 15.



Figure 15: Plot of mean squared error cross-validation scores of the minimum acceptable amount in the ultimatum game for all different models

The best subset selection model was performed. Figure 16 reveals that negative reciprocity explains more than 12% of all the variance and that adding another variable is not useful. Cross-validation also includes negative reciprocity and performs 11 percent better than the base model. The corresponding coefficient is 0.26 and the variable is strongly significant (<0.001).



Figure 16: R2 scores vs. the number and best variables plots for the minimum acceptable offer in the ultimatum game.

The performance of the lasso model is worse than the best subset selection, but more stable. Both variable selection models find the same variable, negative reciprocity, but as in all models the one standard error rule indicates a weak model.

Non-linear relations extensions of the best subset selection model were examined. Adding a second-degree term of negative reciprocity is significant (0.016).[[19]](#footnote-19) In contrast, cross-validation of these extensions were not useful and performed worse than the best subset selection model, therefore a smoothing spline was conducted. Cross-validation shows that three degrees of freedom provides the best fit. This also improved the performance compared to the best subset selection model.

The final boosting model performs better than the other models. The optimal settings are 1400 trees, an interaction depth of 1, and the number of minimal observations in a node is set to 10. The most important variable is the same found in all other methods, negative reciprocity.

Table 7 summarizes the results. The best performing model, boosting, performs 13.8 percent better than the base model. It has a mean squared error of 3.06 and a root mean squared error of 1.75. On the final test set, the base model has a mean squared error of 2.15. The boosting model a mean squared error of 1.98 and improves with 7.9% compared to the base model.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Base | Best subset | Lasso | Smoothing spline | Boosting |
| MSE | 3.55 | 3.18 | 3.30 | 3.13 | 3.06 |
| Percental improvement | 0% | 10.4% | 7% | 11.8% | 13.8% |
| RMSE | 1.88 | 1.78 | 1.82 | 1.77 | 1.75 |
| Standard error | 0.87 | 0.45 | 0.27 | 0.57 | NA |

Table 7: Summarized results of the performance for the minimum acceptable offer in the ultimatum game.

The boosting model was then trained on all available data. The model shows that the relative influence of negative reciprocity is 78%. The bivariate plot in Figure 27 (Appendix A: Tables and figures), shows a complex, positive relationship with the minimum acceptable offer. The predicted value of the minimum acceptable offer stays stable until the negative reciprocity score of 4, and then increases strongly until 6, after which it stays stable again.

### Prisoner’s dilemma

The plot below shows, for all different models, the predictive performance in the prisoner’s dilemma. These are examined in detail below (see Figure 17).



Figure 17: Plot of cross-validation accuracies of the prisoner’s dilemma for all different models

The base model has an accuracy of 69%. A best subset selection method was performed. This model improves performance. The best performing model has three variables: the first variable of altruism, trust, and negative reciprocity. The coefficients show a negative relationship for all variables. Altruism is strongly significant (<0.001). Trust is moderately significant (0.003). Negative reciprocity is weakly significant (0.063).

The lasso model performs slightly worse than the best subset selection method, but results are a lot more stable, the standard deviation decreased with 7.4 percentage points. Moreover, the lasso model includes four variables, and all have a negative relationship. The four variables are the same found in the best subset selection model, but the second variable of positive reciprocity is added. Figure 18 illustrates that altruism is the most important variable, followed trust and negative reciprocity. The least important variable is the second variable of positive reciprocity.



Figure 18: Scaled coefficients and the shrinkage factor plot of the lasso model of dilemma.

Boosting performs slightly worse than the other models. There is no particular important variable, since none on the variables have a relative importance of more than 20 percent. The most important variables are the first variable of altruism, and negative reciprocity. Both have an individual relative importance of 18 percent.

Table 8 summarizes all results. The best performing model is the best subset selection method. However, the lasso model performs just slightly worse, but results are a lot more stable. The accuracy of this model is 72.9 percent compared to 74.1 percent in the best subset selection method. The standard deviation is 4.3 percent compared to 11.7 percent in the best subset selection model and this is the reason that the lasso model is preferred. Compared to the base model, the lasso model improves with 5.5 percent compared and 3.5 percent in percentage points. On the final test set, the base model has an accuracy of 58 percent. The lasso model performs 10 percent better and 6 percent in percentage points and has an accuracy of 64 percent.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Base | Best subset | Lasso | Boosting |
| MSE | 0.694 | 0.741 | 0.729 | 0.724 |
| Percental improvement | 0% | 7.2% | 5.5% | 4.8% |
| Improvement in percentage points | 0% | 4.7% | 3.5% | 3% |
| Standard deviation | 9.9% | 11.7% | 4.3% | NA |

Table 8: Summarized results of the performance for the prisoner’s dilemma

The final lasso model was trained on all available data. The model included the variables: the first variable of altruism, trust, and negative reciprocity. However, contrary to the training set positive reciprocity was not included. In the final model the log coefficients of altruism and trust are around 0.16 and the coefficient of negative reciprocity is 0.04. This shows that in this model the variable of altruism and trust are four times as important as negative reciprocity.

### Chicken game

In the chicken game the base model has an accuracy of 72.4 percent. The best subset selection barely improves on this model with an accuracy of 73.5 percent. The model has three variables: the second variable of altruism and risk, and trust. Only trust is weakly significant (0.083), the second variable of altruism (0.117) and risk (0.126) are not significant. Risk has negative relationship in this model.

The best lambda of the lasso model contains zero variables and is equal to the base model; it is therefore evidence for a weak model. Boosting was performed but accuracy was not improved; Figure 19 emphasizes that only noise is learned.



Figure 19: Accuracy vs. number of trees plot for the chicken game.

All games, except for best subset selection do not perform better than the base model. The improvement of the best subset selection method is neglectable and shows that there is no compelling evidence that this model is substantially better than the base model. On the final test the accuracy of the base model is 70.6 percent. The accuracy of the best subset selection model is 73.5 percent and improves with 2.9 percentage points. The final model shows no significant relation for any variables and trust is no longer weakly significant.

# Discussion

## Discussing and answering the research questions

The goal of this study is to answer the four research questions of this research. The first question was whether the decisions in the game scenarios deviate from the game theoretical solution? Secondly, which and how do the social preferences relate with the decisions in the game scenarios? Thirdly, how much explanatory and predictive value do social preferences and risk-attitude have? Finally, are the concepts generalizable across the games? These questions have been partly answered in previous sections and are discussed in more detail below.

### Deviations

In all the game scenarios, the game theoretical solutions generally do not correspond to the decisions of the respondents. In the public good games, most people tend to invest more than zero dollars. 19.6 percent of all players invest exactly zero, the mode is at 5 dollars and there was also a peak at ten dollars. This indicates that co-operation in the public goods game does exist, which is unexpected from a game theoretical perspective. It does, however, correspond to previous findings (Camerer & Ho, 2015).

In the ultimatum game people tend to offer more than one dollar and therefore it can be reasoned that offerors do not expect responders to play game theoretically. This is a correct assumption, since the minimum acceptable amount for most responders is larger than one dollar. This corresponds to earlier research (Guth, et al., 1983; Forsythe, et al., 1994). Furthermore, the distribution of the minimum acceptable offer was more centred to the lower numbers than the actual offers, showing some inefficiency in the ultimatum game.

In the prisoner’s dilemma the dominating strategy is to always defect; however, 67.7 percent of all the respondents pick co-operate. This also the case for the chicken game, where 72.5 percent of all the respondents choose dare but in a mixed Nash equilibrium only 30% should do so.

### Predicting and explaining

The results of the public goods game showed that the optimal model has moderate predictive value. The best performing polynomial model reduced 13.3 percent of all the variance of the investment in the public goods game. The model also provided explanatory value. The bivariate analysis showed relations between both variables of altruism, trust and the second variable of positive reciprocity. The findings correspond with the variables found in the lasso model. In contrast, the best subset selection model excluded the variables of positive reciprocity and trust but had strong significant relations with the altruism variables. Additionally, the highest coefficients of the lasso model and the most important variables in boosting were also the altruism variables.

Overall, the results indicate that altruism plays a role in the decision-making of the investment in the public goods game. Moreover, all models showed a positive relationship with altruism and the investment. Thus, people that score higher on altruism, are more likely to invest a higher amount in the public fund. In contrast, the relations between trust and positive reciprocity is more ambiguous and could be a confounding relationship.

There is weak predictive value for the offer in the ultimatum game. Boosting reduced the variance with 8.5 percent. The probable cause for this is that the main influence for decision-making in the ultimatum game is the assumption of how an offeror expect a responder to behave considering an unfair offer. This was, however, not measured (Güth & Kocher, 2014).

Nonetheless, there was one social preference that had some predictive and explanatory value: positive reciprocity. In the bivariate analysis and all models, positive reciprocity was the strongest predictor and showed very strong significance. This is a sign of the robustness of this preference. All models showed a positive relationship with the offer in the ultimatum game, where a higher positive reciprocity score relates with a higher offer.

This shows that some people are willing to offer a generous amount. However, it is unlikely that this relation is causal, since there is no opportunity for the responder to show behaviour that should be rewarded. It could be that this generous offer is a form conditional co-operation as was described in (Camerer & Ho, 2015). Noteworthy is that only the second variable of positive reciprocity is included in the models and not the first variable; however, this might be because of collinearity.

Additionally, there is moderate predictive value for the minimum acceptable amount in the ultimatum game. The boosting model reduced the variance with 13.8 percent. The explanatory value is also quite robust, since all models and the bivariate analysis unanimously indicate a strongly significant, positive relationship with negative reciprocity. Thus, a higher negative reciprocity score leads to a higher minimum acceptable amount and increase the chance of rejection. This finding corresponds to previous research (Camerer & Ho, 2015; Güth & Kocher, 2014).

In the prisoner’s dilemma, the lasso model has moderate predictive value. Additionally, there is some explanatory value for the preferences, but results are mixed. All models found positive relationships with the first variable of altruism, trust, and negative reciprocity. The relationship with negative reciprocity is counterintuitive and unexpected. A negative coefficient indicates that the chance of defecting *decreases* when respondents have a higher negative reciprocity score. However, this might be just an indication of the general care of fairness of a respondent. In contrast, there was also no sign of this relationship in the bivariate analysis. It had a relative small coefficient in the lasso model compared to altruism and trust. It was also weakly significant in the logistic regression. Therefore, the influence of negative reciprocity should be interpreted with care. Nonetheless, the variables of trust and altruism show explanatory and moderate predictive value. These preferences have a negative relationship; respondents with lower scores for trust and altruism have a higher chance of defecting.

In contrast with earlier game scenarios, the models for the chicken game had *no* predictive or explanatory value. Indicating that social preferences do not influence decision-making in the chicken game. This might be an artefact of the questionnaire where an abstract, but perhaps a too complicated representation of the game was created. Alternatively, there simply might not be any actual relationships.

### General findings and predictive limitations

In general, for some games scenarios social preferences have decent, but not ground-breaking, predictive value using these statistical learning models. A probable cause for this is that these social preferences do not measure all factors that influence decision-making, for instance: cognitive limitations, strategic sophistication, and economic circumstances may play a role in decision-making. It is also probable that the noise in the dataset, and the psychological nature of the data, prevents better predictions on this dataset.

The explanatory value and understanding of the social preferences that influence decision-making are, in general, useful and some robust and significant relations are found. However, results are mixed and the variables that influence decision-making are dependent on the game scenario. Risk-attitude has no predictive or explanatory value and was not included in any of the models. Since the concepts, predictive value, and the best statistical learning technique is dependent on the game scenario a flexible modelling approach is needed.

## Limitations

Some limitations of above findings and results are discussed and acknowledged. First, the use of Amazon Mechanical Turk provides a convenience sample and does not follow the principles of a random sample. Hence, the results cannot be generalized beyond the sample. Another issue is that online research and the use of Amazon Mechanical Turk coincides with worries about the data quality. This was subdued as much as possible using qualifications and a carefully designed questionnaire, but respondent error bias might still be present as some extremely fast completion times were found. Furthermore, the use of a questionnaire without real monetary pay-offs has implication for the ecological- and measurement validity of this study, since it may adjust decision-making (Gintis, 2005). Moreover, social desirability bias might be an issue, especially since social preferences are sensitive to this bias (Bryman & Bell, 2011). Inspection of the preferences showed the possibility of this bias. These above-mentioned biases might contaminate and influence decision-making and should be taken in consideration when reviewing the results.

## Future research

The statistical learning techniques used in this research, that avoid overfitting and incorporate the heterogeneity of the social preferences, are rare and novel in the field of behavioural game theory (Wright & Leyton-Brown, 2017). These methods are useful to understand human decision-making in game theory. However, if accurate predictions are the goal, different, new methods or cleaner data might be more suitable. On the other hand, future research may use these novel methods for different game scenarios or apply it to other fields of human decision-making in order to gain a better understanding of the influence of social preferences and risk-attitude. Furthermore, the models could be extended by capturing cognitive limitations, level-k reasoning, and strategic reasoning of the respondent. This is useful since rationality assumed in game theory is generally too strong. It has to be adjusted to more realistic assumptions that incorporate cognitive limitations to provide more accurate predictions of decisions in game theory (Brian Arthur, 1994; Gill & Prowse, 2016). Future research could also incorporate economic circumstances and other predictors that may influence decision-making. Moreover, it could try to collect cleaner data by using lab experiments and cash as pay-offs. This would improve predictions since it reduces ecological validity and noise. The reason for this is that respondents are likely more serious and focused in these circumstances and therefore give a better reflection of actual behaviour in these scenarios.

# Conclusion

In this section all research question that were introduced in the Introduction are restated and briefly answered.

1. *Do the findings in the game scenarios deviate from the game theoretical solution?*

No, in all game scenarios the game theoretical solution was generally not played by the respondents, as found in previous research (Camerer & Ho, 2015; Crawford, 2002). People are more co-operative in the public goods game, the prisoner’s dilemma, and the chicken game. In the ultimatum game most responders demand a fairer share and the offerors offer that as well. This was extensively addressed in the chapter describing the Methods in the section Exploratory data analysis.

1. *Which and how do social preferences and risk-attitude relate with the decisions in the game scenarios?*

This question is answered in the Results: in the section Bivariate analysis and Modelling. The social preference altruism plays a role in the decision-making for the investment in the public goods game; this relationship is a positive. The social preference positive reciprocity relates positively with the offer in the ultimatum game. Additionally, as found in earlier research, the social preference negative reciprocity influences the minimum acceptable amount in the ultimatum game (Camerer & Ho, 2015; Guth, et al., 1983). Furthermore, the social preferences: altruism, trust and possibly negative reciprocity relate negatively with the decision-making in the prisoner’s dilemma. In other words, respondents with lower scores for these social preferences are more likely to defect. Finally, in the chicken game there were no relationships found.

1. *How much explanatory and predictive value do social preferences and risk-attitude have?*

The third research question is answered in the Bivariate analysis and Modelling section of the Results chapter and discussed in the Discussions. In general, the predictive value was moderate to weak and dependent on the game scenario. The social preferences have descent predictive value for the public goods game, the minimum acceptable offer in the ultimatum game and the prisoner’s dilemma. This was weaker for the offer in the ultimatum game. For the chicken game the preferences had no predictive value at all. These results show that in general, the social preferences have some predictive value, but that risk-attitude does not have any predictive value in these game scenarios. The social preferences are of substantial value to help explain and understand decision-making in the game scenario’s, robust findings and significant relationships were found, but some results were mixed and reliant on the games. Nonetheless, this is the main strength of the used models.

1. *Are the concepts and models generalizable across the games?*

As mentioned above and in the chapter Discussions, the predictive and explanatory value is dependent on the game scenario. The best predictive algorithm is also dependent on the game scenario. As are the relevant concepts that influence decision-making. This dependency shows that a flexible approach, such as statistical learning techniques is necessary in the understanding and prediction of decision-making in game theory.

To summarize, this research provides a novel methodology in the field of behavioural game theory and is useful for determining what and how variables relate with decision-making. Additionally, this methodology applies to other fields of decision-making whenever inference is more important than predictions. The actual predictive- and explanatory value of these models depend on the noise in the data and the strength of the relations.

# Bibliography

Axelrod, R., 1984. *The evolution of cooperation.* Revised ed. s.l.:Basic Books.

Bardsley, N., 2007. Dictator game giving: altruism or artefact?. *Economic Science Association,* pp. 122 - 133.

Berg, J., Dickhaut, J. & McCabe, K., 1995. Trust, Reciprocity, and Social History. *Games and Economic Behavior,* Issue 10, pp. 122-142.

Bolton, G. E. & Ockenfels, A., 2000. A Theory of Equity, Reciprocity, and Competition. *The American Economic Review,* Volume 90, pp. 166-193.

Brian Arthur, W., 1994. Inductive Reasoning and Bounded Rationality. *The American Economic Review,* 84(2), pp. 406-411.

Bryman, A. & Bell, A., 2011. *Business Research Methods.* 3rd ed. New York: Oxford University Press Inc.

Camerer, C. F., 1997. Progress in Behavioral Game Theory. *Journal of Economic Perspectives,* pp. 167-188.

Camerer, C. F., 2003. *Behavioral Game Theory: Experiments in Strategic Interaction.* s.l.:Princeton University Press.

Camerer, C. F. & Ho, T., 2015. Behavorial Game Theory Experiments and Modeling. In: *Handbook of Game Theory with Economic Applications, vol. 4.* s.l.:North Holland, pp. 517-573.

Colman, M. A., 2003. Cooperation, psychological game theory, and limitations of rationality in social interaction. *Behavioral and brain sciences,* Volume 26, pp. 139-198.

Crawford, V. P., 1997. Theory and experiment in the analysis of strategic interaction. In: *Econometric Society Monographs 26.* s.l.:Cambridge, pp. 206-242.

Crawford, V. P., 2002. Introduction to Experimental Game Theory. *Journal of Economic Theory,* Volume 104, pp. 1-15.

Dohmen, T. et al., 2011. Individual risk attitudes: measurement, determinants and behavioral consequences. *Journal of the European Economic Association,* pp. 522-550.

Engel, C., 2011. Dictator games: a meta study. *Experimental Economics,* Issue 14, pp. 583-610.

Falk, A. et al., 2018. Global evidence on economic preferences. *Quarterly Journal of Economics ,* Issue (forthcoming).

Falk, A. et al., 2016. *The Preference Survey Module: A Validated Instrument for Measuring Risk, Time and Social Preferences.* Bonn, Forschungsinstitut zur Zukunft der Arbeit.

Fehr, E. & Schmidt, K. M., 1999. A theory of fairness, competition and cooperation. *The Quarterly Journal of Economics,* Volume 3, pp. 817-868.

Forsythe, R., Horowitz, J. L., Savin, N. E. & Sefton, M., 1994. Fairness in Simple Bargaining Experiments. *Games and Economic Behavior,* Volume 6, pp. 347-369.

Gill, D. & Prowse, V., 2016. Cognitive Ability, Character Skills, and Learning to Play Equilibrium: A Level-k Analysis. *Journal of Political Economy,* Volume 124, pp. 1619-1676.

Gintis, H., 2005. Behavioral Game Theory and Contempory Economic Theory. *Analyse & Kritik.*

Güth, W. & Kocher, M. G., 2014. More than thirty years of ultimatum bargaining experiments: Motives, variations and a survey of recent literature. *Journal of Economic Behavior & Organization,* Issue 108, pp. 396-409.

Guth, W., Schmittberger, R. & Schwarze, B., 1983. An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization,* Volume 367-388, p. 1982.

Houser, D., Schunk, D. & Winter, J., 2010. Distinguishing trust from risk: An anatomy of the investment game. *Journal of Economic Behavior & Organization,* Issue 74, pp. 72-81.

James, G., Witten, D. H. T. & Tibshirani, R., 2013. *An Introduction to Statistical Learning: with application in R.* 8th printing ed. New York: Springer Science + Business Media.

Johnson, D. N. & Mislin, A. A., 2011. Trust games: A meta-analysis. *Journal of Economic Psychology,* Issue 32, pp. 865-889.

Levitt, S. D. & List, J. A., 2007. What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?. *Journal of Economic Perspectives,* 21(2), pp. 153-174.

Rabin, M., 1993. Incorporating Fairness into Game Theory and Economics. *The American Economic Review,* Volume 83, pp. 1281-1302.

Wright, J. R. & Leyton-Brown, K., 2017. Predicting human behavior in unrepeated, simultaneous-move games. *Games and Economic Behavior,* Issue 106, pp. 16-37.

# Appendix A: Tables and figures



Figure 20: Spearman’s rang correlation plot of all the predictors: preferences and risk-attitude variables. All correlations are coloured gradually from -1 to +1, negative relations are coloured reddish and positive relationship are coloured blueish.



Figure 21: Bar plot of prisoner’s dilemma, co-operate is coloured light blue and defect in coloured red.



Figure 22: Bar plot of the chicken game, chicken is coloured light blue and dare is coloured red.



Figure 23: Bivariate plot of the investment of public goods and the first variable of altruism. The red line corresponds with the linear fit and the blue line with the extended model with a polynomial term of the third degree[[20]](#footnote-20)

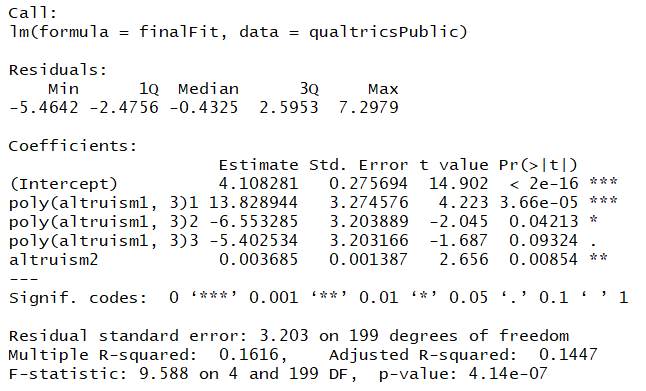


Figure 24: Output of the polynomial model for the public goods game trained on all data.



Figure 25: Bivariate plots of the second variables of positive reciprocity and altruism for the offer in the ultimatum game.



Figure 26: Bivariate plot of the minimum acceptable amount and negative reciprocity. The red line corresponds with the linear fit and the blue line with the smoothing spline with three degrees of freedom.[[21]](#footnote-21)



Figure 27: Bivariate plot of the second variables of negative reciprocity for the minimum acceptable amount in the ultimatum game.

# Appendix B: Questionnaire

## Information sheet

Thank you for considering to partake in this study. This information sheet explains what the study is about and how I would like you to take part in it.

The purpose of the study is to understand human decision-making in simple games. The study is conducted for a research graduation project at Tilburg University. If you agree to take part, a survey must be completed that provides all the necessary data for the research project. The survey will take around 12 minutes.

Please note that this survey will be best displayed on a laptop or desktop computer. Some features may be less compatible for use on a mobile device.

Select the link below to complete the survey[[22]](#footnote-22). At the end of the survey, you will receive a code to paste into the box below to receive credit for taking the survey[[23]](#footnote-23).

*Make sure to leave this window open as you complete the survey***.** When you are finished, you will return to this page to paste the code into the box.

If you have any questions or concerns about the survey, please do not hesitate to contact me.

Jeroen van Buren

Tilburg University

j.vanburen@tilburguniversity.edu

## Informed consent form

* I, the undersigned, have read and understood the study information sheet provided.
* I understand that my answers will be anonymouslyrecorded and used for a scientific graduation project at Tilburg University.
* I understand that the thesis will be presented in public and made available online.
* I understand that all relevant, anonymized research data and code will be open-sourced and made available online.
* I understand that I do not have to take part in this research project. If I agree to take part I can withdraw my participation at any time. However, the work will be considered as incomplete.
* I understand that my work may be rejected by the requester if the HIT was incomplete, not completed correctly or instructions were not followed.
* I have been given adequate time to consider my decision and I agree to partake in the study.
* I agree to all the above statements and hereby consent to the study.
* I do not consent, I do not wish to take part.

## Structure and explanation

Thank you for agreeing to take part in this study.

Firstly, an instruction will follow, explaining the set-up and rules of the games. Followed by, a section consisting out of six different games. In the section thereafter, multiple questions about your preferences will be asked. To finish, three basic demographic questions will be asked.

Please take sufficient time to finish the survey and make sure that you fully understand the question before answering. The survey has *no* 'back button'*,* so earlier answers cannot be changed. Note that the quality of your answers will partly determine the quality and outcome of this study.

For all questions apply that the proper response must be ticked in, or if appropriate filled in, and that *only one answer is possible*. Furthermore, it should be noted that there is *no specific correct answer*.

Thank you for your understanding and good luck with filling in the survey.

## Game scenarios

This section consists of six game scenarios and starts with an explanation of the rules.

**Important rules**

In all games apply the same important rules unless a deviation is mentioned specifically. You will be randomly matched afterwards to play this game against another respondent, which will be a different respondent every game. There is no way of knowing the other player’s decision in advance, however you might have a belief about the other player’s choice and take this into consideration in your own decision. These rules apply to you as well as the other player. All other information in these games (e.g. the description and amount of dollars) are known to both players.

Please, carefully read the explanation of the games and make sure that you fully understand the game before deciding and remember that there is no specific correct answer.

Please only continue if you understand the rules.

### Prisoner’s dilemma

*Consider the following game:*

Imagine that you and another player are both sitting in a different room and must independently decide on sharing an amount of money. You can decide whether you want to *cooperate* or *not cooperate* in sharing an amount of money. You and the other player must both make this decision without knowing what the other will do. The following rules apply:

1. If you and the other player *cooperate*, both get 12 dollars.
2. If you and the other player do *not cooperate*, both get 6 dollars.
3. If *a* player (you or the other) *cooperates* and the other does *not cooperate*, the player that does *not cooperate* receives 20 dollars and the *co-operator* gets nothing.

Will you cooperate or not cooperate?

* Cooperate
* Not cooperate

### Chicken game

*Consider the following game:*

Imagine in this game that you and the other player must anonymously decide whether to pick pot 1 or pot 2. Both pots contain 20 dollars and must be distributed based on your decisions. Your best decision is dependent on the choice of the other player. Furthermore, if both players pick pot 2, the money will not be distributed. The following rules apply:

*If you pick* ***pot 1****:*

* If you and the other player pick pot 1, both share evenly and get 10 dollars. However, if the other player picks pot 2, you will get 3 dollars and the other player gets 17 dollars.

*If you pick* ***pot 2****:*

* If you and the other player pick pot 2, both players get nothing. However, if the other player picks pot 1, you will receive 17 dollars and the other player gets 3 dollars.

Will you pick pot 1 or pot 2?

* Pot 1
* Pot 2

### Ultimatum game: offer

*Consider the following game:*

Imagine that you and the other respondent have been allocated a sum of 10 dollars in total. In this game you are the offeror and the other player is the responder. You, as offeror can offer some, all, or none of these 10 dollars to the other player. The other player can either *accept or reject the offer*. The following rules apply:

1. If the other player accepts the offer, the deal will stand and the amount that you will receive is 10 dollars minus the amount offered, the other player will keep the offered amount.
2. If the other player rejects the offer, the deal will not stand and both players will receive zero.

*An example of a scenario if the other player* ***accepts****:*

*Example 1*: You offer 3 dollars to the other player. The other player accepts and receives 3 dollars, you get 7 dollars.

*An example of a scenario if the other player* ***rejects****:*

*Example 2*: You offer 3 dollars to the other player. The other player rejects, both players receive nothing.

What is the amount of money that you will offer the other player?

* 0 dollars
* 1 dollar
* 2 dollars
* 3 dollars
* 4 dollars
* 5 dollars
* 6 dollars
* 7 dollars
* 8 dollars
* 9 dollars
* 10 dollars

### Ultimatum game: response

*Consider the following game\*:*

*\* Note that the set-up and rules are the same as in the last game, but now you are the responder instead of the offeror and you might pick a different amount than in the last question, or not.*

Imagine that you and the other respondent have been allocated a sum of 10 dollars in total. In this game you are the responder and the other player is the offeror. *You, as responder can either accept or reject* the offer made by the other player, the same rules apply in terms of accepting and rejecting the offer and if rejected, you will get nothing.

What does the other person have to minimally offer you for you to be willing to *accept* the proposal?

* 0 dollars
* 1 dollar
* 2 dollars
* 3 dollars
* 4 dollars
* 5 dollars
* 6 dollars
* 7 dollars
* 8 dollars
* 9 dollars
* 10 dollars

***Important deviation on the rules***

In the following two games you will be randomly matched to play this game against ninety-nine other respondents, instead of one.

### Public goods

*Consider the following game:*

In this game you and ninety-nine other respondents must independently and anonymously decide on how much money to contribute to a pot on the table (up to 10 dollars). Each dollar contributed to the pot will *double in amount*. *You will keep all the money that you have not put in the public pot*. After all players have made the decision the pot will be opened, and you will receive an equal share (1/100) of all the money in the pot, *regardless of your individual contribution.*

*Some examples of scenarios in the game:*

*Example 1*: You contribute zero dollars to the public pot; no money is then doubled. You will keep the rest, which is 10 dollars for certain and share evenly in the amount that is put in the public pot by you and the other players.

*Example 2*: You contribute 10 dollars to the public pot, which is then doubled to 20 dollars. You will keep the rest, which is zero dollars for certain and share evenly in the amount that is put in the public pot by you and the other players.

How much money will you contribute to the public pot?

* 0 dollars
* 1 dollar
* 2 dollars
* 3 dollars
* 4 dollars
* 5 dollars
* 6 dollars
* 7 dollars
* 8 dollars
* 9 dollars
* 10 dollars

### Investment game[[24]](#footnote-24)

*Consider the following game:*

You must decide on how much money to invest in an investment fund (up to 10 dollars). Each dollar contributed, will be invested in the investment fund and you keep all the money that you have not invested. Please invest 6 dollars in the investment fund, this is a check whether you are following instructions correctly and carefully. After all players have made the decision the pot will be opened, and you will receive an equal share (1/100) of all the in the investment fund, *regardless of your individual contribution.*

 How much money will you invest in the investment fund?

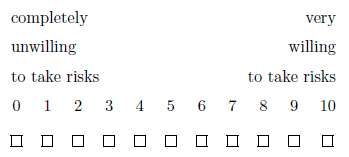
* 0 dollars
* 1 dollar
* 2 dollars
* 3 dollars
* 4 dollars
* 5 dollars
* 6 dollars
* 7 dollars
* 8 dollars
* 9 dollars
* 10 dollars

## The preference survey module[[25]](#footnote-25)

This section of the survey consists out of some brief questions. Note that in these questions you are *not* playing games, against other respondents, anymore.

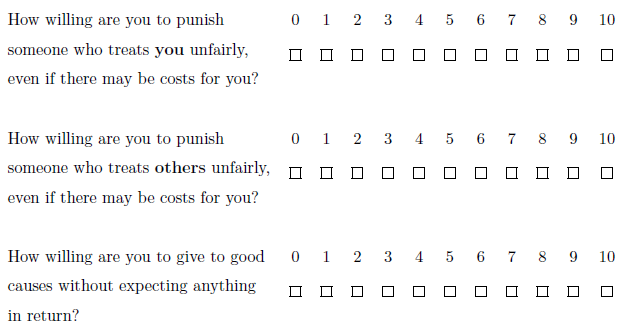
1. Please tell me, in general, how willing or unwilling you are to take risks

*Please use a scale from 0 to 10, where 0 means you are "completely unwilling to take risks" and a 10 means you are "very willing to take risks". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.*



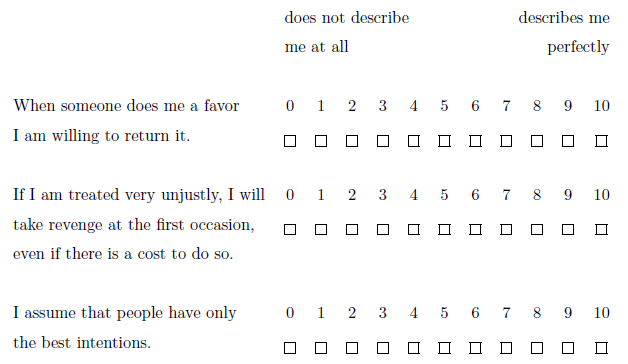
1. We now ask for your willingness to act in a certain way in four different areas.

*Please again indicate your answer on a scale from 0 to 10, where 0 means you are "completely unwilling to do so" and a 10 means you are "very willing to do so". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.*



1. How well do the following statements describe you as a person?

*Please indicate your answer on a scale from 0 to 10. A 0 means "does not describe me at all" and a 10 means "describes me perfectly". You can also use any numbers between 0 and 10 to indicate where you fall on the scale, like 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.*



1. Please imagine the following situation: You can choose between a sure payment

of a particular amount of money, or a draw, where you would have an equal

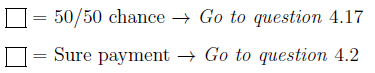
chance of getting 300 dollars or getting nothing. We will present to you five

different situations.

* 1. What would you prefer: a draw with a 50 percent chance of receiving

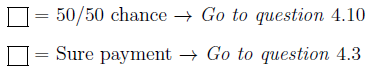
300 dollars, and the same 50 percent chance of receiving nothing, or the

amount of 160 dollars as a sure payment?



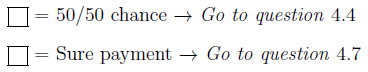
* 1. *(To be answered if replied Sure payment to question 4.1)[[26]](#footnote-26)*

Would you prefer the 50/50 chance or the amount of 80 dollars as sure payment?



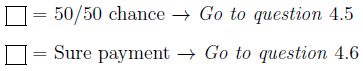
* 1. *(To be answered if replied Sure payment to question 4.2)*

Would you prefer the 50/50 chance or the amount of 40 dollars as sure payment?



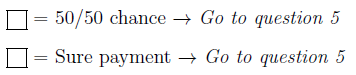
* 1. *(To be answered if replied 50/50 chance to question 4.3)*

Would you prefer the 50/50 chance or the amount of 60 dollars as sure payment?



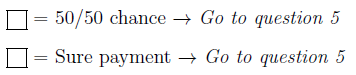
* 1. *(To be answered if replied 50/50 chance to question 4.4)*

Would you prefer the 50/50 chance or the amount of 70 dollars as sure payment?



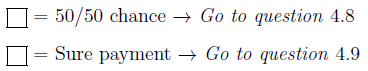
* 1. *(To be answered if replied Sure payment to question 4.4)*

Would you prefer the 50/50 chance or the amount of 50 dollars as sure payment?



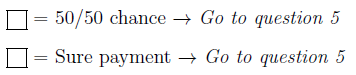
* 1. *(To be answered if replied Sure payment to question 4.3)*

Would you prefer the 50/50 chance or the amount of 20 dollars as sure payment?



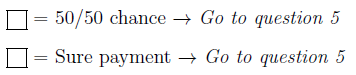
* 1. *(To be answered if replied 50/50 chance to question 4.7)*

Would you prefer the 50/50 chance or the amount of 30 dollars as sure payment?



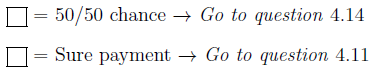
* 1. *(To be answered if replied Sure payment to question 4.7)*

Would you prefer the 50/50 chance or the amount of 10 dollars as sure payment?



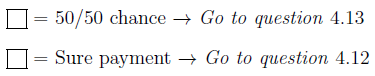
* 1. *(To be answered if replied 50/50 chance to question 4.2)*

Would you prefer the 50/50 chance or the amount of 120 dollars as sure payment?



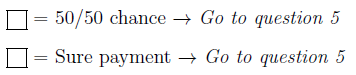
* 1. *(To be answered if replied Sure payment to question 4.10)*

Would you prefer the 50/50 chance or the amount of 100 dollars as sure payment?



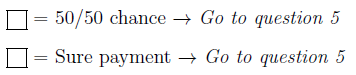
* 1. *(To be answered if replied Sure payment to question 4.12)*

Would you prefer the 50/50 chance or the amount of 90 dollars as sure payment?



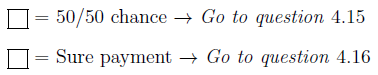
* 1. *(To be answered if replied 50/50 chance to question 4.11)*

Would you prefer the 50/50 chance or the amount of 110 dollars as sure payment?



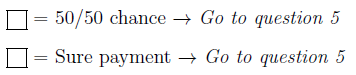
* 1. *(To be answered if replied 50/50 chance to question 4.11)*

Would you prefer the 50/50 chance or the amount of 140 dollars as sure payment?



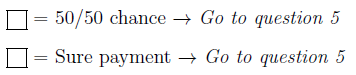
* 1. *(To be answered if replied 50/50 chance to question 4.14)*

Would you prefer the 50/50 chance or the amount of 150 dollars as sure payment?



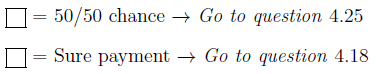
* 1. *(To be answered if replied Sure payment to question 4.14)*

Would you prefer the 50/50 chance or the amount of 130 dollars as sure payment?



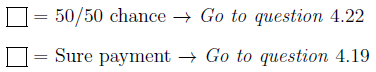
* 1. *(To be answered if replied 50/50 chance to question 4.1)*

Would you prefer the 50/50 chance or the amount of 240 dollars as sure payment?



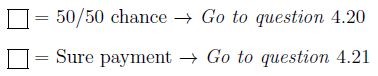
* 1. *(To be answered if replied Sure payment to question 4.17)*

Would you prefer the 50/50 chance or the amount of 200 dollars as sure payment?



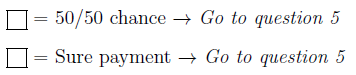
* 1. *(To be answered if replied Sure payment to question 4.18)*

Would you prefer the 50/50 chance or the amount of 180 dollars as sure payment?



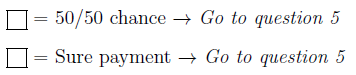
* 1. *(To be answered if replied 50/50 chance to question 4.19)*

Would you prefer the 50/50 chance or the amount of 190 dollars as sure payment?



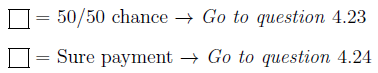
* 1. *(To be answered if replied Sure payment to question 4.19)*

Would you prefer the 50/50 chance or the amount of 170 dollars as sure payment?



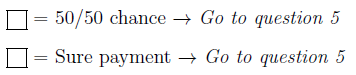
* 1. *(To be answered if replied 50/50 chance to question 4.18)*

Would you prefer the 50/50 chance or the amount of 220 dollars as sure payment?



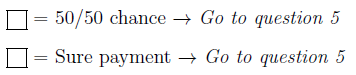
* 1. *(To be answered if replied 50/50 chance to question 4.22)*

Would you prefer the 50/50 chance or the amount of 230 dollars as sure payment?



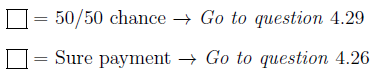
* 1. *(To be answered if replied Sure payment to question 4.22)*

Would you prefer the 50/50 chance or the amount of 210 dollars as sure payment?



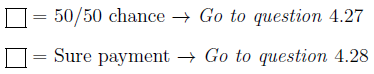
* 1. *(To be answered if replied 50/50 chance to question 4.17)*

Would you prefer the 50/50 chance or the amount of 280 dollars as sure payment?



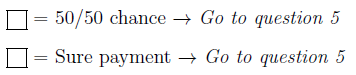
* 1. *(To be answered if replied Sure payment to question 4.25)*

Would you prefer the 50/50 chance or the amount of 260 dollars as sure payment?



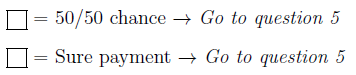
* 1. *(To be answered if replied 50/50 chance to question 4.26)*

Would you prefer the 50/50 chance or the amount of 270 dollars as sure payment?



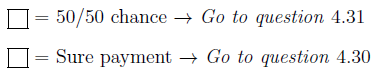
* 1. *(To be answered if replied Sure payment to question 4.26)*

Would you prefer the 50/50 chance or the amount of 250 dollars as sure payment?



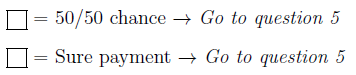
* 1. *(To be answered if replied 50/50 chance to question 4.25)*

Would you prefer the 50/50 chance or the amount of 300 dollars as sure payment?



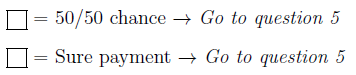
* 1. *(To be answered if replied Sure payment to question 4.29)*

Would you prefer the 50/50 chance or the amount of 290 dollars as sure payment?



* 1. *(To be answered if replied Sure payment to question 4.29)*

Would you prefer the 50/50 chance or the amount of 310 dollars as sure payment?



1. Please think about what you would do in the following situation.

You are in an area you are not familiar with, and you realize that you’ve lost your way. You ask a stranger for directions. The stranger offers to take you to your destination. Helping you costs the stranger about 20 dollars in total. However, the stranger says he or she does not want any money from you. You have 6 presents with you. The cheapest present costs 5 dollars, the most expensive one costs 30 dollars. Do you give one of the presents to the stranger as a "thank-you"-gift? If so, which present do you give to the stranger?

* No present
* The present worth 5 dollars
* The present worth 10 dollars
* The present worth 15 dollars
* The present worth 20 dollars
* The present worth 25 dollars
* The present worth 30 dollars

1. Imagine the following situation: Today you’ve unexpectedly received 1,000 dollars.

How much of this amount would you donate to a good cause?

*(Please pick a whole number between 0 and 1,000, 0 and 1,000 are allowed)*

\_\_\_\_\_\_

\_\_\_\_\_

## Demographics

This closing section of the survey consists out of three questions about your personal information.

1. What is your gender?

* Male
* Female

2. What is your nationality?

(*If multiple, please tick the one you most identify with)*

* American
* Indian
* Other, please specify:

\_\_\_\_\_\_\_\_\_\_\_\_\_\_

3. What is your age?

* Under 18
* 18 – 24 years
* 25 – 34 years
* 35 – 44 years
* 45 – 54 years
* 55 - 64 years
* 65 - 74 years
* 75 – 84 years
* 85 and older

## Debriefing

This is the end of the survey, thank you very much for taking part. Your answers are invaluable for the research project. Multiple concepts are measured, such as social preferences and risk-attitude. These concepts will be studied to test whether these can accurately predict the decisions made in the six games. Your answer will not be matched against other respondents, but will be used for training a model, to explain and to test the accuracy of the predictions made by the model.

Below, a final question about possible, additional interest in the research is asked, and there is an option to give feedback on the survey.

*Note that you will receive your validation code after this section.*

## Additional questions

Interested in receiving the final work, relevant data, and code of this research?

*(The supplied information is independent of the research and will not comprise anonymity)*

* No
* Yes, please fill in your e-mail address

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

### Feedback

For any questions, concerns or feedback about the survey or the research, do not hesitate to leave it below, I appreciate any questions, concerns or feedback you may have.

Jeroen van Buren

Tilburg University

j.vanburen@tilburguniversity.edu

## 

# Appendix C: Pilot questions

**Part 1: Ask after each relevant question**

1. Why have you made this decision?

*Please tell me, how much you disagree or agree with the following statement.*

1. The question was hard to understand

*(please tick in the answer that most corresponds with your view)*

* Very strongly disagree
* Strongly disagree
* Somewhat disagree
* Neither agree or disagree
* Somewhat agree
* Strongly agree
* Very strongly agree

*(If appropriate, ask the following question)*

1. What was harder to understand?
2. How could this question be improved?

**Part 2: Ask after completion of the questionnaire**

1. The questionnaire was hard to understand

*(please tick in the answer that most correspond with your view)*

* Very strongly disagree
* Strongly disagree
* Somewhat disagree
* Neither agree or disagree
* Somewhat agree
* Strongly agree
* Very strongly agree

*(If appropriate, ask the following question)*

1. What was harder to understand in the questionnaire?
2. How could this questionnaire be improved?

# Appendix D: Packages and “thesisFunctions.R”

## Packages

All packages that are used in this script are described and loaded in. The used functions of the packages are noted in brackets (). For more information see the corresponding help files.

* *QuantPsych*: used for extracting the standardized coefficients (lm.beta)
* *dplyr*: used for general data transformation (select, mutate, rename, inner\_join)
* *glmnet*: used for conducting the lasso method (cv.glmnet, glmnet)
* *gam*: used for constructing a generalized additive model containing a smoothing spline (gam)
* *gbm*: used for conducting a gradient boosting machine (gbm)
* *caret*: used for fitting and training the gradient boosting machine (trainControl, expand.grid, train)
* *car*: used for plotting the residuals (residualPlot, redidualsPlot)
* *bestglm*: used for finding the best subset selection of a logistic regression (bestglm)

## thesisFunctions.R file

Other packages and functions are loaded in via “thesisFunctions.R” file. These packages are used in the created functions and described below. The used functions of the package are again noted in brackets (). For more information see the corresponding help files.

* *psych*: used for descriptive statistics, a correlation test and for measuring Cronbach’s alpha (describe, corr.test, alpha)
* *glmnet*: Used for conducting the lasso method and for plotting the shrinkage of the coefficients (glmnet)
* *corrplot*: used for a visualisation of the correlations (corrplot)
* *leaps*: used for conducting a best subset selection method of a linear regression model (regsubsets)
* *splines*: used for conducting a smoothing spline (smooth.spline)
* *ROCR*: used for plotting a ROC curve and accuracy/F-score vs. threshold plot (prediction, performance)

Created functions loaded in via “thesisFunctions.R”. For more information see the corresponding file.

* *plotBoxHist*: plots a standardized boxplot and a histogram in one figure. The histogram includes a mean and median line. Used in the function describeVariable.
* *plotBoxBar*: plots a standardized boxplot and a bar plot in one figure. Used in describeVariable.
* *plotBox*: plots a standardized bar plot. Used in describeVariable.
* *describeVariable*: describes a variable and returns the proper descriptive statistics. The function changes the description and plots based on whether the type of variable is a factor, discrete- or a continuous variable. In case of the latter, the parameter histogram = TRUE needs to be supplied.
* *plotCor*: plots a detailed correlation plot, if pvalues = FALSE, all correlations are coloured gradually from -1 to +1, negative relations are coloured reddish and positive relationships are coloured blueish. If pvalues = TRUE, all significant relations are coloured gradually, all other insignificant relations will have a white background. Returns the correlation table and if pvalues = TRUE also the p-values table.
* *cbAlpha*: calculates the Cronbach’s alpha score, once for the normal input and once for the z-scaled variables of this input. Returns detailed scores and the correlations between the variables.
* *plotBoxes*: plots standardized boxplots for all preferences in one figure, split by a binary factor variable.
* *baseModelcv*: creates a base model and applies cross-validation. If a classification model is appropriate, then the majority class is taken as baseline and accuracy is measured. When a regression model is proper, RMSE is measured and the mean is used as a baseline. Returns the cross-validated score for every fold and the mean of these folds.
* *bestSubsetcv*: applies best subset selection and cross-validates scores for every best subset corresponding with that number of variables. In case of a classification, logistic regression is conducted, and all best models are supplied as a parameter. These are externally found via the bestglm function. In case of a regression model, linear regression is used, and all best models are internally found using regsubsets and the R2 score. Returns a matrix with cross-validated scores for every fold with that number variables, for all number of variables. Also returns the standard deviation, the mean for every number of variables, the ‘true’ best number of variables. and the corresponding score and standard deviation.
* *plotR2Subs*: plots the best R2 score for every number of variables and the variables corresponding with these scores in one figure. Returns the summary of regsubset.
* *plotModels*: plots the output scores of cross-validated models and places a red dot at the best performing model in the plot and returns these scores.
* *plotThreshold*: plots an Accuracy- and F-score- vs. threshold plot in one figure. If plotcontent = “ROC”, a ROC curve is plotted, and the AUC is returned.
* *plotShrinkage*: generates a glmnet model and then plots the shrinkage the coefficients. The red dotted line corresponds with the best lambda and a blue dotted line corresponds with the best lambda plus one standard error.
* *resultsLasso*: returns the results of the lassocv in a standardized way as in the other created cv functions. Returns the ‘true’ best lambda and the corresponding score and standard deviation and the best lambda plus one standard error. Takes a measure input and adjusts the output in case of a general linear model.
* *plotBivariate*: plots a bivariate plot with a x variable and y variable and draws the general linear model or linear model best fit line. The plot and fit are adjusted if y == binary variable.
* *nonLinearcv*: cross-validation is applied on non-linear models. The non-linear models should be supplied. Returns the standardized output as in bestsubsetcv.
* *smoothSplinecv*: cross-validation is applied on smoothing splines for a range of supplied degrees of freedom from 2 until a supplied parameter of maxdegree. Returns the standardized output as in bestsubsetcv, but with the degrees of freedom included.

1. For interested readers, the famous work by (Axelrod, 1984) pursues this issue further and describes scenarios wherein reciprocity creates co-operation in different kind of situations. [↑](#footnote-ref-1)
2. See Appendix B: Questionnaire under the section Game scenarios for the corresponding games and the description, pay-offs and the rules of the game that have been used in the questionnaire. [↑](#footnote-ref-2)
3. Calculation of the mixed Nash equilibrium, where y equals to the probability of picking chicken and (1 – y) to the probability of picking dare:

   Calculate EU (chicken) of player one: 10y + 3 (1 – y) == 10y + 3y – 3 == 7y – 3

   Calculate EU (dare) of player one: 17y + 0 (1 – y) == 17y

   In a mixed Nash equilibrium these are equal: 7y – 3 = 17y == 10y = 3 == y = 3/10

   Therefore, the optimal mixed Nash equilibrium has probability of picking chicken of 0.3 (30%) and dare of 0.7 (70%). [↑](#footnote-ref-3)
4. See the official website of Qualtrics for more information: <https://www.qualtrics.com/> [↑](#footnote-ref-4)
5. See the official website of Amazon Mechanical Turk for more information: <https://www.mturk.com/> [↑](#footnote-ref-5)
6. The packages and the “thesisFunctions.R” file that contains all created functions is described in Appendix D: Packages and “thesisFunctions.R”, the code and R markdown file are dependent on this file and is found on <https://github.com/jeroenzai> [↑](#footnote-ref-6)
7. All data, the R markdown file of the analysis, all other relevant files, code, versions of the questionnaire can be found online at <https://github.com/jeroenzai>. The R markdown document is also available as pdf. [↑](#footnote-ref-7)
8. For more information see the official website: <https://www.briq-institute.org/global-preferences/about> and <http://ftp.iza.org/dp9674.pdf> for the paper. The original version of this survey also measures time discounting, these questions are removed because of irrelevancy. [↑](#footnote-ref-8)
9. For the actual online questionnaire filled in by the respondents via Qualtrics, see the following link:

   <https://tilburghumanities.eu.qualtrics.com/jfe/form/SV_0CfHMSfCcrDzZs1> [↑](#footnote-ref-9)
10. The used qualifications in Amazon Mechanical Turk were:

    (1) HIT Approval Rate (%) for all Requesters’ HITs > 97%,

    (2) Number of HITs Approved > 10000 and

    (3) Location is UNITED STATES (US) [↑](#footnote-ref-10)
11. The were the variable named that contained “willingness” and “description” and were part of matrix table questions in Qualtrics. [↑](#footnote-ref-11)
12. The ordered variables are: public, offer, respond, dilemma, chicken, negRec1, negRec2, negRec3, posRec1, posRec2, altruism1, altruism2, trust, risk1, risk2. [↑](#footnote-ref-12)
13. See Figure 20 in Appendix A: Tables and figures for the correlation plot of the preferences. [↑](#footnote-ref-13)
14. The joined variable is negRecPooled and the removed variables are negRec1, negRec2 and negRec3. [↑](#footnote-ref-14)
15. See Figure 21 in Appendix A: Tables and figures for the frequency bar plot of the prisoner’s dilemma. [↑](#footnote-ref-15)
16. See Figure 22 in Appendix A: Tables and figures for the frequency bar plot of the chicken game. [↑](#footnote-ref-16)
17. For a more detailed description of the social- and risk preferences, see the R markdown file on GitHub. [↑](#footnote-ref-17)
18. The Likert variables and the second variable of positive reciprocity could also be represented via bar plots, but for convenience reasons histograms are chosen. [↑](#footnote-ref-18)
19. See Figure 26 in Appendix A: Tables and figures for the bivariate plot of negative reciprocity and the minimum acceptable amount in the ultimatum games [↑](#footnote-ref-19)
20. The plot is a simplified approximation because the best regression fit it only is based on one variable. [↑](#footnote-ref-20)
21. The plot is a simplified approximation because the best regression fit it only is based on one variable. [↑](#footnote-ref-21)
22. The actual questionnaire was filled in online by the respondents via Qualtrics. The following link leads to the questionnaire and is still available online:

    <https://tilburghumanities.eu.qualtrics.com/jfe/form/SV_0CfHMSfCcrDzZs1> [↑](#footnote-ref-22)
23. Since the used platform was Amazon Mechanical Turk, the input and generated code from Qualtrics had to be verified with the validation code from Amazon. [↑](#footnote-ref-23)
24. This is an attention check question to improve the quality of the responses by avoiding robots and inattentive respondents. [↑](#footnote-ref-24)
25. For more information see the official website: https://www.briq-institute.org/global-preferences/about and for the paper <http://ftp.iza.org/dp9674.pdf>. [↑](#footnote-ref-25)
26. In the online survey these filter questions are automated, and respondents are directed to the next question without intervention. [↑](#footnote-ref-26)