ESSAYS ON FAIRNESS, COOPERATION, AND COORDINATION

by

Devdeepta Bose

A Dissertation Submitted to the Faculty of the

Department of Economics

In Partial Fulfillment of the Requirements
For the Degree of

Doctor of Philosophy

In the Graduate College

The University of Arizona

2 0 1 8

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Devdeepta Bose, titled "Essays on fairness, cooperation, and coordination" and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Julian Romero — Date: (05/07/2018)

Charles Noussair — Date: (05/07/2018)

Andreas Blume — Date: (05/07/2018)

Martin Dufwenberg — Date: (05/07/2018)

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director: Julian Romero — Date: (05/07/2018)

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Devdeepta Bose

Acknowledgments

This work would not have been possible without the advice and guidance of Julian Romero, Charles Noussair, Andreas Blume, and Martin Dufwenberg. I am extraordinarily grateful to them for their help and support. I am especially thankful to my fantastic advisor, Julian, for always giving me a direction forward whenever I was stuck.

This work also benefited from helpful comments and discussions with countless other faculty members and fellow graduate students. In particular, I would like to thank Jackson Dorsey, Sean Inoue, Shariq Mohammed, Christopher Candreva, Bohan Ye, Karla Cordova, Laura Davidoff, Arundhati Tillu, Anatolii Kokoza, Juan Velez, Daehong Min, Sebastian Fleitas, Leslie Dupont, Smriti Sakhamuri, Shalini Sukumar, Ivy Chen, Wolf Gremel, Anooj Prabhoo, Rohan Mann, Khalid Baig, Shivkumar Kaushik, Rahul Reddy, Prasanna Ramachandran, and Anand Bhaskar for their feedback, advice, and support.

Dedication

I dedicate this manuscript to my mother, without whom none of this would have been possible. I also owe a debt of gratitude to my cousins, Babso and Mamondidi, who made a far-away land feel instantly like home. I am especially grateful to Angela, the love of my life, for reminding me that there is a bigger world out there.

Table of Contents

Table of Contents—*Continued*

TABLE OF CONTENTS—*Continued*

List of Figures

List of Figures—*Continued*

List of Figures—*Continued*

List of Tables

Abstract

The primary aim of this dissertation is to identify channels through which economic agents use social preferences such as proclivities towards fairness and inequality to make choices in one-shot and repeated game environments. I am particularly interested in economic situations of strategic interaction that support multiple equilibria. When faced with such situations, economic agents often coordinate on a set of salient outcomes that are smaller than the set of all possible equilibrium outcomes. There is a lack of broad consensus on what makes certain payoff outcomes salient in the minds of agents, while others are ignored. The human decision making involved in this equilibrium selection process interests me, because a better understanding of this mechanism can help us make sharper predictions about the plausible outcomes we should expect to see. I believe models that incorporate social preferences, learning, and bounded rationality are likely to make better predictions in these scenarios than models that assume economic agents are infinitely rational.

To examine the questions of how economic agents go about the equilibrium selection process, I use a combination of theory, computation, and laboratory experiments. Each of these three approaches has unique benefits, and complements each other. 1) Economic theory allows us to focus on a specific problem and prove general theorems that hold in all situations that the assumptions allow. Specifically, I am interested in theoretical models that allow for a combination of personal and social preferences for economic agents that weight notions of efficiency, selfishness, and fairness in personal and joint payoff outcomes when undertaking the equilibrium selection process. 2) Computation allows us to run simulations of models that may not have closed-form solutions, or ones that are difficult to grasp analytically. 3) Experiments allow us to test whether outcomes that agents arrive at after undergoing the equilibrium selection process are context dependent, and change based on the nature of the information available to them. They help examine human behavior in a controlled environment, allowing us to focus on the specific aspect of the decision making process in which we are interested. The combination of theory, computation, and experiments helps

us look into the black box of the decision making of economic agents undertaking the equilibrium selection process.

Chapter 1

Social Preferences in Infinitely Repeated Games

This paper investigates a learning model where players have attractions towards efficiency, equality, fairness, and selfishness in the payoff outcomes of indefinitely repeated two-player two-action games. Using simulated outcomes, the model explores one-state and two-state automata as potential strategies in games such as Battle of the Sexes, Prisoner's Dilemma, Stag-Hunt, and Chicken; and explains the high frequency of alternation between Nash equilibria outcomes in the Battle of the Sexes games, and the high degree of cooperation seen among experimental subjects in other games. The model adjusts naturally to games with asymmetric payoffs; or games where players have limited information about their opponents' payoffs. The tendencies of experimental subjects in various information and payoff settings can only be explained by weighting the considerations of efficiency, equality, fairness, and selfishness differently. Depending on the context, different strategies become focal in the minds of subjects, and any one of these considerations by itself fails to account for many patterns seen in laboratory results across different information or payoff treatments.

JEL Codes: C51, C72, C92, D03, D83.

Keywords: Reinforcement learning, Repeated game strategies, Adaptive models, Finite automata, Prisoner's Dilemma, Battle of the Sexes, Stag-Hunt, Chicken

1.1 Introduction

An extensive body of literature has attempted to explain the learning process in settings with repeated interactions between agents. However, many of these models do not focus on explaining how this learning process adapts to changes in payoff, or information settings. Using simulated outcomes, I explore learning in indefinitely repeated two-player two-action games such as Battle of the Sexes, Prisoner's Dilemma, Stag-Hunt, and Chicken using a model that adapts naturally to changes in payoffs, or changes in information settings. The learning model depends on two separate and equally important phases. Although these phases are motivated by different preferences; they complement each other and cover a range of situations that are common in an economic setting. In the first stage, henceforth referred to as the 'pre-experimental' phase, players form attractions towards repeated-game strategies based on the efficiency, equality, fairness, and selfishness of long-run average payoff outcomes. Since these attractions are innate and pre-existing, this stage is completely independent of the history of play. These attractions instead reflect the players' pre-game 'good intentions.' In the second stage, henceforth referred to as the 'experimental' phase, these attractions are updated after every round of play based on the history of the opponents' actions, and reinforcement dynamics that reflect one's own stage-game payoffs. Thus, this phase can be thought of as the players' pre-game 'good intentions' coming into conflict with the 'reality' of actual play. The first stage is a modified version of the Quantal Response Equilibrium (QRE) idea introduced by McKelvey and Palfrey (1995); the second stage borrows from the reinforcement learning literature pioneered by Roth and Erev (1998).

The behavior of human subjects in repeated games has been widely analyzed (see the references cited in Shubik (1982)). Axelrod (1984) conducted two tournaments that used computerized strategies, submitted by participants from a variety of backgrounds, to explain the behavior of laboratory subjects. His analysis showed that subjects have preferences for strategies that are both effective *and* simple. His analysis is compatible with research that has been conducted since; evidence clearly suggests that subjects adapt to outcomes in the laboratory and learn to play repeated games in an adaptive manner.

Models created to explain *adaptive learning* in games have existed since Cournot (1838), which assumes that players choose the best-response to their opponent's action in the previous round. Brown's (1951) fictitious-play model asserts that players choose the best response to the average of all previous observations of their opponent's play, not just the most recent. Both can be considered examples of 'belief based' models, in which subjects use the history of their opponent's choices to form *beliefs* about what actions their opponent is likely to play next, and choose actions that are the best responses to these updated beliefs. Roth and Erev (1998) capture the idea of reinforcement learning, in which players' propensities to choose particular actions increase or decrease based on the resulting payoff. Both belief-based and reinforcement learning models are special cases of Camerer and Ho's (1999) experience-weighted attraction (EWA) model, which reinforces actions both taken and *not taken*; based on perceived payoffs. All these models can be classified within the 'attractions towards actions' literature, and most papers in this vein differ in their claims about how attractions are initially formed; and updated throughout the game.

Cheung and Friedman (1997) posit a critique of the 'attractions towards actions' literature, claiming that empirical learning models should be sensitive to the institutional context, individual differences among players, and account for the decision process. Learning models should naturally adjust to changes in the payoff matrix (whether symmetric or asymmetric), or the level of information that players have regarding their opponents' payoffs. They ask and attempt to answer the following questions:. Do some payoff matrices encourage players to take a short Cournot view, while others encourage a longer fictitious-play view? Are institutional details important, such as what players can observe about their opponents' strategies and payoffs? They propose a flexible three-parameter model that changes naturally based on the institutional context.

The common thread running through all the adaptive models in the 'attractions towards actions' literature is that initial attractions are provided exogenously (or as in the case of EWA, selected to maximize overall fit), and players only learn about their attractions towards stage-game actions, rather than repeated-game strategies. While these models have been quite successful in accounting for observed behavior in some coordination games, they struggle to replicate behavior in others.[1]

A new class of models in the 'attractions towards strategies' literature extends the action-oriented models by assuming that players are learning about the appeal of repeated-game strategies rather than the appeal of stage-game actions through each round of game play. These models typically employ endogenous initial attractions, and assume that players are exploring tendencies towards a finite set of strategies with bounded complexity. Romero and Ioannou (2014) explore the strategies played in repeated games of Battle of the Sexes, Prisoner's Dilemma, Stag-Hunt, and Chicken. They use a generalized belief-updating model, where each player uses a fitness function based on the historical play of her opponent in order to update her beliefs about what strategies her opponent is using. The fitness function is used 1) to measure how closely an opponents' choices fit their potential strategies and then, given perceived strategies, 2) to choose strategies that maximize expected payoffs. The authors find that 'strategy learning' models outperform 'action learning' models when making predictions in many repeated games.

Hanaki, Sethi, Erev and Peterhansl (2005) have a reinforcement model where the initial attractions towards strategies are updated through reinforcement of payoffs. Relative to the 'action learning' models, they can better account for subject behavior in environments where fairness and reciprocity play a significant role. The authors do not explicitly use the idea of fairness in their model, but think that social preferences could explain behavior such as alternation of Nash equilibria outcomes in the Battle of the Sexes game. They point out that, *"One could conceivably account for the disparity between experimental findings and the predictions of learning models by arguing that subjects care not just about their own monetary payoffs, but also about the payoffs obtained by those with whom they interact. Several recent attempts have been made to identify a richer class of preferences that are able to take such interdependencies into account in a manner consistent with experimental behavior. From this perspective, payoff functions must be appropriately transformed before learning models*

[1] For example, these models cannot explain the well-established pattern of alternation between Nash equilibria outcomes in a Battle of the Sexes game. Arifovic and Ledyard (2015) find that, on average, 46% of subjects exhibit some pattern of alternation between Nash equilibria outcomes. However, this alternation ratio is sensitive to changes in payoff and information settings.

can be properly tested or compared. While this is an important and promising direction for research, there is as yet no consensus on the precise manner in which monetary payoffs should be transformed in order to conform to 'social preferences.' "

The model presented in this paper uses an appropriate transformation of payoff functions in a way that captures social preferences as well as inclinations for efficiency. The parsimonious learning model predicts experimental play routinely observed in the laboratory and also accounts for differences based on changes in payoff structures, and information settings. As such, my attempt is to provide an analog to Cheung and Friedman (1997) in the 'attractions towards strategies' literature. I find limiting attractions on strategies that explain alternations between Nash equilibria outcomes in Battle of the Sexes, and mutual coordination outcomes in Prisoner's Dilemma, Chicken, and Stag Hunt games for symmetric payoff matrices. I also find limiting attractions on strategies for these games once payoffs are asymmetric, or even payoff information is limited.

Is there a sound basis for considering fairness, or social preferences, as a consideration for attractions towards strategies in a repeated game? One-shot games such as Dictator and Ultimatum games that test notions of fairness find that people behave in a manner consistent with other-regarding preferences as expressed in Fehr and Schmidt (1999) among a host of other papers.[2] For repeated games, the evidence is not as comprehensive. Bhaskar (2010) shows that, for finite repetitions of the symmetric Battle of the Sexes game, as well as for the infinitely repeated game with discounting, among all symmetric equilibria of the game, those and only those that promise equal continuation payoffs at the history at which symmetries are broken are (Pareto-) efficient. All other equilibria yield lower ex-ante expected payoffs. Kuzmics, Palfrey, and Rogers (2014) follow up on Bhaskar's results, and ask that when symmetries are broken, how inefficient are other equilibria with asymmetric continuation payoffs, especially as players get infinitely patient? They find significant inefficiencies in the payoffs received by players for any non-egalitarian symmetric equilibria. They also find that when choosing across all symmetric equilibria, players are attracted towards

[2]See outcome-based distribution models such as Charness and Rabin (2002), and Bolton and Ockenfels (2000), or belief-based reciprocity theories explored in Rabin (1993), Dufwenberg and Kirchsteiger (2004), and Falk and Fischbacher (2006).

'focal' strategies, which are both Pareto-efficient and simple. Rabin (1997) introduces the idea of a 'fairness equilibrium' in a finitely repeated game environment, but the idea does not easily generalize to indefinitely repeated games, eliminating some mutually cooperative outcomes in the Prisoner's Dilemma. Duffy and Munoz-Garcia (2012) study a Prisoner's Dilemma game where they allow for social preferences like envy and guilt, and find that a high weight on social preferences can support cooperation for subjects with lower discount factors than agents with purely self-interested preferences. Thus, a preference for inequality can act as a substitute for patience when supporting mutual cooperation as an equilibrium outcome. Finally, Binmore (2010) lays a non-technical claim that while the folk theorem shows the possibility of multiple equilibria, society needs an equilibrium selection device in order to operate successfully. He further argues that fairness is evolution's solution to the equilibrium selection problem, and cites the food-sharing aspect of our ancestral game of life as an example of such a selection process.

The paper is organized as follows. In the next section, I explore the model by first introducing the repeated-game strategies available to agents, then developing the random utilities associated with each strategy, and identifying the equilibrium for agents showing the strategies that agents are attracted towards. Section 3 provides experimental results for the long-run payoffs attained by subjects in a range of different games in different information settings. This section also shows comparisons between simulated and experimental subjects to identify the weights that agents assign to notions of efficiency, absolute fairness (equality), relative fairness, and selfishness in repeated games. Section 4 considers three alternate models to provide robustness checks for the central assumptions of the model. Section 5 concludes.

1.2 Learning Model

In this section, I describe the learning model that aims to capture behavior among experimental subjects in two-player, two-action games such as Battle of the Sexes, Prisoner's Dilemma, Stag Hunt, and Chicken. I follow an 'attractions towards strategies' approach similar to Hanaki et al (2005), and Romero and Ioannou (2014), where an algorithm is used to generate initial attractions towards each of a finite set of strategies, and these attractions are updated during repeated rounds of game-play. The learning taking place in two stages. The 'pre-experimental' phase generates the initial attractions towards strategies, while the 'experimental' phase updates the attractions during rounds of game-play. The exact way in which this learning occurs, and how attractions are updated, are presented in the following sections. The finite set of strategies being considered contain all one-state, and two-state automaton or Moore machines.

1.2.1 Finite Automata

A finite automaton is a mathematical model of a system with discrete inputs and outputs, and can be used to represent repeated-game strategies. The system can be in any one of a finite number of internal configurations or "states," similar to the states in a repeated-game strategy. The state of the system summarizes the information concerning past inputs that is needed to determine the behavior of the system going forward. The specific type of finite automaton used in this paper is a Moore machine. A Moore machine for player i, M_i, in a repeated game $G = (I, \{A_i\}_{i \in I}, \{g_i\}_{i \in I})$ is a four-tuple $(Q_i, q_{0,i}, f_i, \tau_i)$ where Q_i is a finite set of internal states of which $q_{0,i}$ is specified to be the initial state, $f_i : Q_i \to A_i$ is an output function that assigns an action to every state, and $\tau_i : Q_i \times A_{-i} \to Q_i$ is the transition function that takes a players's current state and the opponent's action, and returns the next state in the player's system. It is pertinent to note that the transition function depends only on the present state and the other player's action. This formalization fits the natural description of a strategy as i's plan of action in all possible circumstances that are consistent with i's plans. In contrast, the notion of a game-theoretic strategy for i requires the specification of an action for every possible history, including those that are inconsistent with i's plan of action. It is important to highlight that to formulate the game-theoretic notion of a strategy, one would only have to construct the transition function so that $\tau_i : Q_i \times A \to Q_i$, instead of $\tau_i : Q_i \times A_{-i} \to Q_i$. In the first period, the state is $q_{0,i}$, and the automaton chooses the action $f_i(q_{0,i})$. If a_{-i} is the action chosen by the other player in the first period, then the state of i's automaton changes to $\tau_i(q_{0,i}, a_{-i})$, and in the second period, i chooses the action dictated by f_i in that state. Then, the state changes again according to the transition function given the other agent's action. Thus, whenever the automaton is in some state q, it chooses the action $f_i(q)$, while the transition function τ_i specifies the automaton's transition from q to a new state in response to the action taken by the other player.

1.2.2 Restricting the strategy space

Even though the action-space being considered in the four games is finite, allowing for mixed strategies gives us a very large number of potential strategies that can be chosen in these games. For any 'attractions towards strategies' learning model to have predictive power, the space of potential strategies allowed has to be limited. One way to restrict strategies is only considering those that can be represented as a Moore machine, and in this paper, I only consider strategies that can be represented by one-state and two-state automaton or Moore machines.

There are two possible states for an automaton and two possible actions by the opponent. The transition function thus maps four different possibilities of (own state, the opponent's action) pairs into the set of two states to be taken in the next period. This generates $2^4 = 16$ cases in total. Since each of these 16 cases can have one of two initial states, there are 32 possible automata. Four of these always play the first action, and another four always play the second. Elimination of non-unique automata yields a total of 26, of which two are one-state and the rest are two-state automata. Appendix A1 contains a complete listing of these.

One advantage of using only these 26 one-state and two-state automata to represent potential strategies in a game is that it naturally represents the concept of bounded complexity. It is unlikely that players are able to consider all possible strategies, and will consider a restricted set. Given that it is practically impossible to update attractions if the set of possible strategies is unboundedly large, I have chosen to restrict attention to less-complex ones.

The evidence that subjects in the laboratory only use low complexity one-state and two-state automata as strategies in indefinitely repeated games is somewhat mixed. Fudenberg, Rand and Dreber (2012) estimate that for treatments without noise (i.e. with perfect monitoring), 40% of strategies used needed more than two states to be captured. When public monitoring is imperfect, an even higher percentage of subjects use strategies that condition on more than just the last round to support cooperation. However, these results are much higher than the proportion found in Dal Bo and Frechette (2011), Frechette and Yuksel (2013) or Dreber, Fudenberg, and Rand (2008). Aoyagi and Frechette (2009) study a slightly different setting with imperfect monitoring, but still find that subjects only use memory-zero and memory-one strategies. It is possible that settings with noise and imperfect monitoring require a broader set of strategies, but the restriction on the strategy space is not prohibitive in the settings studied here.

1.2.3 Measurement metrics for strategies

The 26 strategies explored in the model are evaluated on four measurement metrics - efficiency, absolute fairness (equality), relative fairness, and selfishness. These metrics are calculated based on the long-run average payoff outcomes of following these strategies in a repeated game. The strategies that lead to payoff outcomes that maximize the joint payoffs of both players receive high efficiency scores. Strategies whose payoff outcomes minimize the payoff difference between the two players receive high absolute fairness (equality) scores. Many games have payoff functions where the entire convex hull of possible payoffs is very far away from any possible equality of payoffs. I develop a game-specific notion of relative fairness for such games that is described below, and all 26 strategies also receive a relative fairness score. Finally, strategies that maximize an individual's payoffs receive high selfishness scores for that individual.

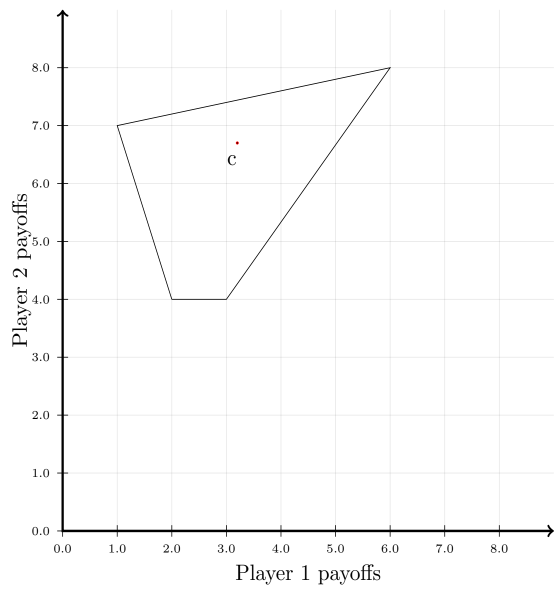
A generalized payoff function for a set of two-person, two-action games, where an individual's action set belongs to $\{A, B\}$, involves payoffs $u^i(A,A) = a^i, u^i(A,B) = c^i, u^i(B,A) = b^i, u^i(B,B) = d^i$. Table 1.1 below shows the generalized payoff function. The restricted strategy set is represented by $M = \{A1, A2, ..., A26\}$. Both players choose strategies from M, and play repeatedly against each other for 1,000 periods in the 'pre-experimental' phase. I calculate the long-run average payoffs received by the player for each strategy pair choice, denoted as $(u^{*,1}, u^{*,2})$. There are 676 such payoff pairs, given that both players have access to 26 strategies each.

Table 1.1: Payoff Table

	A	B
A	a^1, a^2	c^1, c^2
B	b^1, b^2	d^1, d^2

For each long-run average payoff pair, I use the algorithm described in the sections below to generate efficiency, absolute fairness (equality), relative fairness, and selfishness scores. The methodology for generating these scores for a randomly selected generic game (whose payoffs are shown in Figure 1.1) is described below.

Figure 1.1: Game Example

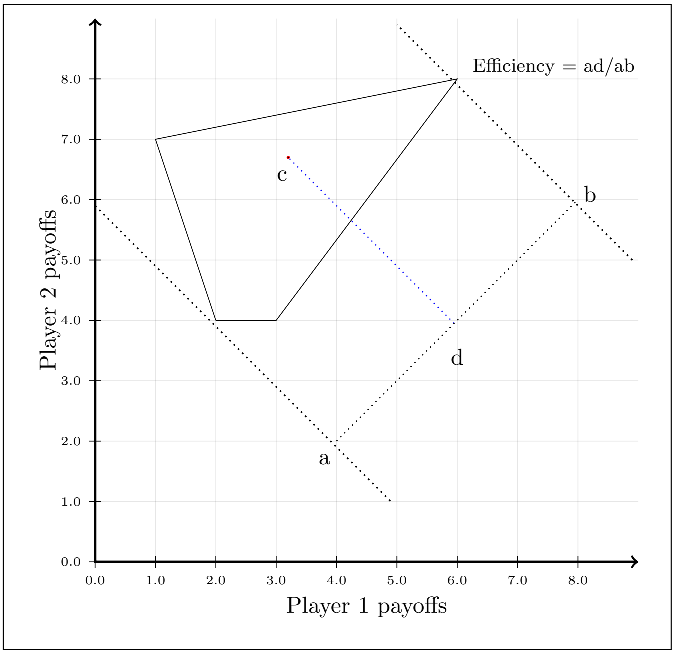


	A	B
A	1, 7	2, 4
B	3, 4	6, 8

Efficiency

The efficiency of a payoff outcome is measured by how close it comes to maximizing the sum of payoffs received by both players. Graphically, we can envision a series of iso-efficiency indifference curves perpendicular to a 45 degree line from the origin. Figure 1.2 shows the iso-efficiency indifference curves passing through the least, and most, efficient payoff outcomes in this generic game. Payoff outcomes on the iso-efficiency indifference curve that passes through the point that minimizes (maximizes) the sum of payoff outcomes for both players receive an efficiency score of zero (100). For any generic payoff outcome (such as point ‘c’ in Figure 1.2), we can define the iso-efficiency indifference curve that it lies on. We then calculate the distance from this curve to the most efficient indifference curve, relative to the distance between the two extreme indifference curves, and generate an efficiency score on a scale of zero to 100.

Figure 1.2: Efficiency

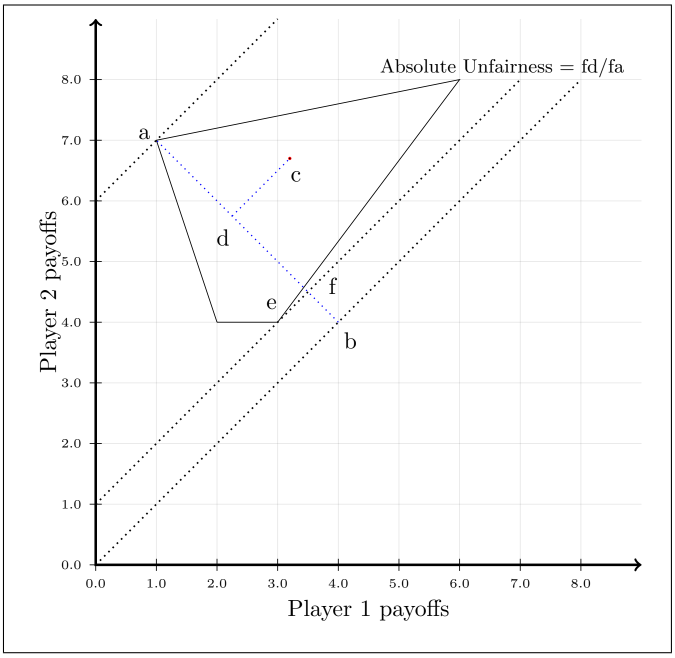


Absolute Fairness (Equality)

The equality of a payoff outcome is measured by how close it comes to minimizing the difference in payoffs received by both players. Graphically, we can envision a series of iso-equality indifference curves parallel to a 45 degree line from the origin. Figure 1.3 shows the iso-equality indifference curves passing through the least, and most, equal payoff outcomes in this generic game. Payoff outcomes on the iso-equality indifference curve that passes through the point that minimizes (maximizes) the difference in payoff outcomes for both players receive an inequality score of zero (100). For any generic payoff outcome (such as point ‘c’ in Figure 1.3), we can define the iso-equality indifference curve that it lies on. We then calculate the distance from this curve to the most equal indifference curve, relative to the distance between the two extreme indifference curves, and generate an inequality score on a scale of zero to 100. Absolute fairness (equality) scores are calculated by subtracting the inequality score from 100.

We use a notion of absolute fairness where outcomes that equalize payoffs are regarded as being perfectly fair, and the payoff outcome within the range of possible stage-game payoffs with the greatest level of inequality labeled as the most unfair outcome. All other payoff outcomes are measured in relative terms between these two extremes. The idea that fairness concerns are partly based on the range of possible outcomes can be justified by patterns seen in Dictator games studied in List (2007), where introducting the possibility of taking away money from receivers made Dictators less generous in their giving. In that one-shot non-strategic setting, expanding the payoff possibility space changes which payoffs agents come to regard as fair. Here, we extend this notion to a repeated game with strategic interactions.

Figure 1.3: Absolute Fairness (Equality)



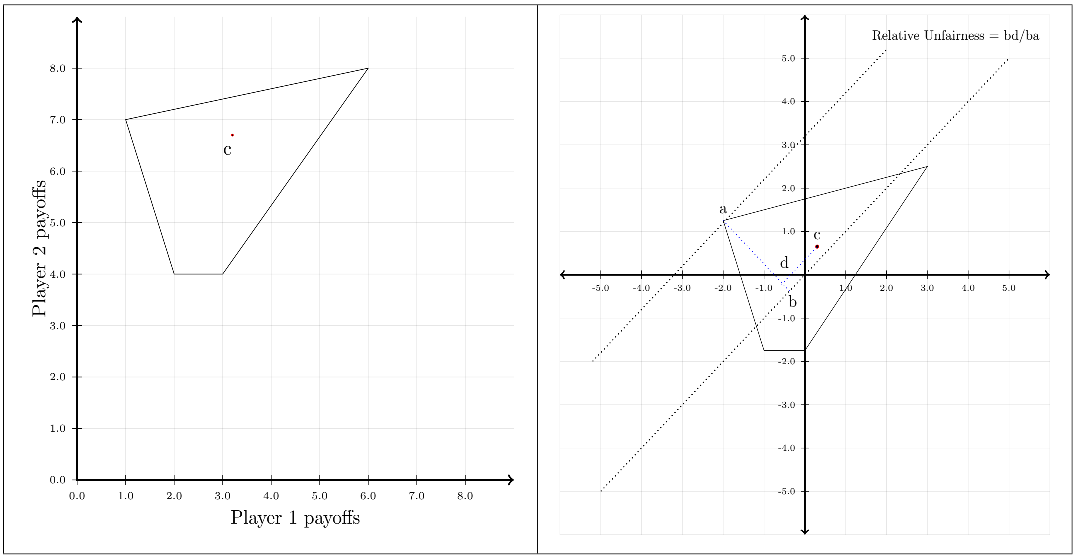
Relative Fairness

There are a number of laboratory studies that provide evidence that people have a natural aversion to inequality. People view equal distribution of resources as a moral good, and are likely to express dissatisfaction towards subjects that benefit from unequal distributions. In fact, subjects are willing to take costly actions to potentially achieve more equitable outcomes in one-shot Dictator, gift exchange, and public good games (Camerer (2003) and Schokkaert (2006) provide an overview). On the other hand, when people are asked about the ideal distribution of wealth in their country, they actually prefer unequal societies, as explored in Norton and Ariely (2011). Thus, people prefer equal distributions between a small group of people in laboratory situations, but prefer unequal distributions between large groups of people in the reaal world. It is difficult to reconcile the strong preference for equality in laboratory studies with a preference for societal inequality in political and behavioral economic research.

Starmans, Sheskin, and Bloom (2017) posit that the puzzle can be reconciled by focusing on the difference between equality and fairness. People will agree to unequal distributions if the underlying mechanism for generating the distribution is perceived to be fair. Akbas, Ariely, and Yuksel (2016) study two factors that subjects may use to determine the fairness of the distribution mechanism. To be perceived as fair, a distribution mechanism must have the following properties: procedural justice (so that subjects face equal prospects ex-ante), and agency (so that subjects have similar ability to determine their payoffs). There are many games where players do not face similar opportunities ex-ante, or similar payoffs. Thus, their notions of the fairness of a payoff outcome within a game may be determined by the range of possible outcomes that are theoretically possible ex-ante.

The notion of relative fairness of payoff outcomes is captured by transforming the entire convex hull of possible ex-ante payoffs. The payoff outcome determined by randomizing play $(u_{1,}u_2) = ((a_1+b_1+c_1+d_1)/4, (a_2+b_2+c_2+d_2)/4))$ is shifted so that it represents the new origin, as depicted in the transformed game represented in Figure 1.4. The fairness of a payoff outcome is measured by how close it comes to minimize the difference in payoffs received by both players in this transformed game. Graphically, we can envision a series of iso-fairness indifference curves parallel to a 45 degree line from the new transformed origin. Figure 1.4 shows the iso-fairness indifference curves passing through the least, and most, fair payoff outcomes in this generic game. Payoff outcomes on the iso-fairness indifference curve that passes through the point that minimizes (maximizes) the difference in payoff outcomes for both players receive an unfairness score of zero (100). For any generic payoff outcome (such as point ‘c’ in Figure 1.4), we can define the iso-fairness indifference curve that it lies on. We then calculate the distance from this curve to the most fair indifference curve, relative to the distance between the two extreme indifference curves, and generate an unfairness score on a scale of zero to 100. Relative fairness scores are calculated by subtracting the unfairness score from 100.

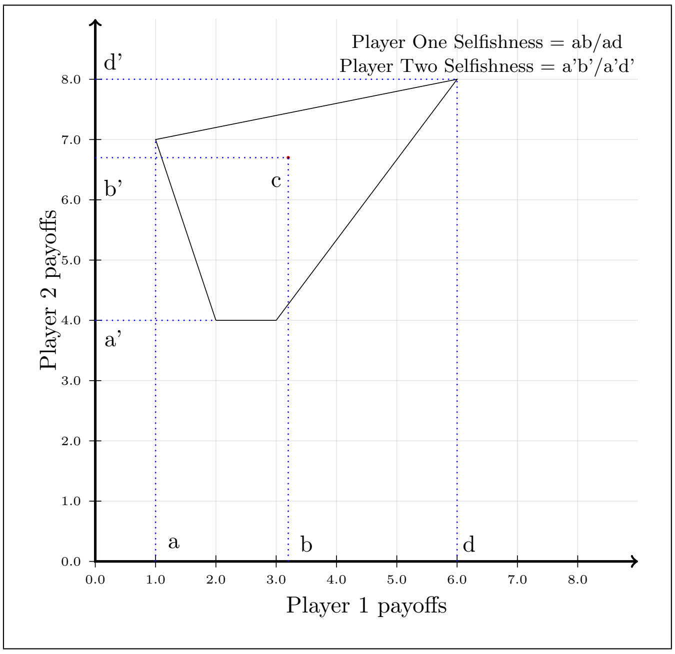
Figure 1.4: Relative Fairness



Selfishness

While the efficiency, absolute fairness (equality), and relative fairness scores of a payoff outcome are the same for both players, selfishness scores are different. The selfishness of a payoff outcome is measured by how close it comes to maximizing the payoff received by each individual player. Graphically, we can envision a series of iso-selfishness indifference curves that are vertical (horizontal) for Player 1 (Player 2). Figure 1.5 shows the iso-selfishness indifference curves for both players passing through the least, and most, selfish payoff outcomes in this generic game. Payoff outcomes on the iso-efficiency indifference curve that passes through the point that minimizes (maximizes) the payoff outcome for each player receive a selfishness score of zero (100). For any generic payoff outcome (such as point ‘c’ in Figure 1.5), we can define the iso-selfishness indifference curve that it lies on. We then calculate the distance from this curve to the most selfish indifference curve, relative to the distance between the two extreme indifference curves, and generate a selfishness score on a scale of zero to 100.

Figure 1.5: Selfishness



The methodology for how the efficiency, absolute fairness (equality), and relative fairness scores are calculated for the symmetric payoff versions of the four games studied in this paper (Battle of the Sexes, Prisoner's Dilemma, Chicken, and Stag Hunt) are presented in Appendix A2. Note that when payoffs are perfectly symmetric, absolute fairness (equality) and relative fairness notions coincide.

1.2.4 Initial attractions towards strategies (Pre-experimental Phase updates)

McKelvey and Palfrey (1995) introduce the idea of a Quantal Response Equilibrium (QRE), which uses statistical methods for predicting subject choices in a game-theoretic setting. Players best-respond to expected payoff utilities received by choosing particular actions, given their beliefs about the probability distribution of their opponent's actions. However, best responses are not played with certainty given that the expected payoff utilities calculated suffer from 'errors' which represent the bounded rationality of players. The best-response function is probabilistic in nature, which eventually converges to a fixed point. The learning model presented here has similar properties, but is slightly non-standard because I only allow for players to access a restricted set of automaton strategies, and so bound the strategy set in a specific way. I construct beliefs that players have regarding the probability distribution of their opponent's strategies using the logistic version of the QRE, and find that the algorithm leads to a fixed point. The intent here is not to explain how players learn to acquire mutually consistent beliefs, but rather to fit laboratory data to an equilibrium model incorporating noisy rational expectations.

As described above, the learning model contains 26 initial strategies considered by both players, represented by one-state and two-state automata. For the ease of exposition, let M_1 and M_2 denote the sets of 26 strategies available to Players 1 and 2 respectively.[3] For all 676 potential strategy pairs, I calculate the long-run average payoff for both players. The 26×26 matrix of potential payoffs has rows that denote the strategy choices of Player 1 and the columns denote the strategy choices of Player 2. Using the formulae for efficiency, absolute fairness (equality), relative fairness, and selfishness, I construct 26×26 matrices for these four indices. These are combined to form a composite matrix, denoted as H_1 which is the composite matrix for Player 1. Similarly, I also construct a 26×26 matrix of composite scores where the rows denote the strategy choices of Player 2 and the columns denote the strategy choices of Player 1, denoted as H_2.

The elements in $H_1$where Player 1 chooses strategy i and Player 2 chooses strategy j, is given by:

$$H_{1,i,j} = \alpha E_{1,i,j} + \beta AF_{1,i,j} + \delta RF_{1,i,j} + (1 - \alpha - \beta - \delta) S_{1,i,j}$$

The parameter α captures the propensity that players have towards efficient strategies, while β captures the propensity towards strategies that lead to equal payoffs, and δ captures the propensity towards fair strategies, while the parameter $(1 - \alpha - \beta - \delta)$ captures propensities towards strategies that maximize selfish payoffs. For a given parameter value of (α, β, δ), the matrices H_1 and H_2 are stationary. Let $p_1 = \{p_{1,1}, p_{1,2}, .., p_{1,i}, p_{1,j}, .., p_{1,26}\}$ and $p_2 = \{p_{2,1}, p_{2,2}, ..p_{2,i}, p_{2,j}, .., p_{2,26}\}$ denote the initial probability distribution (propensities) that Player 1 and Player 2 have towards playing the 26 different strategies. If Player 1 believes that Player 2 is playing strategies based on a probability distribution $p_2 = \{p_{2,1}, p_{2,2}, ..p_{2,i}, p_{2,j}, .., p_{2,26}\}$, then the

[3]This is only for the sake of exposition. Both players actually have access to the same set of strategies.

expected composite score for playing strategy i for Player 1 is given by:

$$\bar{H}_{1,i} = \sum_j (H_{1,i,j} * p_{2,j}).$$

Expected utilities for players are constructed from these expected composite scores. I assume (as in the QRE model) that the bounded rationality of players leads to 'errors' in the calculation of these expected composite scores, such that $\bar{u}_{1,i} = \bar{H}_{1,i} + \epsilon_{1,i}$.[4] I further assume that the errors $\epsilon_{1,i}$ are identical and independently distributed and have an EV(1) distribution. In addition, I have the behavioral assumption that players choose strategies that have higher expected composite scores. Thus, Player 1 has a behavioral tendency to select a strategy i if $\bar{u}_{1,i} \geqq \bar{u}_{1,k} \forall k \in M_1$. Similarly, Player 2 has a behavioral tendency to select a strategy j if $\bar{u}_{2,j} \geqq \bar{u}_{2,k} \forall k \in M_2$. Given the EV1 distribution of the errors, for an initial (p_1, p_2), Player 1 and Player 2's can recalculate their propensities towards their strategies using the following formulae for the logistic version of the QRE:

$$\tilde{p}_{1,i} = \frac{exp(\lambda(\bar{u}_{1,i}))}{\sum_{i \in M_1}(exp(\lambda(\bar{u}_{1,i}))} \text{ and } \tilde{p}_{2,j} = \frac{exp(\lambda(\bar{u}_{2,j}))}{\sum_{j \in M_2}(exp(\lambda(\bar{u}_{2,j}))}$$ [5][6]

The parameter λ, which is greater than or equal to zero, captures the scaling effect, and sensitivity towards the composite scores. In this logistic transformation, it represents the extent to which strategies with higher attractions are favored in strategy choice. When $\lambda = 0$, all strategies are equally likely to be chosen, regardless of their attractions. As λ increases, strategies with higher attractions become disproportion-

[4]There are a number of interpretations of these 'errors' as discussed in Haile, Hortacsu, and Kosnenok (2007). It is not necessary to take these 'errors' in composite score calculations literally. It is simply possible that players use strategies that are 'arbitrarily close' to a best-response rather than 'always fully optimal.' This interpretation has an intuitive appeal, especially given the challenging assumption of completely rational expectations about opponents' behavior that exist in all optimal best-response game-theoretic models. These deviations from 'always fully optimal' best-responses can also be viewed as a way to 'smooth out' best-response functions so as to obtain more robust predictions as in Rosenthal (1989). McKelvey and Palfrey (1995) suggest that these 'errors' have natural economic interpretations. These 'errors' can simply reflect unmodeled costs of information processing involved in calculating 'always fully optimal' best-response strategies. It is also possible that the 'errors' reflect unmodeled determinants of the players' utilities. This interpretation is appealing in many applications since a fully specified theoretical model can, of course, only approximate a real economic environment.

[5]I rescale values such that $max_j \lambda(\sum_i H_{2,j,i} * p_{1,i})$ and $max_i \lambda(\sum_j H_{1,i,j} * p_{2,j})$ has a value of 700, because the algorithm cannot calculate values above exp(700).

[6]It is an empirical question whether the logit, probit, or power forms fit better (adjusting for degrees of freedom). Previous studies show roughly equal fits of logit and power (Tang (1996), Chen and Tang (1998), Roth and Erev (1997)), or better fits for the logit form than the power form (Camerer and Ho (1998).

ately more likely to be chosen. In the limiting case $\lambda \to \infty$, the strategy with the highest attraction is chosen with probability one.

Thus, for any initial point (p_1, p_2) in the $[0,1]^{26}$X $[0,1]^{26}$ simplex that satisfies $\sum_i p_{1,i} = 1$and $\sum_j p_{2,j} = 1$, I have a mapping that generates a new point $(\tilde{p}_1, \tilde{p}_2)$ in the same simplex. The mapping routine continues until the following convergence criterion for the simulations is satisfied.

$max_{i \in A}|p_{1,i}(m) - p_{1,i}(m-1)| < \epsilon$ and $max_{j \in A}|p_{2,j}(m) - p_{2,j}(m-1)| < \epsilon$ for $\epsilon =$ 0.00001 for iteration m.

While the existence of the fixed point is guaranteed[7], in principle, it is possible to have a simulation run that does not satisfy the convergence criterion if ϵ is very small. However, I obtained convergence in all cases, with one caveat. For some parameter groups $(\alpha, \beta, \delta, \lambda)$, the covergence algorithm reaches an alternation between two points, back and forth in succession. For the sake of exposition, let us label these points $(\hat{p}_1, \hat{p}_2)$ and $(\check{p}_1, \check{p}_2)$. For all such situations, I introduce a parameter $\gamma \in [0,1]$, and consider linear combinations of the two probability distributions as the final fixed point. So, the final fixed points I consider are $(\mathring{p}_1, \mathring{p}_2) = \gamma(\hat{p}_1, \hat{p}_2) + (1-\gamma)(\check{p}_1, \check{p}_2)$, normalized such that $(\mathring{p}_1, \mathring{p}_2)$ still satisfies the properties of a probability distribution. I evaluate possible values of γ ranging from $\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.8, 1.0\}$, thus considering linear interpolations of the two probability distributions between which the algorithm is cycling.

It is important to note the mapping is *not* a contraction, so the fixed point is not unique. The fixed point reached at the end of the convergence algorithm depends on the starting probability distribution selected. I choose a natural starting point, where both players assume a uniform distribution over the opponent's likelihood of playing strategies. Thus, the initial starting point for (p_1, p_2) is a uniform distribution across all 26 strategies for both players. A natural question is whether the choice of a uniform distribution as a starting point might be restrictive. I have run simulations using 100,000 randomly chosen starting probability distributions, to check if they converge

[7]As shown in Appendix A3

to different fixed points. A small set of fixed points could suggest that there is a relatively small universe of player 'types', whose 'pre-experimental' preferences are represented by these converged fixed points. For all four games, I find a small set of limiting distributions, and in most cases, the payoff outcomes when starting from the uniform distribution is a reasonable representation of all possible 'types' of players' 'pre-experimental' preferences. The payoff outcomes for both players, when using this set of limiting distributions, is presented in Appendix A4.

The converged fixed point when starting from a uniform distribution probability distribution for both players represents the initial attractions that players have towards strategies, developed in the 'pre-experimental' phase of game-play. These fixed points are similar to a QRE, but with important differences described above.

1.2.5 Updating attractions towards strategies (Experimental Phase updates)

The previous few sections show an algorithm where the players' attractions towards efficiency, equality, fairness, and selfishness help determine the initial attractions they have towards each strategy from the set of 26 strategies representable by one-state and two-state automata. One interpretation of these attractions is the tendencies that players 'bring into the lab', prior to playing any of the rounds in the laboratory. For a repeated game setting, it is important to model how these initial attractions are updated after each round. In any appropriate learning model, players must update their attractions in a way to best-respond to the outcomes of observed play. To incorporate a sense of 'best-response' to observed play during the different rounds, I add the payoffs generated from each round to the expected composite scores being used to calculate attractions towards strategies. Thus, the updating of the initial attractions of strategy learning is done through reinforcement of payoffs as in Hanaki et al (2005).

Before the first round, Player 1 has a 26×26 composite score matrix H_1 and Player 2 has a 26×26 composite score matrix H_2. Let $\ddot{p}_1 = \{\ddot{p}_{1,1}, \ddot{p}_{1,2}, .., \ddot{p}_{1,i}, \ddot{p}_{1,j}, .., \ddot{p}_{1,26}\}$ and $\ddot{p}_2 = \{\ddot{p}_{2,1}, \ddot{p}_{2,2}, ..\ddot{p}_{2,i}, \ddot{p}_{2,j}, .., \ddot{p}_{2,26}\}$ denote the converged fixed point probability distribution (propensities) that Player 1 and Player 2 have towards playing the 26 different strategies. The expected composite score for playing strategy i (strategy j) for Player 1 (Player 2), is given by:

$$\tilde{H}_{1,i} = \sum_j (H_{1,i,j} * \ddot{p}_{2,j}) \text{ and } \tilde{H}_{2,j} = \sum_i (H_{2,j,i} * \ddot{p}_{1,i}) \ .$$

In the repeated game, for any round, let Player 1 choose strategy i and let Player 2 choose strategy j, and let the payoff for both players be a_1 and a_2 respectively. I update the expected composite score for Player 1's strategy i by adding in the payoff: $\tilde{H}_{1,i} = \tilde{H}_{1,i} + a_1$, while all the expected composite scores for the other strategies remain unchanged. Thus, the expected composite score for the strategy *that the player chose in a particular round* is updated to reflect the effects of that choice (i.e the payoffs), while the expected composite scores for all other unchosen strategies remain unchanged. Similarly, for player 2, I update the expected composite score for Player 2's strategy j by adding in the payoff: $\tilde{H}_{2,j} = \tilde{H}_{2,j} + a_2$, while all the expected composite scores for the other strategies remain unchanged. After every round, attractions towards strategies in the next round are then re-calculated using the logistic algorithm described in the earlier section. Updating propensities in this way means that if the payoffs from choosing a particular strategy in a particular round are larger, then it increases the likelihood of players choosing that strategy in future rounds by a larger amount. Draws taken from these constantly updated propensities are used to simulate play between players.

The learning model described here is flexible enough to easily adapt to an asymmetric payoff setting since the values for efficiency, equality, fairness, and selfishness update naturally based on the geometry of the payoff space. The model also deals well with restricted information environments. If a player is only aware of her own payoffs, then she has no notion of efficiency, absolute or relative fairness. In this circumstance, the matrices E, AF amd RF take on a value of zero, and players' composite scores only consider their selfishness concerns.

1.3 Results

The model attempts to capture behavior seen in the laboratory for repeated games such as Battle of the Sexes, Chicken, Prisoner's Dilemma, and Stag Hunt. These games represent a broad class of scenarios with different kinds of cooperative and non cooperative equilibria. A model that attempts to understand behavior should be able to predict behavior in a wide range of games. The four 2 X 2 games, though simple, capture important aspects of everyday experiences, such as cooperation, coordination and reciprocity. In the Prisoner's Dilemma, for most payoff matrices, the dominant action of the player is to defect. Thus (B,B) is theoretically predicted to be the most common outcome of the game. Meanwhile, the Battle of the Sexes has two pure-strategy Nash equilibria denoted by (A,B) and (B,A) respectively. Alternating between the two Nash equilibria in a deterministic way can lead to outcomes that are Pareto efficient but also give both players equal payoffs. The Stag-Hunt game has two pure-strategy equilibria denoted by (A,A) and (B,B). However, outcome (A,A) is the payoff-dominant Nash equilibrium. The Chicken game has a mixed symmetric Nash equilibrium and two pure-strategy Nash equilibria (A,B) and (B,A). However, the mutual co-operation outcome of (A,A) yields higher payoffs than alternating between the two pure-strategy Nash equilibria. Thus, the four games cover a range of different scenarios and the model attempts to capture behaviour in all of them.

1.3.1 Experimental subjects

All four games are played in three treatment settings, by changing the payoffs, or providing subjects with differing levels of information. The three treatments are described below.

Full Information, Symmetric Payoffs

In this treatment, subjects have full information about their own payoffs, as well as their opponents' payoffs. Payoffs are symmetric, so the payoffs for playing a particular strategy depend only on the strategies employed, not on who is playing them. The experimental data for this treatment is sourced from Mathevet and Romero (2012), who studied a total of eight games. The payoffs of the subjects in the four games that are the focus of this paper are reported in Table 1.2. The experimental sessions were run at the Vernon Smith Experimental Economics Laboratory at Purdue University. Subjects interacted on computers using an interface that was programmed using the z-Tree software (Fischbacher, 2007). Subjects' final payoffs consisted of the sum of their earnings from all periods of the experiment. With the completion of the experiment, subjects were paid their cash earnings in private. The average payoff was $16.85.

Table 1.2: Full Information, Symmetrics Payoffs Matrix

	A	B
A	3, 3	1, 4
B	4, 1	2, 2

(a) Prisoner's Dilemma

	A	B
A	3, 3	0, 2
B	2, 0	1, 1

(b) Stag Hunt

	A	B
A	1, 1	2, 4
B	4, 2	1, 1

(c) Battle of the Sexes

	A	B
A	3, 3	1, 4
B	4, 1	0, 0

(d) Chicken

The game-play consisted of a fixed matching protocol. There were a total of 16 sessions consisting of 314 subjects. In each session, subjects played a total of three supergames. Subjects never played the same game twice, and were never matched with another subject twice. At the beginning of each supergame, each subject was randomly paired with one other subject and they played a 2 X 2 game repeatedly with the same partner. After all pairs had finished their supergames, subjects were rematched with a new partner that they had not been matched with before, and played a game that they had not played before. The length of the supergame was determined using the following termination rules :

- T30 - 30 rounds were played with certainty, and then starting in the 31st round, the continuation probability 0.9, so the expected length of each supergame was 40.

- T0 - Starting in the first round, the continuation probability was 0.99, so the expected length of each supergame was 100.

In each round, subjects were asked to select a choice (A or B). The payoffs were displayed in a table, and players roles were always the same (i.e. they didn't have to switch between row and column player). The final experimental data consists of 70, 70, 70 and 50 observation pairs in the Prisoner's Dilemma, Battle of the Sexes, Chicken and Stag Hunt game, respectively.

Full Information, Asymmetric Payoffs

In this treatment, I perturb the payoff matrix in Mathevet and Romero (2012) so that payoffs are no longer symmetric. Subjects are still fully aware of both their own payoffs, as well as those of their opponents. The experimental sessions were run at the Economic Science Laboratory at the University of Arizona. Subjects interacted on computers using a Google Chrome interface that was programmed using Python. Subjects' final payoffs consisted of the sum of their earnings from all periods of the experiment. With the completion of the experiment, subjects were paid their cash earnings in private. The average payoff was $15.2.

Table 1.3: Full Information, Asymmetric Payoffs Matrix

	A	B
A	3, 2.5	1, 3.5
B	4.5, 1	2, 2

(a) Prisoner's Dilemma

	A	B
A	2.5, 3	0, 2
B	2, 0	1, 1

(b) Stag Hunt

	A	B
A	1, 1	2, 4.5
B	3.5, 2	1, 1

(c) Battle of the Sexes

	A	B
A	2.5, 3	1, 4.5
B	3.5, 1	0, 0

(d) Chicken

The game-play consisted of a fixed matching protocol. There were a total of 7 sessions consisting of 100 subjects. In each session, subjects played all four supergames. Subjects never played the same game twice and were never matched with another subject twice. At the beginning of each supergame, each subject was randomly paired with one other subject and they played a 2 X 2 game repeatedly with the same partner. After all pairs had finished their supergames, subjects were rematched with a new partner that they had not been matched with before, and played a game that they had not played before. The length of the supergame was determined using the following termination rules :

- T0 - Starting in the first round, the continuation probability was 0.99, so the expected length of each supergame was 100.

In each round, subjects were asked to select a choice (A or B). The payoffs were displayed in a table, and players roles were always the same (i.e. they didn't have to switch between row and column player). The final experimental data consists of 50 observation pairs in all four games. The instructions for this treatment settings are available in Appendix A6, while a screenshot showing the interface seen by subjects is available in Appendix A7.

Dark Information, Symmetric Payoffs

In this treatment, I use the same payoff matrix as in Mathevet and Romero (2012), but subjects can only see their own payoffs and their opponents' action choices, but cannot see their opponents' payoffs. The experimental sessions were run at the Economic Science Laboratory at the University of Arizona. Subjects interacted on computers using a Google Chrome interface that was programmed using Python. Subjects' final payoffs consisted of the sum of their earnings from all periods of the experiment. With the completion of the experiment, subjects were paid their cash earnings in private. The average payoff was \$15.1.

Table 1.4: Dark Information, Symmetric Payoffs Matrix

	A	B
A	3,	1,
B	4,	2,

(a) Prisoner's Dilemma

	A	B
A	3,	0,
B	2,	1,

(b) Stag Hunt

	A	B
A	1,	2,
B	4,	1,

(c) Battle of the Sexes

	A	B
A	3,	1,
B	4,	0,

(d) Chicken

The game-play consisted of a fixed matching protocol. There were a total of 12 sessions consisting of 100 subjects. In each session, subjects played all four supergames. Subjects never played the same game twice, and where possible, were never matched with another subject twice. At the beginning of each supergame, each subject was randomly paired with one other subject and they played a 2 X 2 game repeatedly with the same partner. After all pairs had finished their supergames, subjects were rematched, where possible, with a new partner that they had not been matched with before, and played a game that they had not played before. The length of the supergame was determined using the following termination rules :

- T0 - Starting in the first round, the continuation probability was 0.99, so the expected length of each supergame was 100.

In each round, subjects were asked to select a choice (A or B). The payoffs were displayed in a table, and players roles were always the same (i.e. they didn't have to switch between row and column player). The final experimental data consists of 50 observation pairs in all four games. The instructions for this treatment settings are available in Appendix A6, while a screenshot showing the interface seen by subjects is available in Appendix A7.

Experimental Data Patterns

There are many interesting patterns observable in the experimental data. The first result is that there is a clear treatment effect on the actions of subjects. The best illustration of this is provided in Figure 1.6 below which shows the heat map with the relative frequency of each payoff combination over the last 20 periods of game-play for the experimental subjects. The relative frequency of a payoff combination is denoted by a circle located on the coordinates that correspond to that combination; the larger the circle, the higher the (relative) frequency of that combination. I also display the set of feasible payoffs. Across the three treatments, the patterns in the final set of outcomes are not entirely captured by the Nash equilibria outcomes, but also do not support the entire set of Folk Theorem payoffs.

There is remarkable coordination on alternation between pure Nash equilibria outcomes in the full information, symmetric payoff setting of Battle of the Sexes (37% of subjects), an outcome that promotes both efficiency and fairness. However, this coordination breaks down when payoffs become asymmetric with only 12% of subjects alternating, and 15% of subjects ending up at each of the two pure Nash equilibria outcomes. Subjects are clearly acting more selfishly when payoffs become asymmetric. Finally, if subjects only have information about their own payoffs, this notion of selfishness becomes even more potent with 36% of subjects ending up at each of the two pure Nash equilibria outcomes, with only 4% of subjects managing to alternate.

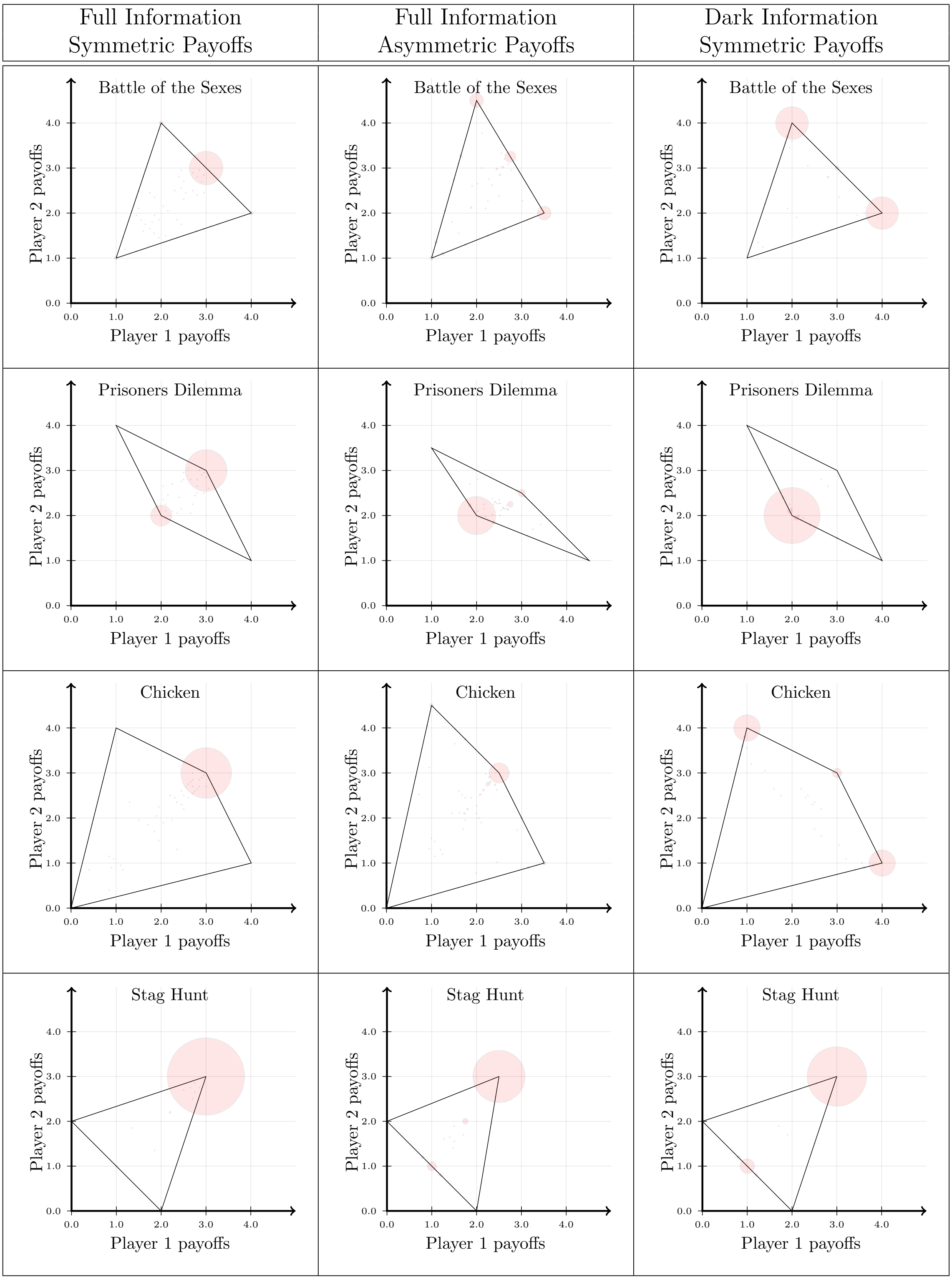
In the Prisoner's Dilemma, there is a tension between the dominant strategy of defection against the payoff-dominant situation of mutual cooperation. In the full information, symmetric payoff setting, although there are some subjects (23%) that end up at the mutual defection outcome, a large number (46%) are able to coordinate on the efficient and fair mutual cooperation outcome. However, this coordination breaks down when payoffs become asymmetric, with 42% of subjects mutually defecting, and only 8% of subjects ending up at the mutual cooperation outcome. This breakdown is further exacerbated once payoff information becomes limited, with 62% of subjects ending up at the mutual defection outcome, and no pair successfully coordinating on the mutual cooperation outcome.

Although the game of Chicken has three potential Nash equilibria outcomes, most subjects in the full information, symmetric payoff setting (56%) end up at the payoff dominant mutual cooperation outcome, which is not a Nash equilibria outcome. Once payoffs become asymmetric, this coordination breaks down (only 22% of subjects), while 4% of subjects end up at each of the two pure strategy Nash equilibria outcomes. This breakdown is further exacerbated once payoff information becomes limited, with only 10% of subjects ending up at the mutual cooperation outcome, and 29% of subjects ending up at each of the two pure strategy Nash equilibria outcomes. The notion of selfishness becomes more potent once payoffs become more asymmetric, or payoff information becomes limited.

Although the game of Stag Hunt has three potential Nash equilibria outcomes, the mutual cooperation outcome is payoff dominant. Not surprisingly, an overwhelming number of subjects in the full information, symmetric payoff setting (86%) end up at this outcome. Once payoffs become asymmetric, this coordination breaks down slightly (only 58% of subjects), while 10% of subjects end up at the mutual defection

Nash equilibria outcome. This breakdown is not as bad in the limited payoff information setting, because selfishness by both players also leads to the mutual cooperation outcome. 66% of subjects ending up at the mutual cooperation outcome, and 16% of subjects ending up at the mutual defection Nash equilibria outcome.

Figure 1.6: Experimental patterns across different treatments



1.3.2 Simulated subjects

The learning algorithm described in Section 2 is used to simulate play between pairs of players. I have run a grid search over different parameter values, with

$\alpha \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

$\beta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

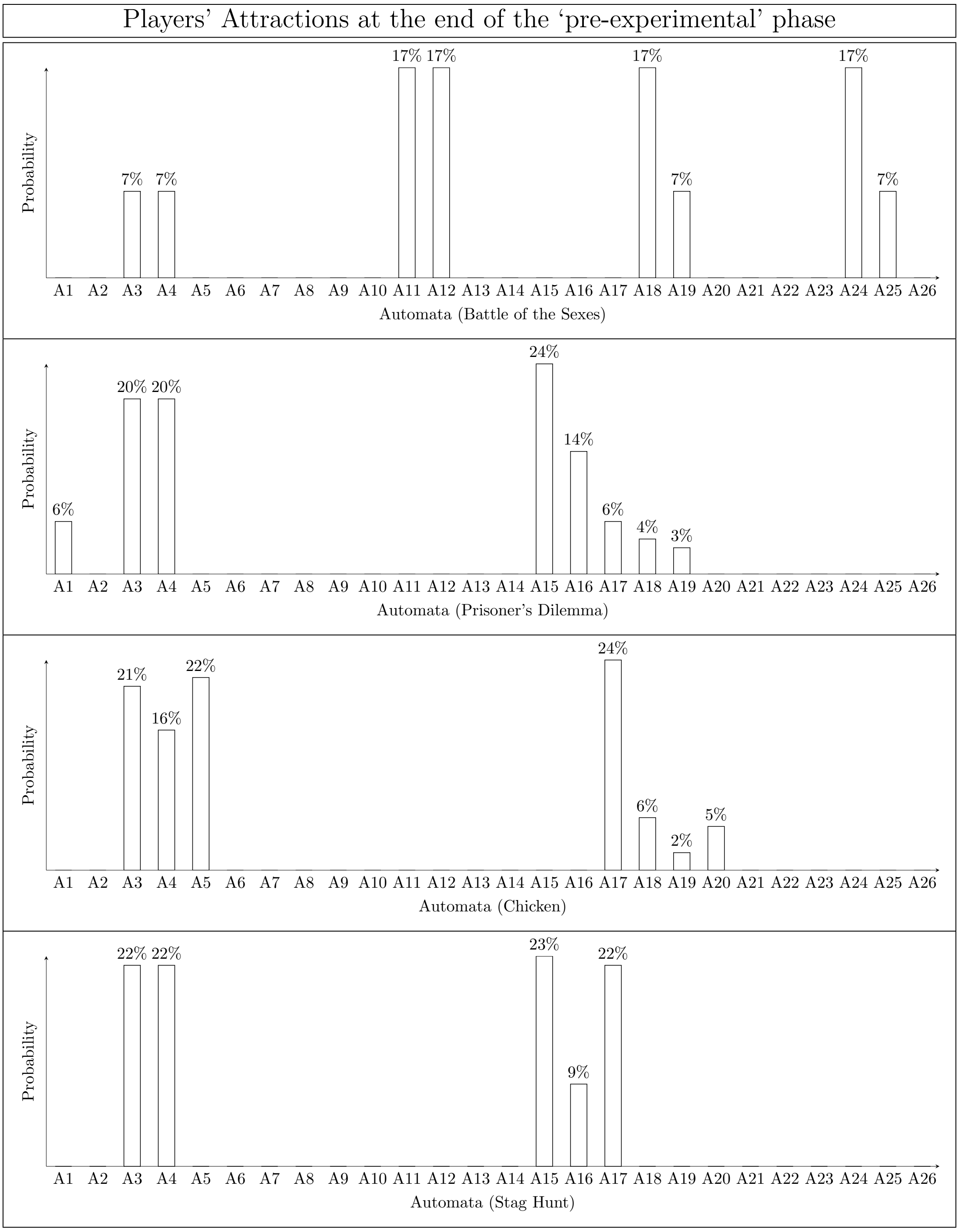
$\delta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

$\lambda \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

and $\gamma \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$.

For most parameter values, the converged fixed points for the initial attractions follow a pattern where most of the probability weight is on a few well-known strategies. For example, Figure 1.7 shows the histogram of the probability distribution for the converged initial attractions for both players across all four games in the full information, symmetric payoffs treatment, when considering parameter values $(\alpha = 0.2, \beta = 0, \delta = 0.8, \lambda = 0.5, \gamma = 0.7)$.

Figure 1.7: Converged initial attractions for the two players in each of the four games (full information, symmetric payoffs treatment)



To simplify the discussion, I focus on strategies with limiting probability weights greater than or equal to 5%. These strategies are, in a sense, the principal strategies that players 'bring to the laboratory' for the 'experimental' phase.

Battle of the Sexes

Strategies A11 and A24 mechanically alternate between two states, but have different initial states. Strategy A12 or the 'co-ordinate once' strategy starts by playing action A and regardless of the opponent's action, switches to state B. It stays in this state if the opponent plays B; otherwise it returns to state A for one round before returning to state B. Strategy A18 or the 'punish once' strategy is qualitatively identical to strategy A12, with the role of the two actions reversed. Strategies A4 and A19 are 'tit-for-tat' strategies, where players match the action taken by the opponent. Strategy A3 is a 'punish once' strategy, which starts by playing action A and remains in that state until the opponent plays B. At this point, the strategy plays B once before returning to play A irrespective of the opponents' action. Strategy A25 or the 'co-ordinate once' strategy is qualitatively identical to strategy A3, with the role of the two actions reversed.

Note that strategies A3, A12, A18 and A25, when matched with any of the other strategies having significant weights, eventually achieve perfect alternation between the two pure strategy stage-game equilibria. Since these 'flexible' strategies are among the ones that have the highest weights, players in the 'experimental' stage will eventually learn how to use them even if initial actions are not coordinated on the equitable and efficient outcome.

Prisoner's Dilemma

For both players, strategies obtaining high limiting probability weights are 'always play A,' 'punish once,' 'tit-for-tat,' 'quick B, then A,' 'BA triggers A,' and 'BB triggers A', respectively. The 'always play A' strategy prescribes playing the action A forever in all scenarios. The 'quick B, then A' strategy starts at action B and and then mechanically moves to state A and stays there forever after. The 'BA triggers A' strategy starts at action B and stays at at that state as long as the opponent plays B. As soon as the opponent plays A once, the strategy dictates playing A forever after.

The ‘BB triggers A’ strategy starts at action B and stays at at that state as long as the opponent plays A. As soon as the opponent plays B once, the strategy dictates playing A forever after. When most of these six strategies are played amongst themselves, the observed history of actions will involve mutual cooperation in all periods. The set of strategies with limiting probability weights are similar to the ones predicted in Hanaki et al (2005) and Dal Bo and Frechette (2015), which consider finite repetition of rounds, and bounded complexity of strategies respectively.

Chicken

For both players, the strategies ‘punish once,’ ‘tit-for-tat,’ ‘win-stay-lose-shift,’ ‘BB triggers A,’ and ‘suspicious punish once’, respectively are the ones with the high probability weights. The ‘win-stay-lose-shift’ strategy dictates that players will continue playing their current action if they receive a payoff higher than their opponent, and will shift to the other action if they receive a payoff lower than their opponent. If these strategies are played only among themselves, we expect to mostly observe only the efficient and fair outcome as both players continue to play action A.

Stag Hunt

For both players, strategies obtaining high limiting probability weights are ‘punish once,’ ‘tit-for-tat,’ ‘quick B, then A,’ ‘BA triggers A,’ and ‘BB triggers A’ respectively. If these five strategies are played amongst themselves, the observed history of actions will likely involve mutual cooperation in all periods.

After examining the strategies with high limiting probability weights from the first phase, I expect that efficient and fair outcomes are preferred in the ‘pre-experimental’ phase in the full information, symmetric payoff setting. The limiting probability weights for the other two treatments are provided in Appendix A5.

1.3.3 Comparison between simulations and experimental subjects

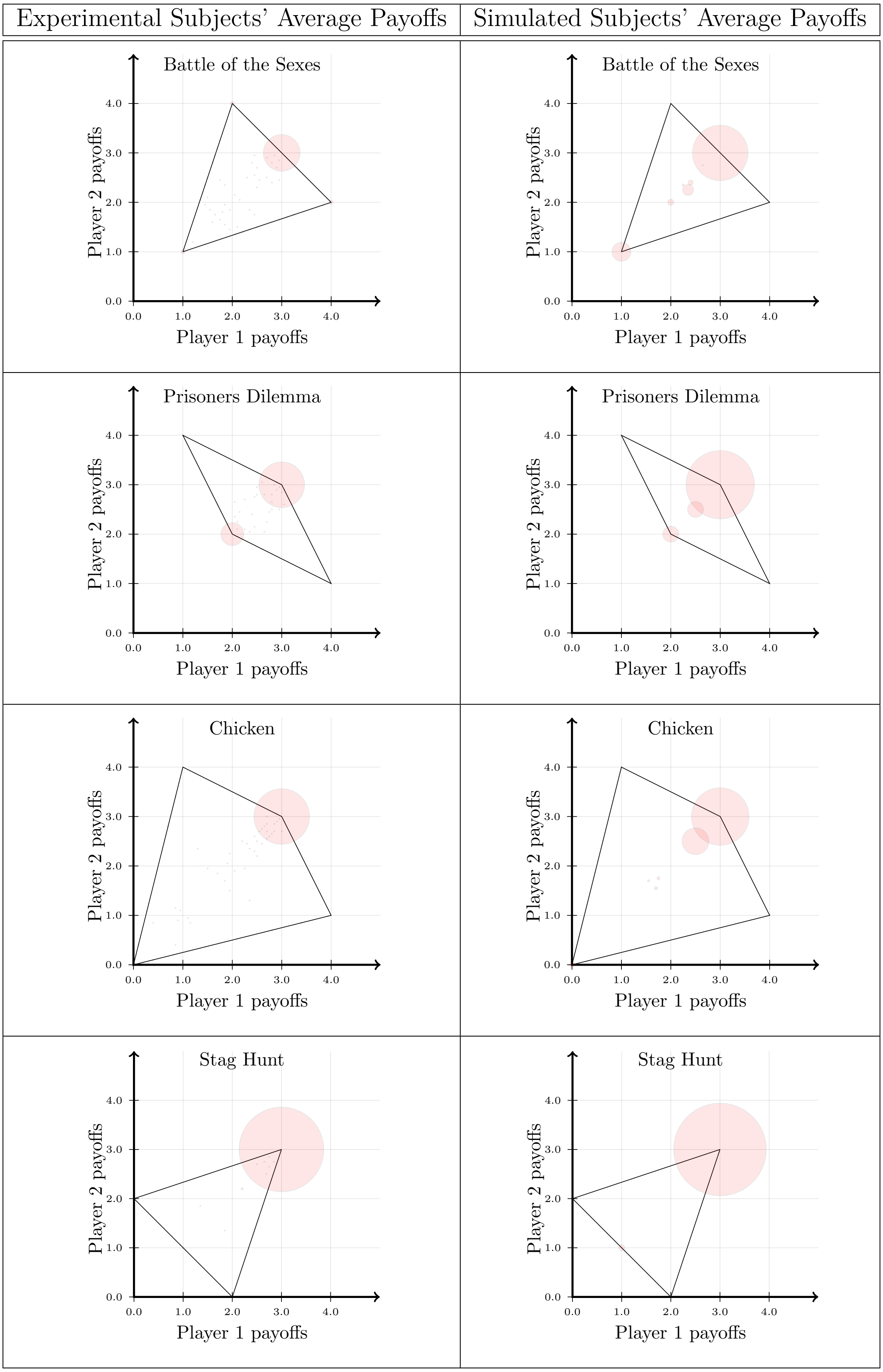
Full Information, Symmetric Payoffs

The experimental data has 70 observation pairs for the Battle of the Sexes game. For each observation pair, I note the number of rounds played, and simulate play for the same number of rounds between a 'simulated pair,' using a parameter group $(\alpha, \beta, \delta, \lambda, \gamma)$ to generate propensities towards strategies. I then calculate average payoffs for the experimental subject pair in the last 20 rounds, and also calculate average payoffs for the simulated subject pair in the last 20 rounds. Finally, I calculate the Euclidean distance between the two average payoffs. Note that for the Battle of the Sexes game, I have 70 experimental subject pairs and 70 simulated subject pairs. These pairs have to be matched to each other in a way that globally minimizes the sum of the Euclidean distance across all 70 pairs. I use the Kuhn-Munkres algorithm (also known as the Hungarian assignment problem algorithm) described in Kuhn (1955) to perform this objective.

The grid search over the different parameter groups $(\alpha, \beta, \delta, \lambda, \gamma)$ attempts to find parameters that minimizes this Kuhn-Munkres distance across all four games. The parameter values $(\alpha = 0.2, \beta = 0, \delta = 0.8, \lambda = 0.5, \gamma = 0.7)$ minimize the errors between average payoff points between experimental subject pairs and simulated subject pairs across **all four** games. Figure 1.8 below shows the heat map for the computational simulations and the experimental results with human subjects using parameter values $(\alpha = 0.2, \beta = 0, \delta = 0.8, \lambda = 0.5, \gamma = 0.7)$, and shows the relative frequency of each payoff combination over the last 20 periods of game-play. The relative frequency of a payoff combination is denoted by a circle located on the coordinates that correspond to that combination; the larger the circle, the higher the (relative) frequency of that combination. I also display the set of feasible payoffs.

The model slightly over-predicts the level of alternation between pure Nash equilibria outcomes in the Battle of the Sexes, and slightly underpredicts the mutual defection outcome in Prisoner's Dilemma, but for the most part, does a good job of predicting the outcome patterns of subjects in the four games.

Figure 1.8: Full Information, Symmetric Payoffs



The grid search shows that both efficiency and fairness measures are important, but selfishness is not a significant concern when attempting to minimize errors across all four games in this treatment. Parameter values that highlight selfishness concerns may perform well for one or two of the games individually, but perform very poorly when attempting to explain behavior in all four games simultaneously. Table 1.5 below present the error rate for different parameter combinations when placing either no weight, or all the weight, on the considerations of efficiency, fairness, and selfishness[8]. The error rates are significantly higher when efficiency, or fairness, measures are ignored. The closest fit to the experimental dataset places no weight on selfishness, and a much higher weight on fairness relative to efficiency.

Table 1.5: Full Information, Symmetric Payoffs Error Rate

	Efficiency (α)	Equality (β)	Fairness (δ)	Selfishness ($1-\alpha-\beta-\delta$)	QRE Scale (λ)	Gamma (γ)	Global Euclidean Distance
1	0	0	0.8	0.2	0.9	0.6	154
2	1	0	0	0	0.3	0.2	291
3	0	1	0	0	0.7	1	323
4	0	0	1	0	0.9	0.4	238
5	0	0	0	1	0.1	0	508
6	**0.2**	**0**	**0.8**	**0**	**0.5**	**0.7**	**133**

Full Information, Asymmetric Payoffs

The experimental data has 50 observation pairs for the Battle of the Sexes game. For each observation pair, I note the number of rounds played, and simulate play for the same number of rounds between a 'simulated pair,' using a parameter group $(\alpha, \beta, \delta, \lambda, \gamma)$ to generate propensities towards strategies. I then calculate average payoffs for the experimental subject pair in the last 20 rounds, and also calculate average payoffs for the simulated subject pair in the last 20 rounds. Finally, I calculate the Euclidean distance between the two average payoffs. Note that for the Battle of the Sexes game, I have 50 experimental subject pairs and 50 simulated subject pairs. These pairs have to be matched to each other in a way that globally minimizes the sum of

[8]Note that in the full information, symmetric payoff setting, absolute and relative fairness measures are the same, so we should think of $(\beta + \delta)$ as the combined weight on fairness

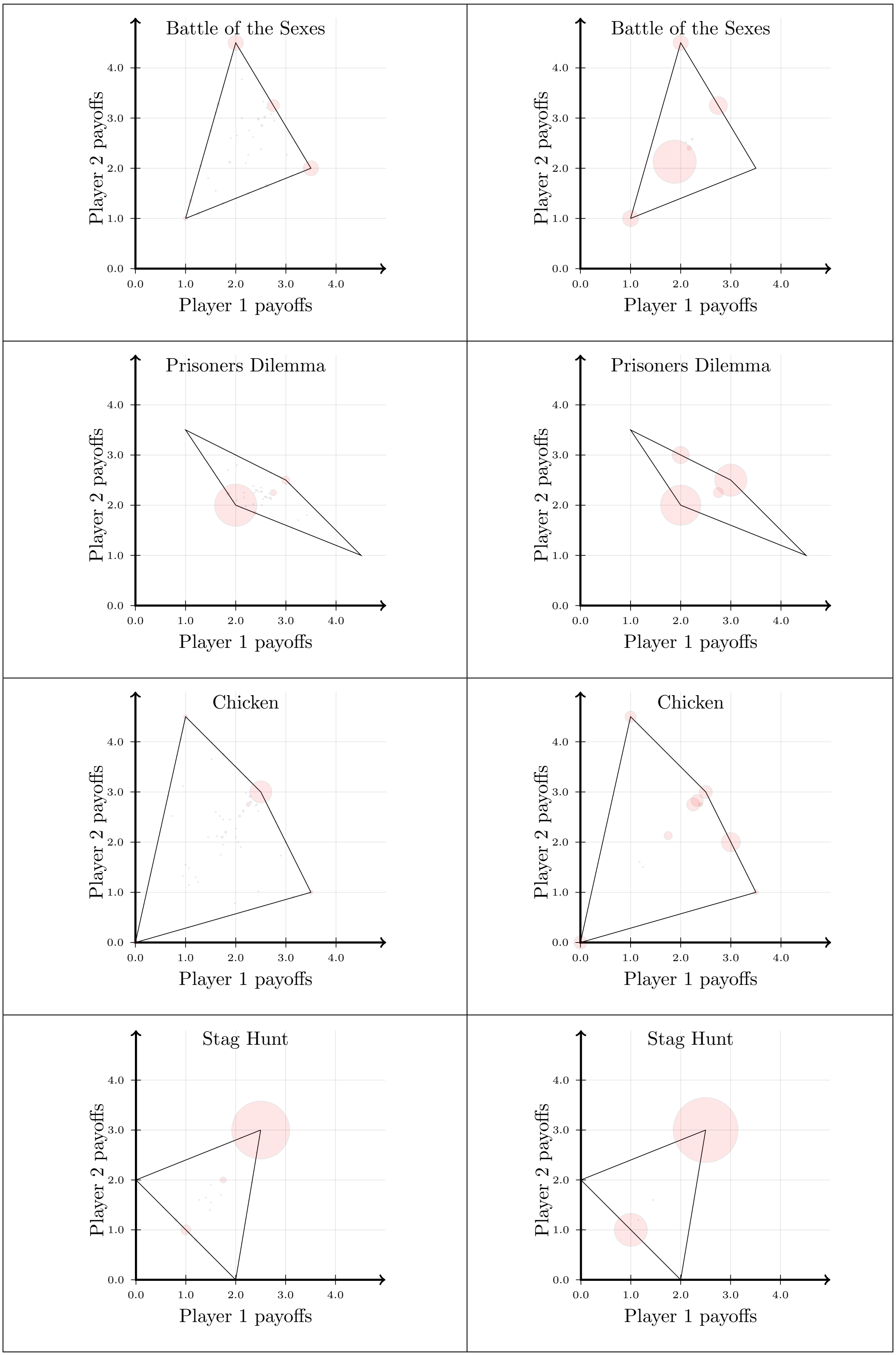
the Euclidean distance across all 50 pairs. I use the Kuhn-Munkres algorithm (also known as the Hungarian assignment problem algorithm) described in Kuhn (1955) to perform this objective.

The grid search over the different parameter groups $(\alpha, \beta, \delta, \lambda, \gamma)$ attempts to find parameters that minimizes this Kuhn-Munkres distance across all four games. The parameter values $(\alpha = 0.3, \beta = 0, \delta = 0.6, \lambda = 0.3, \gamma = 1.0)$ minimize the errors between average payoff points between experimental subject pairs and simulated subject pairs across all four games. Figure 1.9 below shows the heat map for the computational simulations and the experimental results with human subjects using parameter values $(\alpha = 0.3, \beta = 0, \delta = 0.6, \lambda = 0.3, \gamma = 1.0)$, and shows the relative frequency of each payoff combination over the last 20 periods of game-play. The relative frequency of a payoff combination is denoted by a circle located on the coordinates that correspond to that combination; the larger the circle, the higher the (relative) frequency of that combination. I also display the set of feasible payoffs.

The experimental data for Battle of the Sexes is spread out pretty uniformly over the entire convex hull of potential payoffs, and the model does a poor job of capturing this, and overpredicts payoff outcomes close to uniform randomization. The model also slightly overpredicts the mutual cooperation pattern in Prisoner's Dilemma. However, the model does a good job of predicting a high rate of mutual defection in Prisoner's Dilemma, a small rate of mutual cooperation in Chicken, and both the mutual cooperation, as well as the alternation of the mis-coordination payoff outcomes in Stag Hunt.

Figure 1.9: Full Information, Asymmetric Payoffs

Experimental Subjects' Average Payoffs	Simulated Subjects' Average Payoffs



The grid search shows that efficiency, selfishness and fairness measures are all important, but equality is not a significant concern when attempting to minimize errors across all four games in this treatment. Parameter values that highlight equality concerns may perform well for one or two of the games individually, but perform very poorly when attempting to explain behavior in all four games simultaneously. Table 1.6 below present the error rate for different parameter combinations when placing either no weight, or all the weight, on the considerations of efficiency, equality, fairness, and selfishness. The error rates are significantly higher when efficiency, fairness, or selfishness measures are ignored. The closest fit to the experimental dataset places no weight on equality, and a much higher weight on fairness relative to efficiency and selfishness.

Table 1.6: Full Information, Asymmetric Payoffs Error Rate

	Efficiency (α)	Equality (β)	Fairness (δ)	Selfishness ($1-\alpha-\beta-\delta$)	QRE Scale (λ)	Gamma (γ)	Global Euclidean Distance
1	0	0.3	0.3	0.4	0.6	0.1	194
2	1	0	0	0	0.1	0.3	254
3	**0.3**	**0**	**0.6**	**0.1**	**0.3**	**1**	**164**
4	0	1	0	0	0.2	0.4	355
5	0.3	0.5	0	0.2	0.4	0.9	207
6	0	0	1	0	0	0.9	358
7	0.3	0.1	0.6	0	0.3	0.5	192
8	0	0	0	1	0	0	362

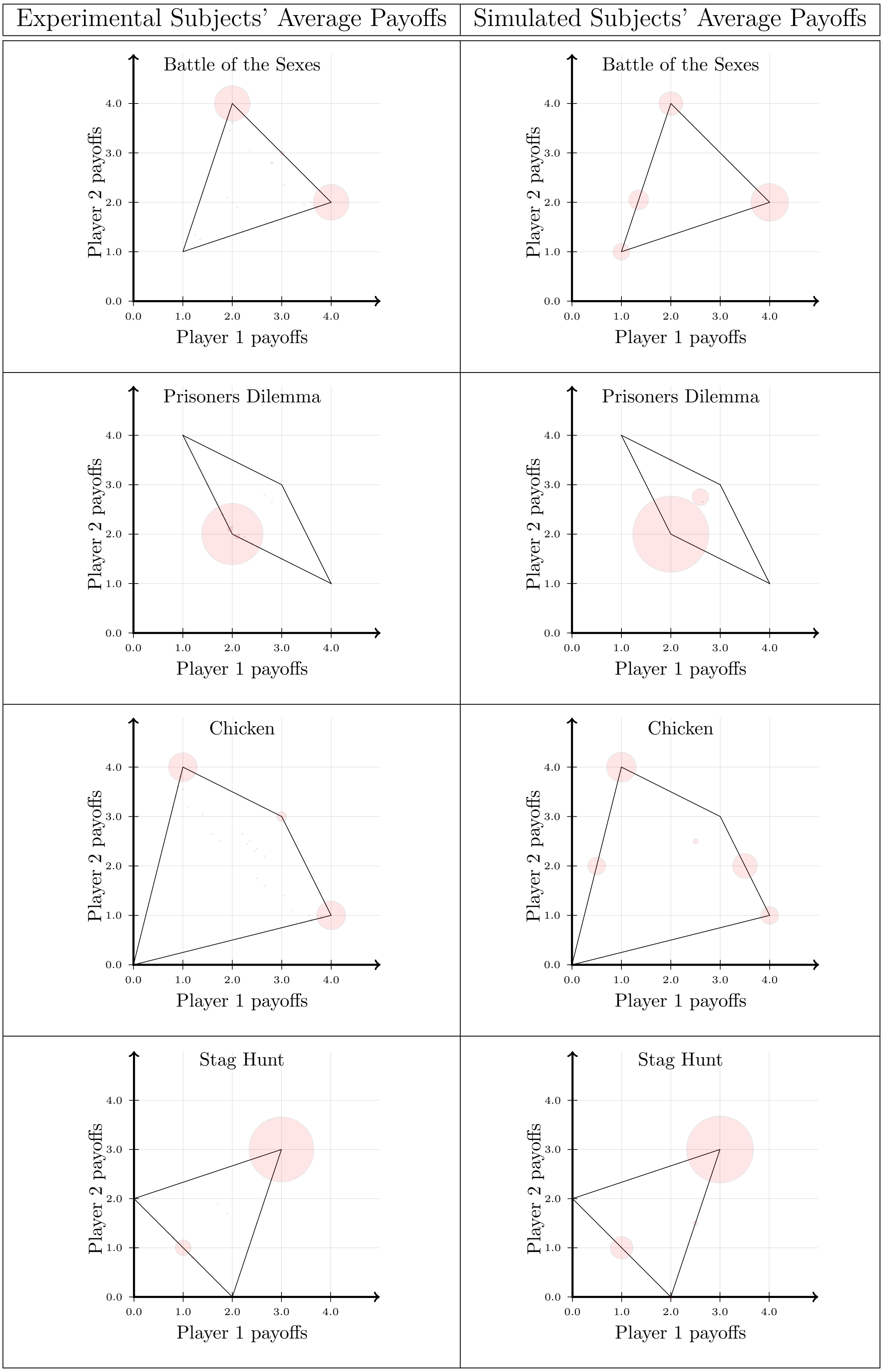
Dark Information, Symmetric Payoffs

The experimental data has 50 observation pairs for the Battle of the Sexes game. For each observation pair, I note the number of rounds played, and simulate play for the same number of rounds between a ‘simulated pair,’ using a parameter group $(\alpha, \beta, \delta, \lambda, \gamma)$ to generate propensities towards strategies. I then calculate average payoffs for the experimental subject pair in the last 20 rounds, and also calculate average payoffs for the simulated subject pair in the last 20 rounds. Finally, I calculate the Euclidean distance between the two average payoffs. Note that for the Battle of the Sexes game, I have 50 experimental subject pairs and 50 simulated subject pairs. These pairs have to be matched to each other in a way that globally minimizes the sum of the Euclidean distance across all 50 pairs. I use the Kuhn-Munkres algorithm (also known as the Hungarian assignment problem algorithm) described in Kuhn (1955) to perform this objective.

The grid search over the different parameter groups $(\alpha, \beta, \delta, \lambda, \gamma)$ attempts to find parameters that minimizes this Kuhn-Munkres distance across all four games. Since the matrices E, AF amd RF take on a value of zero in this treatment, the exact values of α, β and δ are not relevant, and the only parameter of interest is the weight on selfishness concerns $\mu = 1 - \alpha - \beta - \delta$. The parameter values $(\mu = 0.1, \lambda = 0.7, \gamma = 0.6)$ minimize the errors between average payoff points between experimental subject pairs and simulated subject pairs across all four games. Figure 1.10 below shows the heat map for the computational simulations and the experimental results with human subjects using parameter values $(\mu = 0.1, \lambda = 0.7, \gamma = 0.6)$, and shows the relative frequency of each payoff combination over the last 20 periods of game-play. The relative frequency of a payoff combination is denoted by a circle located on the coordinates that correspond to that combination; the larger the circle, the higher the (relative) frequency of that combination. I also display the set of feasible payoffs.

The model does a good job in predicting the mass of subjects ending up at the two pure Nash equilibria outcomes in the Battle of the Sexes, mutual defection in Prisoner's Dilemma, the two pure Nash equilibria outcomes in Chicken, and both the mutual cooperation, as well as the alternation of the mis-coordination payoff outcomes in Stag Hunt.

Figure 1.10: Dark Information, Symmetric Payoffs



The grid search shows that lower weights on selfishness perform better when attempting to minimize errors across all four games in this treatment. Table 1.7 below presents the error rate for different weights on selfishness. The error rates are significantly higher with higher weights on selfishness.

Table 1.7: Dark Information, Symmetric Payoffs Error Rate

	Selfishness $(1-\alpha-\beta-\delta)$	QRE Scale (λ)	Gamma (γ)	Global Euclidean Distance
1	0	0.6	0.8	255
2	**0.1**	**0.7**	**0.6**	**232**
3	0.2	0.3	0.4	239
4	0.3	0.3	0.6	272
5	0.4	0.2	0.3	259
6	0.5	0.3	0.7	292
7	0.6	0.2	0.5	329
8	0.7	0.2	0	361
9	0.8	0.2	0.4	368
10	0.9	0.2	0.5	413
11	1	0.1	0	445

1.4 Robustness Checks

The learning model described in the sections above captures tendencies that players have towards efficiency, absolute fairness (equality), relative fairness, and selfishness of long-run average payoffs in choosing initial attractions in the 'pre-experimental' phase before players actually start playing the game, and augments them by considering preferences where players react to stage-game payoffs received in every round to update tendencies towards strategies. During the rounds of the game, players get more and more attracted towards strategies that grant them higher stage-game payoffs, even if they initially started with 'good intentions', where they based their initial attractions towards strategies taking fairness and efficiency concerns into consideration. It is clearly a modelling choice to consider efficiency, equality, fairness, and selfishness concerns of long-run average payoffs to calculate initial attractions in the 'pre-experimental' phase, and then augment them with stage-game payoff related concerns in the 'experimental' phase. To ensure that this modelling choice is not the key driver behind the results in the following section, I also three alternate specifications.

- Standard QRE - This specification eliminates the efficiency, equality, fairness, and selfishness considerations, and instead only uses the long-run average payoffs for the players from following the automata strategies to generate attractions towards strategies. The rest of the model, with the fixed points as the ‘pre-experimental’ attractions augmented by stage-game payoffs in the ‘experimental’ phase, is kept exactly the same as the original model.

- No Second Stage - This specification eliminates the stage-game payoff related concerns in the ‘experimental’ phase. In this specification, only the ‘pre-experimental’ phase using efficiency, equality, fairness, and selfishness considerations of long-run average payoffs are used to generate attractions towards strategies.

- Attractions towards Actions - This specification uses attractions towards the two actions, rather than attractions towards the 26 repeated-game strategies. The rest of the model using efficiency, equality, fairness, and selfishness considerations of long-run average payoffs to generate attractions towards actions, augmented by stage-game payoffs in the ‘experimental’ phase, is kept exactly the same as the original model.

I compare all three alternate specifications against the findings of the original model for the Full Information, Symmetric Payoffs setting. I find that all three alternate specifications perform worse than the original model used, with only the ‘no second stage’ model coming close to the performance level of the original model. Table 1.8 shows the comparison between the global Euclidean distance for the different models. Thus, we can verify that three of the key ingredients of the model, namely 1) using attractions towards repeated-game strategies rather than attractions towards stage-game actions, 2) utilities that weigh considerations of efficiency, fairness, and selfishness of long-run average payoffs rather than simply using the long-run average payoffs, and 3) augmentation of stage-game payoffs to update attractions towards strategies, are all required to improve the accuracy of the overall model.

Table 1.8: Global Euclidean Distance Table

	Global Euclidean Distance
Standard QRE	293
No Second Stage	168
Attractions towards actions	268
Model	**133**

1.5 Conclusion

The experimental patterns about subject behavior in repeated games found in Mathevet and Romero (2012) show trends that are similar to many earlier studies such as Rapoport and Chammah (1965) and Rapoport et al (1976). Most learning models, especially the ones that follow an 'attractions towards actions' method to analyse initial attractions have had limited success in explaining phenomena such as the high degree of cooperation in the Prisoner's Dilemma and the alternation between two pure-strategy equilibria in the Battle of the Sexes game. As discussed earlier, an 'attractions towards strategies' method to analyse initial attractions is likely to have greater success. However, given the large number of possible initial strategies, an estimation of initial attractions can be challenging. Both Hanaki et al (2005) and Romero and Ioannou (2014) overcome this by restricting their analysis to strategies that can be analysed using one-state and two-state automata.

The learning model studied here follows a similar restriction on the strategy space, but uses weighted versions of efficiency, equality, fairness, and selfishness scores to evaluate attractions that players have towards strategies. The model adjusts naturally between symmetric and asymmetric payoffs, and also adjusts naturally if there are changes in the information structure of the game (e.g if players only know their own payoffs in the payoff matrix but do not know the payoffs of the opponent). Experimental treatments show that asymmetry in payoffs, or lack of information about the opponent's payoffs, leads to a much higher rate of mis-coordination among players. The learning model captures these trends by placing greater emphasis on efficient and fair strategies when considering experimental subjects in the full information, symmetric payoff setting, and placing more weight on selfishness in treatment settings where payoffs are asymmetric, or information is limited. The tendencies of experimental subjects in various information and payoff settings can only be explained by weighting the considerations of efficiency, equality, fairness, and selfishness differently. Depending on the context, different strategies become focal in the minds of subjects, and any one of these considerations by itself fails to account for many patterns seen in laboratory results across different information or payoff treatments.

One of the remaining challenges is isolating the motivations behind individual actions in these 2 X 2 games. For example, players mutually playing the cooperative action in Chicken leads to payoff outcomes that are simultaneously efficient, equal, and fair. Similarly, players mutually cooperating in Stag Hunt also leads to payoff outcomes that are simultaneously efficient, equal, fair, and selfish. It is difficult to disentangle the motivation that preempted a particular action taken by a player in this class of games. It will be interesting to study this learning model in a setting with more actions, where each action highlights each of these motivations individually. I am aiming to study this in greater detail in future work.

Chapter 2

Infinitely Repeated Games : Selfishness, Efficiency, or Fairness?

This paper investigates a learning model where players have attractions towards efficiency, fairness and selfishness in the payoff outcomes of an infinitely repeated two-player three-action game. The payoffs are chosen so that one of the players' actions emphasizes their attempt to maximize their own selfish payoffs, another action emphasizes the maximization of joint payoffs, and the third action emphasizes payoff parity between both players. We consider both an attractions towards stage-game actions, and an attractions towards repeated-game strategies learning model to ascertain the relative weight that players place on these motivations when making decisions in infinitely repeated games. We find that 32% of all subject pairs settle on the Nash equilibrium outcome, while another 12% of subjects pick outcomes that equalize payoffs between subjects. The learning models used here show that players value notions of efficiency and fairness more highly than selfishness.

JEL Codes: C51, C72, C92, D03, D83.

Keywords: Reinforcement learning, Repeated game strategies, Adaptive models, Finite automata.

2.1 Introduction

Infinitely repeated games have been studied in a wide variety of settings. Games such as Prisoner's Dilemma, Battle of the Sexes, Chicken, and Stag Hunt, which have different kinds of cooperative and non cooperative equilibria, have been studied extensively. Such games help capture important aspects of everyday experiences such as cooperation, coordination and reciprocity. While the Folk Theorem shows that a broad range of payoffs can be supported as equilibrium outcomes in such infinitely repeated games given that players are sufficiently patient, a host of experimental studies show that players usually coordinate on only a few of the feasible payoffs. Some of the payoff outcomes are 'focal', and players use certain criteria as an equilibrium selection device to arrive at these payoff outcomes. The identification of these criteria is an open research question. In this paper, I posit efficiency (maximization of joint payoffs), fairness (parity in payoffs), and selfishness (maximization of personal payoffs) as potential criteria that players use when selecting equilibrium outcomes. Using a three-parameter model, I attempt to identify the relative weight that players place on these three criteria when choosing stage-game actions, or repeated-game strategies, in an infinitely repeated game.

In order to make these criteria salient in the minds of players, I construct a two-player three-action game where each action highlights one of the three concerns: A 'selfish' action that maximizes each players' individual payoffs relative to the other two actions, a 'fair' action that equalizes the payoffs for both players, and an 'efficient' action that maximizes the joint payoffs for both players, relative to the other two actions. Table 2.1 below shows the payoffs for all three actions available to both players. Given the construction of the payoffs, the game has a single Nash equilibrium where both players play the 'selfish' action. A key point of note is that for Player 1, her 'efficient' action is more unfair than her 'selfish' action because it generates a higher mismatch in payoffs between the two players. On the other hand, for Player 2, her 'selfish' action is more unfair than her 'efficient' action. This highlights the asymmetry in the incentives for the two players as they progress through the game.

Table 2.1: Selfishness, Efficiency and Fairness Game Payoff Matrix

	S	F	E
S	40, 30	25, 25	50, 22
F	35, 35	20, 20	35, 35
E	32, 40	25, 25	45, 35

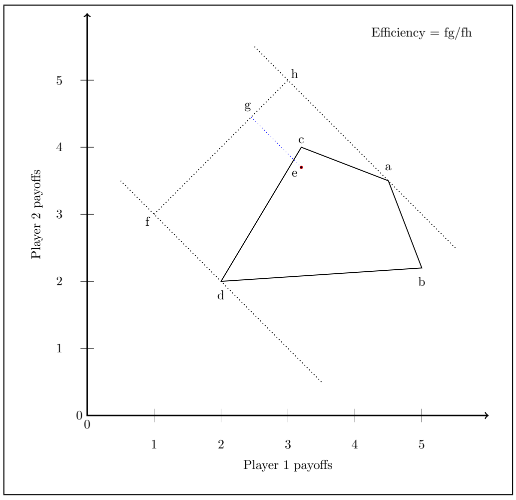
SFE Game

2.2 Efficiency, Fairness, and Selfishness

2.2.1 Efficiency

The efficiency measure utilizes the payoffs of both players. If both players choose their most efficient action, their joint payoffs are maximized at 80, denoted on the diagram below as point (a). On the other hand, point (d) represents the most inefficient joint payoff for both players, with a joint payoff of 40, when both players choose the fair action. Graphically, we can envision a series of iso-efficiency indifference curves perpendicular to a 45 degree line from the origin. Figure 2.1 shows the iso-efficiency indifference curves passing through the least, and most, efficient payoff outcomes in this generic game. Payoff outcomes on the iso-efficiency indifference curve that passes through the point that minimizes (maximizes) the sum of payoff outcomes for both players receive an efficiency score of zero (100). For any generic payoff outcome (such as point 'e' in Figure 2.1), we can define the iso-efficiency indifference curve that it lies on. We then calculate the distance from this curve to the most efficient indifference curve, relative to the distance between the two extreme indifference curves, and generate an efficiency score on a scale of zero to 100.

Figure 2.1: Efficiency Score

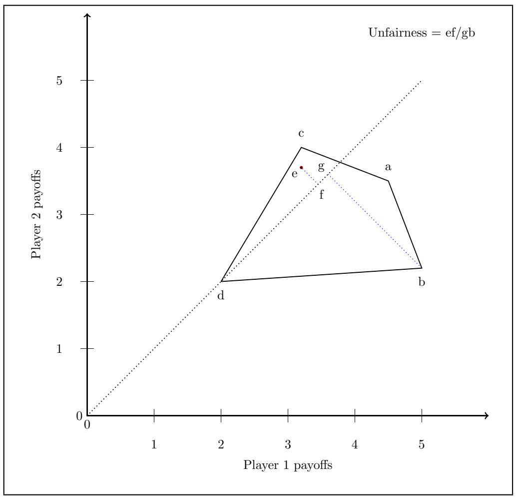


2.2.2 Fairness

The fairness measure utilizes the payoffs of both players. Actions that perfectly equalize payoffs between players, represented by the 45 degree line, are considered fair. On the other hand, when Player 1 chooses her selfish action and Player 2 chooses her efficient action, the realized payoffs are the most unequal. The distance from this outcome, represented as point (b) in Figure 2.2 below, to the 45 degree line represents the largest level of unfairness that can be achieved. The unfairness measure is the ratio of the distance of any payoff combination to the 45 degree line, divided by the distance of point (b) from the 45 degree line. The fairness measures are calculated on a scale from 0 to 100, with point (b) representing a fairness score of 0, and points on the 45 degree line representing a fairness score of 100.

We use a notion of fairness where outcomes that equalize payoffs are regarded as being perfectly fair, and the payoff outcome within the range of possible stage-game payoffs with the greatest level of inequality labeled as the most unfair outcome. All other payoff outcomes are measured in relative terms between these two extremes. The idea that fairness concerns are partly based on the range of possible outcomes can be justified by patterns seen in Dictator games studied in List (2007), where introducting the possibility of taking away money from receivers made Dictators less generous in their giving. In that one-shot non-strategic setting, expanding the payoff possibility space changes which payoffs agents come to regard as fair. Here, we extend this notion to a repeated game with strategic interactions.

Figure 2.2: Fairness Score



2.2.3 Selfishness

The selfishness measure for each player only considers their own payoff. For each player, the selfishness measure is the ratio of the difference between their received payoff and the minimum of all their possible payoffs, divided the difference between the maximum and the minimum of all their possible payoffs. Selfishness measures are calculated on a scale from 0 to 100, with point (b) representing a selfishness score of 100 for Player 1, and point (c) representing a selfishness score of 100 for Player 2.

Figure 2.3: Selfishness Score

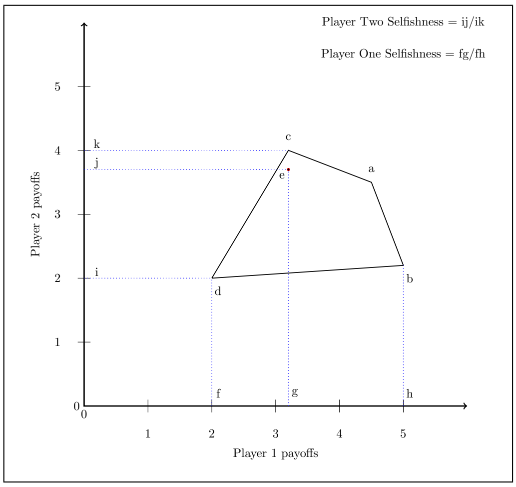


Table 2.2 below shows the matrix, with payoffs transformed into efficiency, fairness, and selfishness scores.

Table 2.2: Selfishness, Efficiency and Fairness Game Transformed Matrix

	S	F	E
S	75	25	80
F	75	0	75
E	80	25	100

(a) Efficiency

	S	F	E
S	64.3	100	0
F	100	100	100
E	71.4	100	64.3

(b) Fairness

	S	F	E
S	66.7	16.7	100
F	50	0	50
E	40	16.7	83.3

(c) Player One Selfishness

	S	F	E
S	50	25	10
F	75	0	75
E	100	25	75

(d) Player Two Selfishness

2.3 Attractions towards actions

In this section, I describe the learning model that follows an 'attractions towards actions' approach, where an algorithm is used to generate initial attractions towards each of a finite set of actions, and these attractions are updated during repeated rounds of game-play.[1] The learning taking place in two stages. The 'pre-experimental' phase generates the initial attractions towards actions, while the 'experimental' phase updates the attractions during rounds of game-play. The exact way in which this learning occurs, and how attractions are updated, are presented in the following sections. Note that there are three possible actions available to each player in the game studied here.

2.3.1 Initial attractions towards actions

The learning model helps identify and update the propensities that players have towards each of the three actions available to them through the course of the game. For the ease of exposition, let M_1 and M_2 denote the sets of three actions available to Players 1 and 2 respectively.[2] For all nine potential action pairs, the payoffs are

[1] The learning model used here is similar to the one used in Bose (2018).

[2] This is only for the sake of exposition. Both players actually have access to the same set of actions, $\{S, F, E\}$.

transformed using the formulae for efficiency, fairness and selfishness. The transformed 3×3 matrices of efficiency, fairness and selfishness indices are combined to form a composite matrix, denoted as H_1 for Player 1 where the rows denote the actions of Player 1 and the columns denote the actions of Player 2. Similarly, I also construct a 3×3 matrix of composite scores where the rows denote the action choices of Player 2 and the columns denote the action choices of Player 1, denoted as H_2.

The elements in $H_1$where Player 1 chooses action i and Player 2 chooses action j, is given by:

$$H_{1,i,j} = \alpha E_{1,i,j} + \beta F_{1,i,j} + (1 - \alpha - \beta) S_{1,i,j}$$

The parameter α captures the propensity that players have towards efficient actions. The parameter β captures the propensity that players have towards fair actions, while the parameter $(1 - \alpha - \beta)$ captures propensitives towards selfish actions. For a given parameter value of α and β, the matrices H_1 and H_2 are stationary. Let $p_1 = \{p_{1,1}, p_{1,2}, p_{1,3}\}$ and $p_2 = \{p_{2,1}, p_{2,2}, p_{2,3}\}$ denote the initial probability distribution (propensities) that Player 1 and Player 2 have towards playing the three actions they have access to. If Player 1 believes that Player 2 is playing actions based on a probability distribution $p_2 = \{p_{2,1}, p_{2,2}, p_{2,3}\}$, then the expected composite score for playing action i for Player 1 is given by:

$$\bar{H}_{1,i} = \sum_j (H_{1,i,j} * p_{2,j}).$$

Expected utilities for players are constructed from these expected composite scores. I assume (as in the Quantal Response Equilibrium (QRE) model discussed in McKelvey and Palfrey (1998)) that the bounded rationality of players leads to 'errors' in the calculation of these expected composite scores, such that $\bar{u}_{1,i} = \bar{H}_{1,i} + \epsilon_{1,i}$.[3] I further

[3]There are a number of interpretations of these 'errors' as discussed in Haile, Hortacsu, and Kosnenok (2007). It is not necessary to take these 'errors' in composite score calculations literally. It is simply possible that players use actions that are 'arbitrarily close' to a best-response rather than 'always fully optimal.' This interpretation has an intuitive appeal, especially given the challenging assumption of completely rational expectations about opponents' behavior that exist in all optimal best-response game-theoretic models. These deviations from 'always fully optimal' best-responses can also be viewed as a way to 'smooth out' best-response functions so as to obtain more robust predictions as in Rosenthal (1989). McKelvey and Palfrey (1995) suggest that these 'errors' have natural economic interpretations. These 'errors' can simply reflect unmodeled costs of information processing involved in calculating 'always fully optimal' best-response actions. It is also possible that the 'errors' reflect unmodeled determinants of the players' utilities. This interpretation is appealing

assume that the errors $\epsilon_{1,i}$ are identical and independently distributed and have an EV(1) distribution. In addition, I have the behavioral assumption that players choose actions that have higher expected composite scores. Thus, Player 1 has a behavioral tendency to select an action i if $\bar{u}_{1,i} \geqq \bar{u}_{1,k} \forall k \in M_1$. Similarly, Player 2 has a behavioral tendency to select an action j if $\bar{u}_{2,j} \geqq \bar{u}_{2,k} \forall k \in M_2$. Given the EV1 distribution of the errors, for an initial (p_1, p_2), players can recalculate their propensities towards their actions using the following formulae:

$$\tilde{p}_{1,i} = \frac{exp(\lambda \,.\, \bar{u}_{1,i})}{\sum_{i \in M_1} exp(\lambda \,.\, \bar{u}_{1,i})} \text{ and } \tilde{p}_{2,j} = \frac{exp(\lambda \,.\, \bar{u}_{2,j})}{\sum_{j \in M_2} exp(\lambda \,.\, \bar{u}_{2,j})} \ ^{45}$$

The parameter λ, which is greater than or equal to zero, captures the scaling effect, and sensitivity towards the composite scores. In this logistic transformation, it represents the extent to which actions with higher attractions are favored in action choice. When $\lambda = 0$, all actions are equally likely to be chosen, regardless of their attractions. As λ increases, actions with higher attractions become disproportionately more likely to be chosen. In the limiting case $\lambda \to \infty$, the action with the highest attraction is chosen with probability one.

Thus, for any initial point (p_1, p_2) in the $[0,1]^3$X $[0,1]^3$ simplex that satisfies $\sum_i p_{1,i} =$ 1and $\sum_j p_{2,j} = 1$, I have a mapping that generates a new point $(\tilde{p}_1, \tilde{p}_2)$ in the same simplex. The mapping routine continues until the following convergence criterion for the simulations is satisfied.

$max_{i \in A}|p_{1,i}(m) - p_{1,i}(m-1) < \epsilon$ and $max_{j \in A}|p_{2,j}(m) - p_{2,j}(m-1) < \epsilon$ for $\epsilon = 0.0001$ for iteration m.

While the existence of the fixed point is guaranteed[6], in principle, it is possible to have a simulation run that does not satisfy the convergence criterion if ε is very small.

in many applications since a fully specified theoretical model can, of course, only approximate a real economic environment.

[4]I rescale values such that $max_j \lambda(\sum_i H_{2,j,i} * p_{1,i})$ and $max_i \lambda(\sum_j H_{1,i,j} * p_{2,j})$ has a value of 700, because the algorithm cannot calculate values above exp(700).

[5]It is an empirical question where the logit, probit, or power forms fit better (adjusting for degrees of freedom). Previous studies show roughly equal fits of logit and power (Tang (1996), Chen and Tang (1998), Roth and Erev (1997), or better fits for the logit form than the power form (Camerer and Ho (1998)).

[6]As shown in Appendix B3

However, I obtained convergence in all cases, with one caveat. For some parameter groups (α, β, λ), the covergence algorithm reaches an alternation between two points, back and forth in succession. For the sake of exposition, let us label these points $(\hat{p}_1, \hat{p}_2)$ and $(\check{p}_1, \check{p}_2)$. For all such situations, I introduce a parameter $\gamma \in [0, 1]$, and consider linear combinations of the two probability distributions as the final fixed point. So, the final fixed points I consider are $(\mathring{p}_1, \mathring{p}_2) = \gamma(\hat{p}_1, \hat{p}_2) + (1 - \gamma)(\check{p}_1, \check{p}_2)$, normalized such that $(\mathring{p}_1, \mathring{p}_2)$ still satisfies the properties of a probability distribution. I evaluate possible values of γ ranging from $\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.8, 1.0\}$, thus considering linear interpolations of the two probability distributions between which the algorithm is cycling.

The initial starting point for (p_1, p_2) is a uniform distribution across all three actions for both players. The converged fixed point when starting from a uniform distribution probability distribution for both players represents the initial attractions that players have towards actions, developed in the 'pre-experimental' phase of game-play. These fixed points are similar to a QRE, but with a restricted action-space.

2.3.2 Updating attractions towards actions

For a repeated game setting, it is important to model how initial attractions are updated after each round. In any appropriate learning model, players must update their attractions in a way to best-respond to the outcomes of observed play. To incorporate a sense of 'best-response' to observed play during the different rounds, I add the payoffs generated from each round to the composite scores being used to calculate attractions towards actions.

Before the first round, Player 1 has a 3×3 composite score matrix H_1 and Player 2 has a 3×3 composite score matrix H_2. Let $\ddot{p}_1 = \{\ddot{p}_{1,1}, \ddot{p}_{1,2}, \ddot{p}_{1,3}\}$ and $\ddot{p}_2 = \{\ddot{p}_{2,1}, \ddot{p}_{2,2}, \ddot{p}_{2,3}\}$ denote the converged fixed point probability distribution (propensities) that Player 1 and Player 2 have towards playing the three different actions. The expected composite score for playing action i (action j) for Player 1 (Player 2), is given by:

$\tilde{H}_{1,i} = \sum_j (H_{1,i,j} * \ddot{p}_{2,j})$ and $\tilde{H}_{2,j} = \sum_i (H_{2,j,i} * \ddot{p}_{1,i})$.

In the repeated game, for any round, let Player 1 choose action i and let Player 2 choose action j, and let the payoff for both players be $a_1$and a_2 respectively. I update the expected composite score for Player 1's action i by adding in the stage-game payoff: $\tilde{H}_{1,i} = \tilde{H}_{1,i} + a_1$, while all the expected composite scores for the other action remain unchanged. Thus, the expected composite score for the action *that the player chose in a particular round* is updated to reflect the effects of that choice (i.e the payoffs), while the expected composite scores for all other unchosen actions remain unchanged. Similarly, for player 2, I update the expected composite score for Player 2's action j by adding in the payoff: $\tilde{H}_{2,j} = \tilde{H}_{2,j} + a_2$, while all the expected composite scores for the other actions remain unchanged. After every round, attractions towards actions in the next round are then re-calculated using the logistic algorithm described in the earlier section. Updating propensities in this way means that if the payoffs from choosing a particular strategy in a particular round are larger, then it increases the likelihood of players choosing that strategy in future rounds by a larger amount. Draws taken from these constantly updated propensities are used to simulate play between players.

The initial propensity that players have towards their actions is a tuple (p_1, p_2) in the $[0,1]^3$X $[0,1]^3$ simplex that satisfies $\sum_i p_{1,i} = 1$and $\sum_j p_{2,j} = 1$. After every round, I have a mapping that generates a new propensity $(\tilde{p}_1, \tilde{p}_2)$ in the same simplex. Thus, the attractions that players have towards each of their three actions are being updated after every round of game play.

The learning model described in the sections above captures tendencies that players have towards efficiency, fairness and selfishness in choosing actions during the course of a repeated-game. During the rounds of the game, players get more and more attracted towards actions that grant them higher payoffs.

2.3.3 Simulated subjects (Grid Search)

The learning algorithm described above is used to simulate play between pairs of players. I run a grid search over different parameter values, with

$\alpha \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

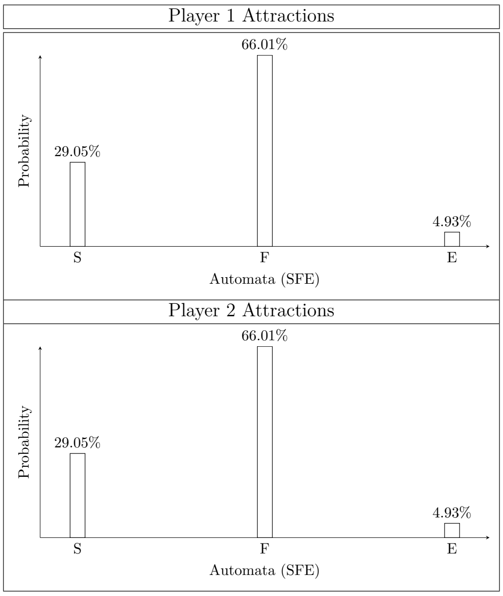
$\beta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

$\lambda \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

and $\gamma \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$.

For most parameter values, the converged fixed points for the initial attractions towards actions follow a pattern where most of the probability weight is on a subset of the three actions. Figure 2.4 shows the histogram of the probability distribution for the converged initial attractions towards actions for both players, when considering parameter values ($\alpha = 0.1, \beta = 0.9, \delta = 0.0, \lambda = 0.1, \gamma = 0.9$).

Figure 2.4: Converged initial attractions for the two players in the SFE game (Actions)



2.3.4 Simulated subjects (Maximum Likelihood Estimation)

In addition to using the grid-search technique to identify parameters that best describe the experimental data patterns (described in later sections), I also use a maximum likelihood estimation model. The maximum likelihood estimation method for estimating the α, β, λ and γ parameter values in the attractions towards actions model will include the following steps:

1. Players are denoted as $i \in \{1, 2, ..., N\}$, where $-i$ denotes the matched partner for player i.

2. Each player has J_i actions available to her, with actions denoted as $j \in J_i$.

3. There are $t \in \{1, 2, ..., T\}$ total rounds

4. We are attempting to identify the likelihood of Player i choosing action j in round t.

5. The log likelihood function is given by $log\ L(\lambda, \alpha, \beta) = \sum_{i=1}^{N} \sum_{j=1}^{J_i} \sum_{t=T-20}^{T} log(\sigma^*_{i,j,t}(\lambda, \alpha, \beta))^{1(s_{i,j,t})}$ where $\sigma^*_{i,j,t}(\lambda, \alpha, \beta)$ is the probability that Player i chooses action j in round t, and the indicator function ensures that only actions that were actually chosen are added when calculating the log likelihood, and we only consider the likelihoods for the last twenty rounds, to keep consistent with the grid search method.

6. The probability that Player i chooses action j in round t is given by the multinominal logit probability distribution adapted from the QRE in McKelvey and Palfrey (1998). $\sigma^*_{i,j,t}(\lambda, \alpha, \beta) = \frac{exp\,\lambda(\bar{x}_{i,j,t})}{\sum_{k=1}^{J_i} exp\,\lambda(\bar{x}_{i,k,t})}$

7. The average utilities for Player i choosing action j in round t is given by $\bar{x}_{i,j,t} = \sum_{k=1}^{J_{-i}} A^k_{i,j,t} \,.\, \sigma^*_{-i,k,t-1}(\lambda, \alpha, \beta)$ where $A^k_{i,j,t}$ is the attraction that Player i has towards action j in round t when Player i's matched partner plays action k, and $\sigma^*_{-i,k,t-1}(\lambda, \alpha, \beta)$ is the probability that Player i's matched partner played action k in the previous round. Let $A_{i,j,t}$ denote the vector of attractions that that Player i has towards action j in round t across all k possible actions played by Player i's matched partner.

8. The attractions that Player i has towards action j in round t is given by $A_{i,j,t} = A_{i,j,t-1} + 1(s_{i,j,t-1}) \,.\, \pi_i(s_{i,j,t-1}, s_{-i,k,t-1})$ where attractions that Player i has towards action j in round t is the same as the attraction that Player i has towards action j in the previous round, unless Player i selected action j in the previous round, in which case the payoff received by Player i in the previous round is added to the attraction for action j from the previous round. The indicator function indicates that Player i chose action j in the previous round, while the function $\pi_i(.)$ denotes the payoff that Player i received in the previous round after choosing action j.

9. Since attractions towards each action in every round is dependent on the attrac-

tions towards each action in the previous round, we must define the initial attractions formed before the first round. The attraction that Player i has towards action j at the start of the game is given by $A_{i,j,0} = \alpha\bar{E}_{i,j,0} + \beta\bar{F}_{i,j,0} + (1-\alpha-\beta)\bar{S}_{i,j,0}$ where the parameters α and β capture the relative weight that players give to the average efficiency and fairness metrics, relative to the average selfishness metric.

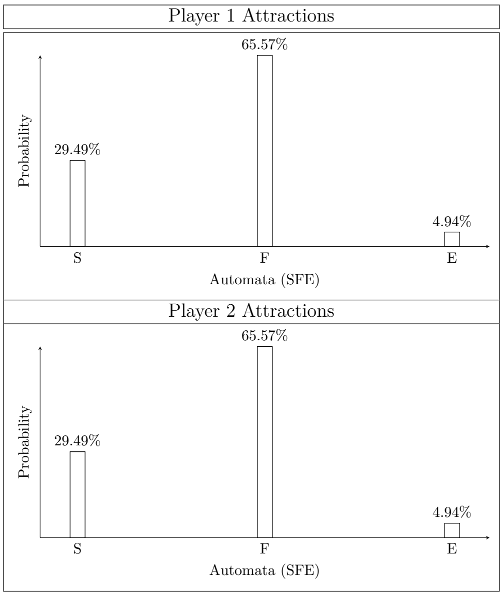
10. The average efficiency and fairness metrics are determined in the following way: $\bar{E}_{i,j,0} = \sum_{k=1}^{J_{-i}} E^k_{i,j,0} \cdot \sigma^*_{i,j,0}(\lambda, \alpha, \beta)$ where $E^k_{i,j,0}$ is the efficiency score for Player i choosing action j when her matched partner chooses action k, and $\sigma^*_{i,j,0}(\lambda, \alpha, \beta)$ is the attraction that Player i's matched partner has towards playing action k at the start of the game. Similarly, $\bar{F}_{i,j,0} = \sum_{k=1}^{J_{-i}} F^k_{i,j,0} \cdot \sigma^*_{i,j,0}(\lambda, \alpha, \beta)$ and $\bar{S}_{i,j,0} = \sum_{k=1}^{J_{-i}} S^k_{i,j,0} \cdot \sigma^*_{i,j,0}(\lambda, \alpha, \beta)$ provide similar analogs for fairness and selfishness metrics.

11. I assume that both players assume that their opponent is using a uniform distribution as starting probabilities across all three actions available to them. Thus $\sigma^*_{i,j,0}(\lambda, \alpha, \beta)$ is just $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

12. The values of $E^k_{i,j}$, $F^k_{i,j}$, and $S^k_{i,j}$ are different depending on whether Player i is a row player or a column player, as displayed in Table 5.

Figure 2.5 shows the histogram of the probability distribution for the initial attractions towards actions for both players that has the highest maximum likelihood when matched against the experimental data, with parameter values ($\alpha = 0.11, \beta = 0.89, \delta = 0.00, \lambda = 0.10, \gamma = 1.00$).

Figure 2.5: Converged initial attractions for the two players in the SFE game (Actions MLE)



2.4 Attractions towards strategies

In this section, I describe the learning model that follows an 'attractions towards strategies' approach similar to Hanaki et al (2005), and Romero and Ioannou (2014), where an algorithm is used to generate initial attractions towards each of a finite set of strategies, and these attractions are updated during repeated rounds of game-play. The learning taking place in two stages. The 'pre-experimental' phase generates the initial attractions towards strategies, while the 'experimental' phase updates the attractions during rounds of game-play. The exact way in which this learning occurs, and how attractions are updated, are presented in the following sections. The finite set of strategies (represented by automaton or Moore machines) is described below.

2.4.1 Finite Automata

A finite automaton is a mathematical model of a system with discrete inputs and outputs, and can be used to represent repeated-game strategies. The system can be in any one of a finite number of internal configurations or "states," similar to the states in a repeated-game strategy. The state of the system summarizes the information concerning past inputs that is needed to determine the behavior of the system going forward. The specific type of finite automaton used in this paper is a Moore machine. A Moore machine for player i, M_i, in a repeated game $G = (I, \{A_i\}_{i\in I}, \{g_i\}_{i\in I})$ is a four-tuple $(Q_i, q_{0,i}, f_i, \tau_i)$ where Q_i is a finite set of internal states of which $q_{0,i}$ is specified to be the initial state, $f_i : Q_i \to A_i$ is an output function that assigns an action to every state, and $\tau_i : Q_i \times A_{-i} \to Q_i$ is the transition function that takes a players's current state and the opponent's action, and returns the next state in the player's system. It is pertinent to note that the transition function depends only on the present state and the other player's action. This formalization fits the natural description of a strategy as i's plan of action in all possible circumstances that are consistent with i's plans. In contrast, the notion of a game-theoretic strategy for i requires the specification of an action for every possible history, including those that are inconsistent with i's plan of action. It is important to highlight that to formulate the game-theoretic notion of a strategy, one would only have to construct the transition function so that $\tau_i : Q_i \times A \to Q_i$, instead of $\tau_i : Q_i \times A_{-i} \to Q_i$. In the first period, the state is $q_{0,i}$, and the automaton chooses the action $f_i(q_{0,i})$. If a_{-i} is the action chosen by the other player in the first period, then the state of i's automaton changes to $\tau_i(q_{0,i}, a_{-i})$, and in the second period, i chooses the action dictated by f_i in that state. Then, the state changes again according to the transition function given the other agent's action. Thus, whenever the automaton is in some state q, it chooses the action $f_i(q)$, while the transition function τ_i specifies the automaton's transition from q to a new state in response to the action taken by the other player.

2.4.2 Restricting the strategy space

Even though the action-space being considered in the game is finite, allowing for mixed strategies gives us a very large number of potential strategies that can be chosen in these games. For any 'attractions towards strategies' learning model to have predictive power, the space of potential strategies allowed has to be limited. One way to restrict strategies is only considering those that can be represented as a Moore machine, and in this paper, I only consider a subset of strategies that can be represented by three-state, three-action automaton.

I consider automata where there are three possible states and three possible actions by the opponent. The transition function thus maps twelve different possibilities of (own state, the opponent's action) pairs into the set of three states to be taken in the next period. This generates $3^{12} = 531{,}441$ cases in total. Since each of these 531,441 cases can have one of three initial states, there are 1,594,323 possible automata. A large number of these are non-unique since they prescribe the same set of outcomes under any combination of opponent's actions. Eliminating non-unique automata leaves us with a total of 173,913 unique possible automata.

Constructing a probabilility distribution that measures the attractions that players have towards each of 173,913 possible repeated-game strategies (modeled as automata) is both analytically difficult, and also does not provide any significant weights for any automata, thereby providing us with little insight into the strategies that players are actually attracted towards. I use a clustering technique to identify a smaller set of automata that are representative of the behavior of the larger set of automata. The following sections describe the clustering process used to identify this smaller set of representative automata.

Similarities between automata

The overall goal in any clustering technique is to maximize the similarities between each automata and the exemplar (or cluster centroid) it is assigned to, so that each automata is assigned to a cluster to which it is most 'similar'. A natural question in this context is how are we defining similarities? If we look at data points that can be visualized in R_n, and can construct a metric that helps capture the notion of 'distance' between data points, then the similarity could simply be the negative value of this 'distance'. For example, the goal of the clustering technique can be set to minimize the Euclidean distance between each data point and the exemplar it is assigned to. In this case, the similarity is equal to the negative of the Euclidean distance between each data point and the exemplar it is assigned to. Here, we do not have a metric that denotes 'distance' between automata. Instead, we use an input string of actions taken by the opponent, and have each automata 'play' against this input string to generate a series of output actions recommended by the automata when faced by this input string of opponent actions. Since we are considering three-state, three-action automata here, an input string of length nine is required to ensure that the automata will traverse all its three states. An input string of length nine will generate an output string of length ten, with the first element displaying the action recommended by the output function of the start state, and the following nine elements displaying the results of the transition and output functions based on the input string. I test each automata against all possible combinations of opponent action input strings of length nine. There are $3^9 = 19{,}683$ possible combinations of these input strings of length nine, generating a total output string of length 196,830.

The similarity of two automata is a measure between zero and one, based on how many of the elements out of 196,830 on which they recommend the same action. For example, if two automata have output functions that recommend different actions in the start state, then it is necessarily true that they will differ on at atleast 19,683 elements, and cannot have a similarity score greater than 0.9. Based on the similarity measure, we can construct a 173,913 X 173,913 similarity matrix where each element $S_{i,j}$ takes on a value between zero and one denoting how similar automata i is to automata j in terms of reactions to various possible input combinations.

Affinity propagation clustering technique

The affinity propagation clustering technique[7] described below uses the measure of 'similarity' described in the previous section. Let the set of all 173,913 automata be denoted as K. The affinity propagation clustering technique includes the the following steps:

1. The diagonal elements of the similarity matrix $S_{k,k}$ will be equal to 1 for all values of $k \in K$ because each automata is perfectly similar to itself. When the values on the diagonal are equal to 1, each automata becomes its own exemplar, and the final number of clusters is equal to the total number of automata (173,913), and there is no clustering. To avoid this, all diagonal values in the similarity matrix are replaced by a value chosen at the discretion of the modeler. If the chosen value is close to the minimum value of the elements of the similarity matrix (close to zero in this case), then each automata has a very small likelihood of ending up as an exemplar, and there will be a fairly small number of final clusters. If

[7] Introduced by Frey and Dueck (2007) to identify a clustering technique where the choice of the final number of clusters is endogenous.

the chosen value is close to the maximum value of the elements of the similarity matrix (close to one in this case), then each automata has a strong likelihood of ending up as an exemplar, and there will a large number of final clusters. The default mechanism is to select a value equal to the median value of the elements of the similarity matrix, which has been followed here. Each automata initially has an equal chance at becoming an exemplar.

2. Once the diagonal values of the similarity matrix are updated with the median values of the elements of the similarity matrix, 'responsibility' messages are sent from each automata to candidate exemplar automata that are competing to become exemplars. The message goes from each automata i to the candidate exemplar automata $k \in K$ and helps exemplar automata compete for ownership of automata i. The messages sent from automata i to the candidate exemplar automata and itself respectively are provided below: $R(i,k) = S(i,k) - max_{k' \neq k}\{A(i,k') + S(i,k')\}$ and $R(i,i) = S(i,i) - max_{i' \neq i}\{S(i,i')\}$. Initially, the value of $A(i,k')$ is zero for all automata, so we will ignore that for now. At the start of the algorithm, most automata will send negative responsibility messages to automata that are far away from it, and will send a responsibility message of zero to only the automata most 'similar' to itself. Since all automata initially start with the same value of $S_{i,i}$, the responsibility message each automata sends itself will be positive only for automata that are relatively far away from all other automata.

3. The next set of messages, known as 'attractions' messages. is sent from each candidate exemplar automata k to all the other automata $i \in K$ announcing their legitimacy, and their candidacy. The messages sent from aspiring exemplar automata k to every other automata and itself respectively are provided below: $A(i,k) = min\{0, R(k,k) + \sum_{i' \notin \{i,k\}} max\{0, R(i'k)\}$ and $A(k,k) = \sum_{i' \neq k} max\{0, R(i',k)\}$. Thus, for each candidate exemplar k, the message it sends to all the other automata i is the initial level of responsibility it has towards itself, along with the sum of any positive responsibilities it has received from other automata $i' \in K$. Since most responsibility messages are initally close to zero, these attraction messages take on values that are usually nega-

tive. Only candidate automata that have been receiving many responsilibility messages from other automata will take on enough attractions to get to an attractions value close to zero. The attraction that candidate exemplar automata k has for itself is simply the sum of all the positive responsibilities it has received from other automata.

4. Thus, candidate exemplar automata that do not initially receive a lot of responsibilities start to have negative attraction values, and start dropping out as candidates in the 'maximum' race in the responsibility message formula, and eventually stop remaining viable candidates as exemplars. On the other hand, candidate exemplars who initially receive many positive responsibilities have attractions that have values that are close to zero, so they are more likely to keep winning the 'maximum' race in the responsibility messages formula for automata that are 'similar' to them, and the responsibility scores between them and automata that are 'similar' to them remain somewhat constant. These automata start accruing many positive responsibilities, and have higher attraction scores for themselves.

5. After every round of responsibility and attraction messages, for every automata i, we can find the exemplar automata k that maximizes the value of $\{R(i,k) + A(i,k)\}$. For some automata, i will be equal to k, and these automata will become their own exemplars, and thus the centroids of their own cluster.

6. After every iteration, the set of exemplars is identified and the automata that belong to each cluster are identified. If the set of exemplars and the automata in each cluster remain the same from one iteration to the next, the algorithm is said to have converged, and the final set of exemplars identify the clusters into which the 173,913 automata have been classified. Here, we undertake this entire procedure twice. In the first instance, 173,913 automata are classified into 8,039 clusters. Then we perform the entire procedure again using only the exemplar centroids from these 8,039 clusters, and finally identify 440 clusters.

Thus, I restrict the analysis to the 440 repeated-game strategies (modeled as automata) that are the representative automata of the clusters identified after the affinity propagation clustering technique. I follow the same algorithm to identify the initial attractions and the same methodology for attractions in later rounds as the 'attractions towards actions' model, but use the 440 repeated-game strategies (modeled as automata) instead of the three actions.

2.4.3 Simulated subjects (Grid Search)

The learning algorithm described for the attractions towards strategies estimation method is used to simulate play between pairs of players. I have run a grid search over different parameter values, with

$\alpha \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

$\beta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

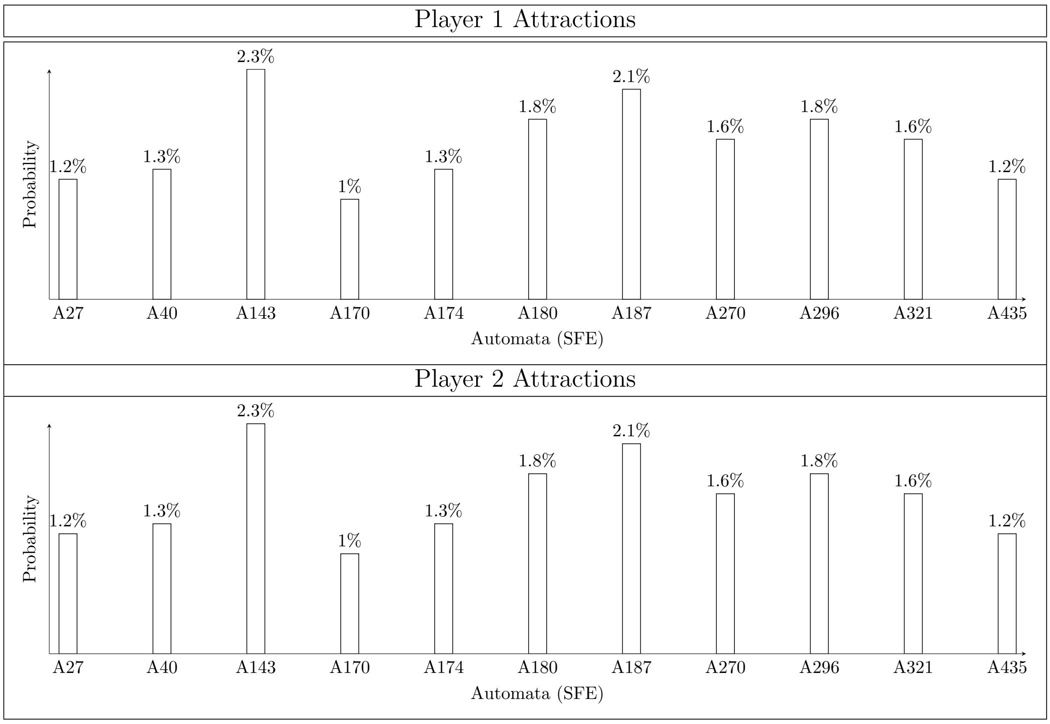
$\delta \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

$\lambda \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$,

and $\gamma \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$.

For most parameter values, the converged fixed points for the initial attractions towards strategies follow a pattern where most of the probability weight is on a few strategies. For example, Figure 2.6 shows the histogram of the probability distribution for the converged initial attractions towards strategies for both players, when considering parameter values ($\alpha = 0.3, \beta = 0.7, \delta = 0.0, \lambda = 1.0, \gamma = 0.8$). To simplify the discussion, I only show strategies with limiting probability weights greater than or equal to 1%. These strategies are, in a sense, the principal strategies that players 'bring to the laboratory' for the 'experimental' phase. The automata that have a weight greater than 1% are presented in Appendix B1.

Figure 2.6: Converged initial attractions for the two players in the SFE game (Strategies)



2.5 Experimental subjects

The payoff matrix used in the experiments in the same as shown in Table 2.1. Subjects can see the entire payoff matrix as they play along, so they are fully aware of both their own payoffs, as well as those of their opponents. The experimental sessions were run at the Economic Science Laboratory at the University of Arizona. Subjects interacted on computers using a Google Chrome interface that was programmed using Python. Subjects' final payoffs consisted of the sum of their earnings from all periods of the experiment. With the completion of the experiment, subjects were paid their cash earnings in private. The average payoff was $13.2 for sessions that lasted approximately an hour.

The game-play consisted of a fixed matching protocol. There were a total of 5 sessions consisting of 100 subjects. In each session, subjects played the SFE supergame a total of ten times, but here we are only analysing the choices made in the first supergame. Subjects were never matched with another subject twice. At the beginning of each supergame, each subject was randomly paired with one other subject and they played the 3 X 3 game repeatedly with the same partner. After all pairs had finished each supergame, subjects were rematched with a new partner that they had not been matched with before, and played the game again. The length of the supergame was determined using the following termination rules :

- T0 - Starting in the first round, the continuation probability was $\eta = 0.98$, so the expected length of each supergame was 50.[8]

In each round, subjects were asked to select a choice (A, B or C). The payoffs were displayed in a table, and players roles were always the same (i.e. they didn't have to switch between row and column player). The final experimental data consists of 50 observation pairs in the game.

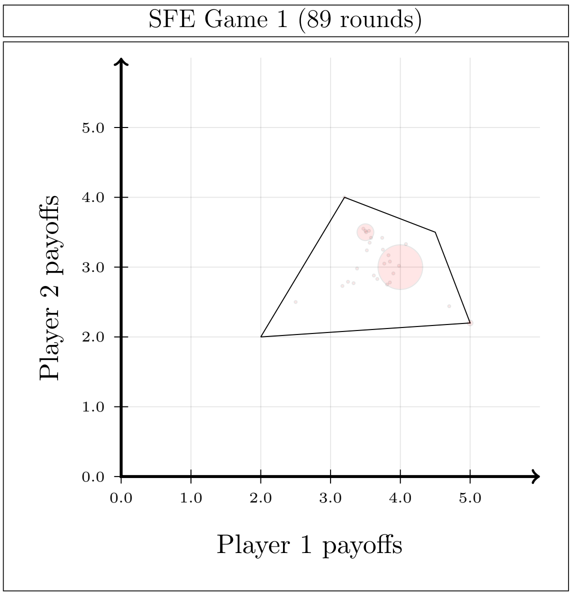
Experimental Data Patterns

There are some interesting patterns observable in the experimental data. Figure 2.7 shows the heat map with the relative frequency of each payoff combination over the last 20 periods of game-play for the experimental subjects. The relative frequency of a payoff combination is denoted by a circle located on the coordinates that correspond to that combination; the larger the circle, the higher the (relative) frequency of that combination. I also display the set of feasible payoffs. The patterns in the final set of outcomes are not entirely captured by the Nash equilibria outcomes, but also do not support the entire set of Folk Theorem payoffs.

About 32% of subjects end up playing the Nash equilibrium outcome with both players selecting their 'selfish' action. Another 12% of subjects end up at the payoff outcome of (35, 35) which is reached either by a combination of Player 1 playing the 'fair' action while Player 2 plays the 'selfish' action, or by a combination of Player 1 playing the 'fair' action while Player 2 plays the 'efficient' action. The other subjects end up at a broad range of outcomes, but no more than one pair of subjects end up at any other payoff outcome.

[8] The actual draw for the length of the supergames was made prior to the start of the sessions using the geometric distribution with $X \sim Geom(1-\eta)$, and the first supergame lasted 89 rounds.

Figure 2.7: Experimental Patterns



2.6 Results

The model attempts to capture behavior seen in the laboratory for repeated games where the three actions are motivated by efficiency, fairness, and selfishness concerns. The payoffs are structured such that the game has a single Nash equilibrium with both players playing their 'selfish' action.

2.6.1 Comparison between simulations and experimental subjects

The experimental data has 50 observation pairs for the SFE game. For each observation pair, I note the number of rounds played (89 rounds), and simulate play for the same number of rounds between a 'simulated pair,' using a parameter group $(\alpha, \beta, \lambda, \gamma)$ to generate propensities towards actions and strategies. I then calculate average payoffs for the experimental subject pair in the last 20 rounds, and also calculate average payoffs for the simulated subject pair in the last 20 rounds. Finally, I calculate the Euclidean distance between the two average payoffs. These pairs have to be matched to each other in a way that globally minimizes the sum of the Euclidean distance across all 50 pairs. I use the Kuhn-Munkres algorithm (also known as the Hungarian assignment problem algorithm) described in Kuhn (1955) to perform this objective.

The parameter values that minimize the errors between average payoff points between experimental subject pairs and simulated subject pairs in the SFE game using the three estimation methods are presented below:

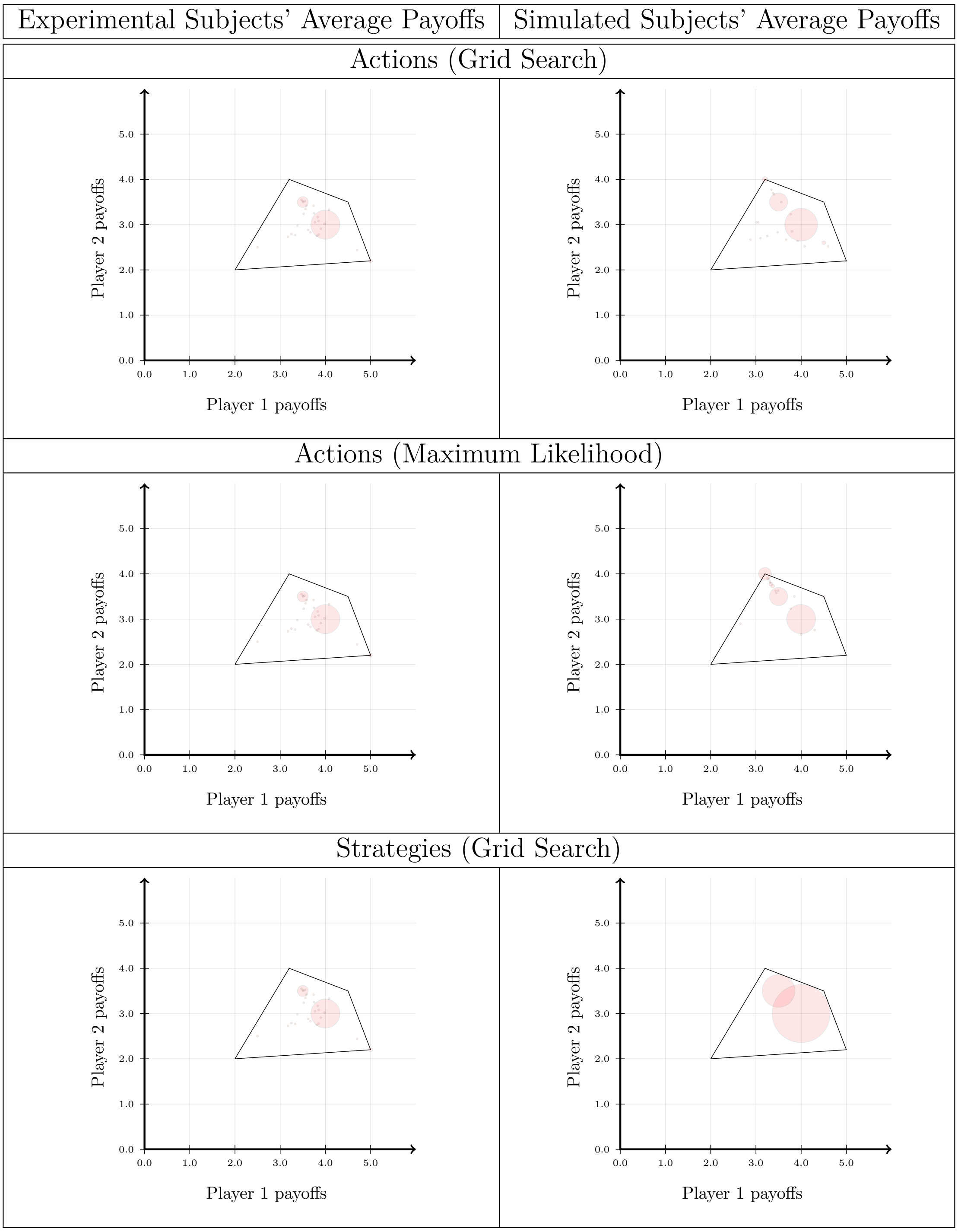
Attractions towards actions (grid search): $(\alpha = 0.1, \beta = 0.9, \delta = 0.0, \lambda = 0.1, \gamma = 0.9)$.

Attractions towards actions (maximum likelihood estimation): $(\alpha = 0.11, \beta = 0.89, \delta = 0.00, \lambda = 0.10, \gamma = 1.00)$.

Attractions towards strategies (grid search): $(\alpha = 0.3, \beta = 0.7, \delta = 0.0, \lambda = 1.0, \gamma = 0.8)$.

Figure 2.8 below shows the heat map for the computational simulations and the experimental results with human subjects using these parameter values, and shows the relative frequency of each payoff combination over the last 20 periods of gameplay. The relative frequency of a payoff combination is denoted by a circle located on the coordinates that correspond to that combination; the larger the circle, the higher the (relative) frequency of that combination. I also display the set of feasible payoffs.

Figure 2.8: Full Information, Symmetric Payoffs



The strategies model slightly over-predicts the level of Nash equilibria and the (35, 35) outcome, while the actions model using both methods does a very accurate job. We can present some examples of repeated-game strategies using only two-state two-action automata that can sustain the (35, 35) outcome. For example, if Player 1 responds to Player 2 playing the fair action by playing the selfish action, but plays the fair action if Player 2 plays either the selfish action. This is just one example of a repeated-game strategy that can help explain the (35, 35) outcome, and there are many other three-state three-action automata that can sustain this outcome.

The grid search for the attractions towards actions model shows that both efficiency and fairness measures are important when attempting to minimize errors relative to experimental data. Table 2.3 below presents the error rate for different parameter combinations when placing either no weight, or all the weight, on the considerations of efficiency, fairness, and selfishness. The error rates are significantly higher when all the weight is on only efficiency or only selfishness. The closest fit to the experimental dataset places high weights on efficiency and fairness.

Table 2.3: SFE Game (Actions) Error Rate

	Efficiency (α)	Fairness (β)	Selfishness ($1-\alpha-\beta$)	QRE Scale (λ)	Gamma (γ)	Global Euclidean Distance
Only efficiency	1	0	0	0.1	0.6	15
Only fairness	0	1	0	0.6	0.8	20
Only selfishness	0	0	1	0	0	26
No efficiency	0	0.8	0.2	0.1	0.1	10
No fairness	1	0	0	0.1	0.6	15
No selfishness	**0.1**	**0.9**	**0**	**0.1**	**0.9**	**7**

The grid search for the attractions towards strategies model shows that both efficiency and fairness measures are important when attempting to minimize errors relative to experimental data. Table 2.4 below present the error rate for different parameter combinations when placing either no weight, or all the weight, on the considerations of efficiency, fairness, and selfishness. The error rates are significantly higher when all the weight is on only efficiency, fairness or selfishness. The closest fit to the experimental dataset places high weights on efficiency and fairness.

Table 2.4: SFE Game (Strategies) Error Rate

	Efficiency (α)	Fairness (β)	Selfishness ($1-\alpha-\beta$)	QRE Scale (λ)	Gamma (γ)	Global Euclidean Distance
Only efficiency	1	0	0	0	0	25
Only fairness	0	1	0	0	0.8	22
Only selfishness	0	0	1	0	1	23
No efficiency	0	0.5	0.5	0.7	0	13
No fairness	0.5	0	0.5	0.5	0	15
No selfishness	**0.3**	**0.7**	**0**	**1**	**0.8**	**12**

2.7 Conclusion

There are a number of experimental papers (Mathevet and Romero (2012), Rapoport and Chammah (1965), Rapoport et al (1976) among others) which show that certain payoff outcomes become focal in the minds of subjects when playing indefinitely repeated games in the laboratory. I see a similar pattern in the SFE game, with 44% of all subject pairs coordinating on only two payoff outcomes possible in the game. To better account for the experimental patterns seen among laboratory subjects when playing this game, I use three different learning models to test if incorporating social preferences into the utility function for subjects over long-run payoff outcomes helps account for experimental behavior that cannot be captured by just the Nash equilibrium outcomes. In addition, I allow subjects to respond to their beliefs about opponent actions with noise, using a quantal response equilibrium framework, rather than assuming that players are infinitely rational. Finally, the model allows subjects to learn through the reinforcement of state-game payoffs as they progress through the game. Combining these three elements allows us to use three estimation techniques - 'attractions towards actions' models using the grid search technique and a maximum likelihood estimation technique, as well as an 'attractions towards strategies' model using the grid search technique.

The learning model studied here has a restriction on the strategy space, but uses weighted versions of efficiency, equality, fairness, and selfishness scores to evaluate attractions that players have towards actions and strategies. The tendencies of experimental subjects can only be explained by placing high weights on efficiency and fairness, but no weight on selfishness. The consistency of this pattern across all three models suggests that subjects use actions and strategies that lead to efficient and fair outcomes, relative to selfish outcomes in this game.

Chapter 3

The legitimacy of status, and its effects on other-regarding preferences

This paper investigates how legitimacy of status affects other-regarding preferences towards peers in the absence of ex-ante inequality. We use modified Dictator games to study the impact that players' beliefs regarding the legitimacy of their own and their partner's endowments have on other-regarding preferences. While there is a consensus that some conditionally altruistic motives like reciprocity are driven by players' beliefs about their partner's actions and utilities, we investigate whether alternate channels of conditional altruism are driven by players' beliefs about the legitimacy of their status. We find that the connection between prosocial behavior and 'deservingness' of receivers seen in the presence of ex-ante inequality is not replicated when that inequality is absent. However, there is a negative relationship between the legitimacy of one's own claim to an endowment (or one's own 'deservingness') and the level of prosocial behavior towards others.

JEL: C70; C91; D63; D64

Keywords: Experimental Economics; Dictator Games; Altruism

3.1 Introduction

Identifying underlying primitives that determine the magnitude of prosocial behavior in economic agents has long been a focus for economists. A large portion of the literature encompassing social preferences addresses situations in which agents have unequal status (usually through unequal endowments), and investigates generosity from the rich to the poor. Two broad classes of motives have been used to explain redistribution behavior in such cases. The first is unconditional altruism, where payoff inequalities motivate the rich to help the poor, irrespective of receiver intentions or other characteristics. Models in this class include simple altruism (Becker (1974)), warm-glow altruism (Andreoni (1989)), and preferences for egalitarianism (Fehr and Schmidt (1999)). The second class refers to conditional altruism, through which giving is based on the giver's perception of what the receiver 'deserves'. Past actions of the receiver (Oxoby and Spraggon (2008)) or demonstrated need (Fong (2001)) affect the extent of conditional altruism displayed by privileged subjects (in payoff terms) to disadvantaged receivers.

Consider the following example: Imagine a situation where we are approached by a stranger who claims to a veteran and asks for a handout. In another situation, we are approached by a stranger who makes a similar claim but is dressed in Army fatigues. Are we equally likely to give the handout in both situations? If yes, are we likely to give similar amounts? Conditional altruism models predict higher levels of giving to a disadvantaged veteran in fatigues, as compared to a disadvantaged stranger with claims that cannot be visually verified. Conditional altruism suggests that antecedents can directly affect giving behavior from advantaged populations to disadvantaged populations. Moreover, the example illustrates that one way antecedents affect this giving behavior is by changing our perceptions regarding the 'deservingness' of the disadvantaged receiver.

In contrast to the example above, the focus of this paper is on investigating the determinants of conditional altruism towards peers, rather than disadvantaged populations. In the absence of inequality, unconditional altruism is unlikely in situations where agents have non-zero levels of endowment. However, conditional altruism could still play a role in determining prosocial behavior towards peers. Consider the following motivating example: Peers working at the same designation level within a firm are usually paid the same wage levels. This is especially true in white-collar jobs in established offices, oftentimes mandated by laws such as the Fair Labor Standards Act of 1938, or the Equal Pay Act of 1963. However, levels of productivity are likely heterogeneous across these peers at the same designation level. More productive workers may feel that a larger share of their overall wage income is being determined by skill and effort, while their peers who are not as productive are enjoying the same levels of fruit, but largely as a result of luck rather than skill because the pay at a certain designation level is uniform by mandate. This perception about the 'deservingness' of their peers can lead to ill-will within groups, and it is useful for firms to understand (and possibly measure) the level of dissatisfaction that this perception causes among the more productive employees. It should be noted that many firms mitigate this concern either by using bonuses that are proportional to productivity levels, or by firing employees within the designation level whose productivity levels are below certain thresholds. However, both these methods require management to have clarity and observability of effort and skill, and the direct connection between effort provision and productivity levels. This is often not the case in firms with many employees, and also cannot address situations where the gaps between employees' productivity levels are not significant enough to be obvious to management.

This paper studies how the altruism of economic agents is driven by their beliefs about the legitimacy of their own endowment status, as well as those of their peers. Specifically, we study how these beliefs affect their other-regarding preferences towards peers in the absence of ex-ante payoff inequality. We propose a theoretical model of incomplete information, where Dictators are unsure about certain parameters of their own utility functions. In this framework, players' utilities are influenced not only by final payoff outcomes, but also by players' beliefs at various history nodes as they progress through the game. Specifically, we use a utility functional form borrowed from Cox, Friedman, and Gjerstad (2007), that imagines a marginal rate of substitution between one's own and other players' welfare in the utility function. Our model assumes that agents are uncertain about this marginal rate of substitution in their utility functions, and their beliefs about this parameter generates differing levels of other-regarding preferences. The model predicts that signals which can influence these beliefs, such as signals showing the ratio of initial endowments that are a result of effort rather than luck, will have an influence on altruistic behavior in later scenarios where agents can express generosity towards their peers. This is also true when considering a version of the model that incorporates elements of reference dependence and gain-loss aversion like Koszegi and Rabin (2006).

We test the predictions of the model by running experiments involving modified Dictator games. The experiment consists of two stages. In the first stage, all subjects are involved in a combination of real effort tasks. Their effort levels in the first tournament stage determine their role as either the Dictator or the receiver in the second stage. Through an effective randomization, subjects are paired in a way that their first stage earnings are a combination of luck and effort, and subjects within a pair have the same total first stage earnings. Previous experiments are commonly structured so that income is either completely earned in a contest, or attained solely through lucky randomization, but is rarely a combination of the two.[1] An interesting facet of the Alesina and Angeletos (2005) model exploring cross-cultural generosity trends is that luck and effort both play a role in the determination of income. Inspired by this, the experiment is set up so that earnings at the end of the first stage are a result of the combination of luck and effort, with differing ratios for subjects within a pair that have equal amounts of total first stage earnings. In the second stage, Dictators can distribute a new set of payoffs between them and their matched partner.

The main results of the paper show that giving levels in the Dictator stage of the experiment are not correlated with the ratio of receiver's first stage earnings that are a result of effort, rather than luck. The relationship is statistically insignificant, suggesting that the effect of antecedents on giving levels witnessed in setups where advantaged groups give to disadvantaged groups is not replicated among subjects that

[1] An exception is Stoddard, Lefgren, and Sims (2016) where income is a combination of luck and effort/skill. However, their experiments do not allow effort levels to be a choice variable, while we do not place any such constraints in our sessions. Some other studies allowing for income to be a combination of luck and effort include Cappelen, Hole, Sorensen, and Tungodden (2007); Erkal, Gangadharan, and Nikiforakis (2011), Rubin and Sheremeta (2016), and Ray-Biel, Sheremeta and Uler (2018).

are peers in payoff terms, in the absence of ex-ante inequality. However, there is a statistically significant negative relationship between giving levels and the ratio of the Dictator's first stage earnings that are a result of effort rather than luck. Dictators who can legitimize their own claims to an endowment are less likely to show benevolence towards matched partners, even when distributing a new set of payoffs. The experiment is carefully calibrated to control for receiver's beliefs about her first stage earning levels, and also controls for the Dictator's beliefs about the receiver's expectations. Thus, our results are not driven by psychological game theory channels like reciprocity or guit aversion. This is supported by the fact that we find no relationship between receivers' performance levels in the first stage, and their expectations of receipt amounts in the second stage. There is also no statistically significant relationship between Dictators' beliefs about receiver's expectations of giving levels, and the receiver's expectations of receipt amounts in the second stage. Thus, we claim that prosocial behavior towards peers that is independent of beliefs about peers' expectations of prosocial behavior, is also independent of subjects' perceptions of the 'deservingness' of their peers, but does depend on their own level of 'deservingness'. Finally, we also find that wealthier Dictators (subjects with higher total endowment levels in the first stage) give lower amounts in the second stage.

Section 2 summarizes the literature that is closely related to the concepts studied here, Section 3 covers the details of the experimental setup, and Section 4 covers our main hypothesis and the theoretical framework underlying the predictions. Sections 5 presents all the experimental results, and Section 6 concludes.

3.2 Related Literature

The paper makes a contribution to a few different strands of literature. Foremost, it contributes to the vast and ever-expanding literature regarding social preferences by agents in non-strategic interactions. An extensive number of experimental studies have found that subjects in the laboratory make choices inconsistent with the maximization of purely material self-interest. Such actions have been rationalized in theoretical models by assuming other-regarding preferences (in addition to self-regarding preferences), where subjects are thought to care not only about their own payoffs, but also the payoffs of their partners. Various modifications of Dictator games have been used to study the existence and properties of these other-regarding preferences since the game was first popularized by Kahneman, Knetsch, and Thaler (1986). The canonical Dictator game has a simple structure where one party (the Dictator) is asked to split a sum of money between themselves and another party (the receiver). While self-interested subjects should keep the entire pie for themselves, experimental subjects regularly give about 20 to 30 percent of the pie to the other party. Maximization of one's own material payoffs is a poor predictor of how experimental subjects play in a number of other contexts, such as Ultimatum games, Investment games, Public Goods games with voluntary contributions, and experimental labor markets. Fehr and Gachter (2000) summarize the experimental evidence on the economic motives beyond purely self-interested payoffs.

The laboratory data, and some suggestive field data, have led to the development of theoretical models that explicitly incorporate other-regarding preferences as a source of utility for agents. These models have appealed to concepts such as altruism, 'warm-glow' effects, fairness, reciprocity, justice, inequality aversion, and adherence to social norms, among others to capture the nature of other-regarding preferences. The literature covering these models falls broadly into two main categories. First, there are the outcome-based models that focus on the relative payoff, or distributional outcomes, of the material payoffs in the game. Theories of fairness and inequity aversion have theorized that players feel a disutility from receiving amounts that are disproportionately higher (or lower) than the amounts received by their matched partners (e.g., Fehr and Schmidt (1999), Charness and Rabin (2002), and Bolton and Ockenfels (2000)). These models posit that players only have preferences over final payoff outcomes.

The main alternatives to the outcome-based distributional preference models are intentions-based reciprocity models, where subjects are motivated by their beliefs about the expectations of their matched partners and are also influenced by their perceptions about the intent behind their partners' actions, irrespective of the final outcomes. Intention-based reciprocity models predict that reciprocal behavior can help explain giving behavior in Ultimatum games, where the receivers are allowed to reject the offered split and generate a zero-payoff for both parties (e.g., Rabin (1993), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006)).

Both outcomes-based distributional preference models and intentions-based reciprocity models are driven by actions taken during the course of the game or beliefs updated through the course of the game. More recently, research has shown that the importance of other-regarding preferences is often influenced by antecedents. This is another strand of literature to which this paper helps contribute. Antecedents such as the exogenous variation in the action set (i.e., expanding or decreasing the possible amounts that the Dictator can give or take) as in List (2007), as well as exogenous variations in the endowment set, lead to differences in levels of giving. These variations have sometimes been attributed to framing effects. Dufwenberg, Gachter, and Hennig-Schmidt (2011) show that labeling a game with the same payoffs as a “Community” or “neutrally framed” game changes the subjects’ behavior. Similarly, they find that public goods games where subjects are endowed with money and can contribute to a public good, when described as “give-some” or “take-some” games, generate differing levels of contribution. Hoffman and Spitzer (1985), and Hoffman, McCabe, and Smith (1994) compare situations where individuals earn their roles in an experiment (e.g., earning the role of seller or first mover) to standard Dictator games. They find that subjects that earn their advantaged positions are more likely to retain greater amounts for themselves. Participants appear to legitimize their entitlement to better outcomes once they feel they have earned the right to influence these outcomes, which in other circumstances might be considered unfair.

Cherry, Frykblom, and Shogren (2002); Cherry (2001); and Ruffle (1998) find that levels of giving in Dictator and Ultimatum games are also influenced by the source of wealth: individuals that earn assets and can claim to have legitimate claims to those assets keep larger shares for themselves when compared to baseline cases where assets are unearned. Oxoby and Spraggon (2008) consider treatments where the endowment to be distributed in Dictator games is earned by the Dictator or earned by the matched partner, and find that giving behavior is higher if the source of the endowment is the earnings of the matched partner. In summary, individuals’ perceptions of fairness and extent of concern about the payoffs of others are influenced by the factors leading up to a decision environment and the characteristics of that decision environment. This is also true when the receiver of the redistribution is a charity, rather than another subject in a laboratory environment. Reinstein and Riener (2012) show that the levels of giving in the laboratory towards other subjects is replicated when considering charities as receivers. Similar outcomes are seen in Fong (2001), as well as Fong and Oberholzer-Gee (2011) where different levels of information regarding the receiver's need or 'deservingness' has strong effects on the level of charitable giving.

The specific antecedent we study in this paper is the nature of earnings earned through effort, as opposed to luck. Benabou and Tirole (2006) and Alesina and Angeletos (2005) posit models that suggest that differences between support for redistribution rates between the United States and Western Europe is driven by divergent beliefs about the role of luck vs effort in the accrual of earnings. As a result, the paper also contains aspects of the literature exploring how differences in the sources of income affects support for redistribution using lab experiments. The key takeaway from this literature is that the results are mixed, and very context dependent. Papers such as Balafoutas, Kocher, Putterman, and Sutter (2013), Durante, Putterman, and Weele (2014), and Stoddard, Lefgren, and Sims (2016) find that there is greater support for redistribution when income results from luck rather than performance or effort. Stoddard, Lefgren, and Sims (2016) compare a treatment where inequality of payoff is directly proportional to effort with a treatment where inequality of payoff is not tied to effort level, and find that support for redistribution is higher when inequality is unrelated to effort. When redistribution decisions are made before individuals know their actual ex-post income, they find greater support for redistribution when income is randomly assigned than when income is assigned by zip code, Tetris skill or score on a quiz. Ku and Salmon (2013) find the opposite result where there is less support for redistribution when income in randomly assigned than when income is based on a skill task, specifically the SAT test.

Finally, this paper also contributes to the literature on personnel economics and specifically, the impact of pay compression within hierarchical levels within an organization. An element of the choice of compensation is whether, within a firm, the pay for individuals should have less variance than the variance in their performance levels. The often-stated objective of a compressed pay structure is to make pay more equitable, and create a fair environment which promotes harmony among workers. While people do seem to value equity and fairness as shown in Baron and Kreps (1999), what happens in situations where output results from choices about effort rather than inherent differences in ability? If co-workers are exhibiting lower levels of effort, but enjoying similar levels of pay, this can cause dissension within the ranks. In general, pay compression reduces the incentives to engage in sabotage or other noncooperative behavior that is induced by the prisoner's dilemma style payoff structure (Lazear, 1989). However, this result could be context-specific depending on how closely teams have to work together to provide joint output, which requires a higher level of bonhomie within members of similar ranks.

Our study is most closely related to two recent papers - Gee, Migueis, and Parsa (2017), and Sheremeta and Uler (2018). Gee, Migueis, and Parsa (2017) study the interaction effect between levels of inequality and the support for redistribution when incomes are generated by effort, rather than luck. They find that support for redistribution is higher when incomes are a result of effort as compared to luck when initial levels of inequality are low, but support for redistribution is lower when comparing effort and luck incomes when initial levels of inequality are high.[2] Our experiment complements their findings by studying settings where there are no initial levels of inequality, and improve the design by controlling for higher-order beliefs of subjects that may affect giving levels. Sheremeta and Uler (2018) study how giving depends on income and luck, and cross-cultural differences in how this relationship changes when subjects are given information about how much of their incomes are a result of effort and luck. They find that US and Spanish subjects give similar amounts when they are fully informed about the source of income (effort vs luck), but US subjects give significantly smaller amounts in the absence of this information. Our study complements theirs by covering situations where only Dictators are fully informed about the source of income for them and receivers, so we additionally control for any expectations that receivers may form based upon this information. In addition, we find results similar to theirs in situations with no ex-ante inequality.

[2]In contrast to their results, Krawczyk (2010) finds that going from a low inequality setting to a high inequality setting leads to an increase in support for redistribution when the income is a result of effort as compared to luck.

We explore the effect of players' beliefs about the legitimacy of their relative status by changing the legitimacy of their own and their matched partner's endowments. By leveraging a modified version of the reference-dependent utility framework explored in Koszegi and Rabin (2006), we explore how antecedents shift a 'fairness' reference point that participants place on the marginal rate of substitution between one's utility and the matched partner's utility. Our hypothesis in this paper is that antecedents change the importance participants give to the utility of others by shifting this 'fairness' reference point. Given this effect, antecedents affect the reference point of allocations against which Dictators evaluate all potential payoffs.

3.3 Experimental Setup

In this section, we discuss the experimental setup used to study the predictions of our model discussed in Section 4. All participants received a basic show-up fee of \$5. In addition, their performance in the experiment was rewarded with Experimental Currency Units (ECUs), and the conversion rate for ECUs was 20 ECUs = \$1. The experiment was conducted in two stages, with a combination of effort tasks in the first stage used in a tournament style. Subjects were randomly matched at the end of the first stage, and then a modified Dictator game was played in the second stage.

The first stage involved all participants participating in effort tasks. The nature of the effort tasks was chosen so as to successfully separate the effect of effort from ability. We used two different kinds of effort tasks that can help identify effort (as separate from ability) with varying degrees of success. The first effort task used is the *Slider Task* proposed by Gill and Prowse (2011) where subjects were presented with forty-eight sliders on their computer screen, with a range of integer values from 0 to 100, initially at 0. They were asked to use their mouse to position the slider at 50. Once the slider was adjusted, the value of its position was shown to the right of the slider. If the number did not read 50, subjects were able to adjust the position again. There was no limit on the number of attempts per each slider. Subjects had two minutes to correctly position as many sliders as they could. Subjects' performance was measured by the number of sliders correctly positioned in two minutes, and they received 1 ECU for each slider correctly positioned. The second effort task was the

Grid Task, where a 15 by 15 grid with randomly distributed 0s and 1s appeared on the screen, and subjects were asked to count the number of zeros in the table. They had two minutes to solve as many grids as they could, and if they had three unsuccessful attempts, a new table was generated. In this task, which is similar to the one used by Abeler, Falk, Gotte, and Huffman (2011), subjects' performance was measured by the number of grids correctly solved, and subjects received 3 ECUs for each successful grid. The subjects alternated between the two effort tasks five times each. Thus, the total length of completing all effort tasks was 20 minutes. This stage was constructed with a special emphasis on requiring a significant effort level from participants, in order to create strong feelings of property rights to endowment levels, and stimulate notions of 'deservingness'.

After the first stage, subject were randomly matched into pairs. The subject within a pair with the higher score in stage one was assigned to Group A (who played the role of Dictators in the second stage), while the subject within a pair with the lower score was assigned to Group B (the receivers in the second stage). Subjects then received a stage one bonus, determined partly through randomness, and partly based on the group they were assigned to.[3] A random draw was made over a uniform distribution from zero to twenty ECUs to determine the stage one bonus for Group A participants. The difference between the total stage one earnings for the Group A subject within a pair (their stage one earnings through the effort task plus the stage one bonus) and the stage one earnings through the effort task for the Group B subject within the pair was assigned as the stage one bonus to the Group B participant in the pair. This ensured that both subjects within a pair had the same level of earnings at the end of the first stage, but had different ratios of how much of their stage one earnings was earned through effort, rather than luck.

In the second stage, participants in Group A were provided a table showing the breakdown of the stage one earnings through the effort task and bonus stage one earnings earned by them, and their matched partner. They were also provided a

[3]Subjects were informed of this setup for the stage one bonus in their written instructions before starting the session, to avoid any notions of deception. The exact instructions were "The Stage One Bonus has a completely random component, and a deterministic component that depends directly on whether you are in Group A or Group B."

new endowment of 120 ECUs that they were asked to split between them and their matched partner. They were reminded that their partner did not know the breakdown of their stage one earnings, and also would not learn their stage two earnings. The only information that Group B participants received was the total ECUs earned as a combination of stage one and stage two earnings. This design helped control any expectations that receivers may have had given their effort levels in the first stage, and ensured that Dictators did not need to take these expectations into consideration when distributing ECUs in the second stage. This setup also helped Dictators obfuscate the level of generosity that they were showing to their matched partners, which provided an avenue for displaying more selfish levels of giving than situations where receivers would know the exact level of generosity shown by their matched partners.

After the second stage ended, Group B participants were incentivized to provide an accurate belief of how many ECUs they expected to receive in the second stage, and Group A participants were incentivized to provide an accurate belief of how many ECUs their matched partner expected to receive. This helps us accurately determine what beliefs and expectations subjects had, and what effect the accuracy of their beliefs had on the levels of giving in the second stage.

The action set available to the Dictators in Group A was the same for all subject pairs. They had to allocate 120 ECUs between themselves and their matched partner. Once the effort task was over, subjects in Group B could no longer take any further actions to influence the beliefs about reciprocity that subjects in Group A may have, so these beliefs were the same across all subject pairs. Since subjects in Group A were told explicitly that their partners in Group B had no knowledge about their performance in the effort task, we explicitly controlled for their beliefs about the expectations that their matched partner may have had. Beliefs about guilt-aversion were the same across all subject pairs. However, across subject pairs, there were differing levels of stage one receiver earnings that resulted from effort rather than luck.

Both outcome-based distributional preference models and intentions-based reciprocity models predict the same level of giving across all subject pairs. The theory of fairness and reciprocity proposed by Cox, Friedman, and Gjerstad (2007) also predicts the same level of giving across all subect pairs because the relative status of the two groups remains the same for all subject pairs. However, the beliefs that participants in Group A have about the legitimacy of the endowments of the matched partner (how much of their matched partner's endowment was *earned*, rather than given) varied across all subject pairs. Similarly, beliefs that participants in Group A have about their own legitimacy also varied across subject pairs. We hypothesize that the updated beliefs about the legitimacy of the endowments shifts the 'fairness' reference point of players, and their giving behavior reflects these new 'fairness' reference points. If we imagine three subject pairs where all the Dictators earned 40 ECUs in the effort task and received a draw of 10 ECUs as the first stage bonus, but they face receivers who earned 20, 30, and 40 ECUs in the effort task, we hypothesize that giving behavior will increase as we move across subject pairs. The conditions summarizing such hypothetical subject pairs are presented in Table 3.1 below.[4]

Table 3.1: Subject pairs with differing levels of "deservingness"

Subject Pair #	Dictator's Earned ECUs	Dictator's Bonus ECUs	Receiver's Earned ECUs	Receiver's Bonus ECUs	Receiver's Earned ECU %
1	40	10	20	30	40%
2	40	10	30	20	60%
3	40	10	40	10	80%

In Section 4, we propose a utility form similar in spirit to the one proposed in Cox, Friedman, and Gjerstad (2007), and augment it with the reference-dependence framework explored in Koszegi and Rabin (2006). The updating process used to explore shifting 'fairness' reference points helps explain the behavior we expect to see across subject pairs.

[4]Alternately, we could also imagine three subject pairs where the level of earned and bonus ECUs of the matched partners are equal, but the level of earned ECUs for the Dictator is increasing across the subject pairs.

3.4 Theoretical Framework and Predictions

3.4.1 Utility Form

The utility form proposed here accounts for other-regarding preferences but, crucially, relies on uncertainty. Dictators in the game are uncertain about the relative importance they place on the consumption of their matched partners. Since Dictators do not know the partners they are being matched against, this uncertainty is likely. Experimental data in the past, such as Andreoni and Bernheim (2009), show that audience effects and social norms influence the importance that Dictators place on others' welfare. This data somewhat justifies the consideration that players may be initially uncertain of the value they place on other players' welfare. Formally, we introduce the instantaneous consumption utility of the Dictator as:

$m(c) = f(c_d) + \theta * f(c_r)$ where c_d represents the percentage of second-stage wealth (the ECUs to be distributed) that the Dictator keeps for himself while c_r represents the percentage of second-stage wealth that the Dictator gives to the receiver, and θ represents the relative importance that the Dictator places on the receiver's outcomes. Because Cox, Friedman, and Gjerstad (2007) propose θ as a function of relative status and reciprocity, their model is not inconsistent with the claim that θ is not initially fully known to the agent. Similar to CJG, I propose that θ is a function of relative status, reciprocity, and beliefs about legitimacy of the status, giving us the functional form $\theta(s, r, l)$. $f(.)$ is a continuously differentiable, increasing, and concave function that provides diminishing marginal returns in higher shares of the total endowment. Note that if θ is known to the Dictator, the optimal choice c_r^* must satisfy the following first-order conditions: $\theta = \frac{f'(1-c_r^*)}{f'(c_r^*)}$

θ is unknown to the Dictator, except that she has a prior on the probability distribution that θ can take. The model is specified as below:

$\theta \sim N(\mu, \sigma^2)$, $\mu \sim Unif[a, b]$[5]

$f_\theta(\theta|\mu) = \frac{1}{\sqrt{2\Pi\sigma^2}} exp\left[\frac{-1}{2\sigma^2}(\theta - \mu)^2\right]$, $g(\mu) = \frac{1}{b-a}$

[5]In Appendix B1, we also consider a more informative case where μ has a normal distribution, and find similar results.

In this representation, the Dictator enters the game with a prior probability distribution over the relative importance she gives to the receiver's earnings. The expected value of this prior has a uniform distribution between the values of a and b. We propose that a and b take on the values of zero and two. Thus, we impose that, in expectation, the Dictator cares equally about her and the receiver's allocations. At the start of the game, a signal is sent to the Dictator which influences her beliefs over l, and consequently $\theta(.)$. The signal helps the Dictator form beliefs over the legitimacy of her own or her partner's endowments, which affects her beliefs over θ. Given this new potential datapoint of the realization of θ, the Dictator uses Bayesian updating to change her prior distribution over μ.

$$s(\mu|\theta = \hat{\theta}) = \frac{f_{\theta|\mu}(\hat{\theta}|\mu)g(\mu)}{\int f_{\theta|\mu}(x|\mu)g(\mu)du}$$

$$s(\mu|\theta = \hat{\theta}) \propto \frac{1}{\sigma}exp(\frac{-1}{2\sigma^2}(\hat{\theta} - \mu)^2)$$

$$s(\mu|\theta = \hat{\theta}) \propto \frac{1}{\sigma}exp(\frac{-1}{2\sigma^2}(\mu - \hat{\theta})^2)$$

The posterior distribution has a mean equal to the value of the signal received x. As the signal received $\hat{\theta}$ increases, the distribution of μ shifts to the right, which shifts the expected distribution of θ to the right. Given the first-order conditions, the distribution for the optimal choice of c_r^* shifts to the right as the signal $\hat{\theta}$ increases (as it does across subject pairs in the experiment). We label this condition Proposition 1.

Proposition 1 : As the percentage of earned ECUs in the Receiver's endowment increases (or as the percentage of earned ECUs in the Dictator's endowment decreases), the probability distribution for the optimal choice of c_r^* *that maximizes the instantaneous consumption utility shifts to the right, and Dictators are likely to give larger shares to receivers.*

In Appendix C1, we consider a more general version of utility by incorporating gain-loss Koszegi & Rabin type preferences for the Dictator, and find a similar prediction about levels of giving.

3.4.2 Hypothesis

Given Propositions 1 and 2,[6] we have the following hypothesis that is the main question explored.

Hypothesis 1 : As the percentage of earned ECUs in the Receiver's endowment increases across subjects (or as the percentage of earned ECUs in the Dictator's endowment decreases), Dictators give a larger share of the endowment to the Receivers.

We contrast this hypothesis with predictions under outcome-based theories of fairness and inequity aversion, as well as intention-based theories of reciprocity. Since outcome-based theories remain indifferent to the sources of endowments, the prediction is that the level of giving should be the same for all subject pairs in the experiment. Dictators that are 'fair' or 'inequity-averse' should remain equally fair or inequity-averse for all pairs. Intention-based theories of reciprocity consider actions taken by the matched partner, which are perceived as levels of 'kindness'. However, in the experiment, the only choice available to the receiver is how much effort to put into the task, which has no bearing on the amount of ECUs that the Dictator is splitting between parties, or the level of total first stage earnings by Dictators. Thus, there is no notion of reciprocity in our experimental setting.

The belief-based theory of guilt aversion explored in Dufwenberg and Battigalli (2007) and Charness and Dufwenberg (2006) could potentially affect the levels of giving in our experiment. If receivers expect higher amounts when the earned share of their endowments is higher, and Dictators correctly anticipate these increased expectations, then Dictators will give higher amounts to match those expectations and avoid feelings of 'guilt'. However, our experimental structure explicitly controls for the first-order beliefs of the receiver and the second-order beliefs of the Dictator. Receivers are not aware of their effort-task performance levels at any stage (including payoff), and Dictators are explicitly reminded of this fact. Thus, receivers' expectations should not differ across subject pairs, and Dictators maintain the same second-order beliefs about receivers' expectations. Since guilt-aversion theory predicts that levels of giving should

[6] As presented in Appendix C1

depend only on outcomes, receiver strategies, and Dictator's second-order beliefs of about receiver expectations, guilt-aversion predicts the same level of giving across all subject pairs.

3.5 Results

A total of 124 registered undergraduate subjects participated in seven experimental sessions. The sessions were run at the Economic Science Laboratory at the University of Arizona in March and April, 2018. The ESL system sent standardized invitations to students who had voluntarily registered before the experiment. The experiments were conducted using an interface programmed using Z-Tree (Fischbacher, 2007). Subjects were assigned computers randomly, instruction papers were handed out, and read out loud at the start of each session. Subjects had to undertake a quiz in order to ensure that they understood the instructions, and no one was able to start the experiment without everyone correctly answering all the quiz questions. The final payoffs consisted of the sum of earnings from the two stages of the experiment, as well as a $5 show-up fee. With the completion of the experiment, subjects were paid their cash earnings in private. The average payoff was $18.3 for sessions that lasted approximately an hour, with a minimum of $10.2, and a maximum of $25.9.

Table 3.2 below provides summary statistics for the earnings and demographics averages of the 62 subject pairs that participated in the sessions, with details provided separately for Dictator and receiver subjects. On average, subjects made 206 ECUs in the first stage, 194 of which were 'earned' through the effort task by subjects that became Dictators, but only 121 of which were 'earned' by subjects that became receivers[7]. On average, subjects that became Dictators earned a total of 305 ECUs (about $15) through two stages, but there is a significant amount of variation in earnings. Receivers were expecting an average of 41 out of 120 ECUs in the second stage, but received a significantly lower 22 ECUs. Both Dictator and receiver groups are about equally represented by male and female subjects (slighly more females in both groups). Subjects of both groups are on average, juniors who are about 21 years of age.

[7]The conversion rate used in the experiment was 20 ECUs = $1.

Table 3.2: Summary Statistics

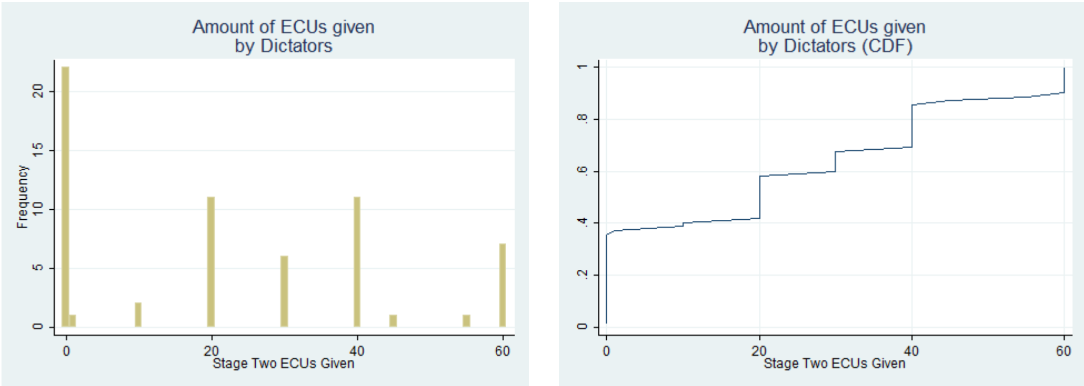
ECU Variables	Means (S.D.)	Demographic Variables	Means (S.D.)
Total Stage One ECUs	206.2 (56.39)	Dictator's Gender	0.532 (0.503)
Dictator's Earned Stage One ECUs	194.3 (57.15)	Receiver's Gender	0.548 (0.502)
Receiver's Earned Stage One ECUs	121.0 (48.47)	Dictator's Age	21.18 (1.625)
Stage Two ECUs Given	22.27 (21.07)	Receiver's Age	21.03 (1.639)
Receiver's Stage Two Expectations	41.05 (21.22)	Dictator's Academic Year	3.452 (1.210)
Dictator's Total ECUs	305.7 (64.96)	Receiver's Academic Year	3.065 (0.921)
Observations	62		

Dictators had 120 ECUs available to them for distribution in the second stage, of which they distributed about 22 ECUs on average, while keeping the rest. Thus, on average, subjects gave about a sixth of their available endowment in the second stage, which is similar to giving levels in past experiments where subjects earned the right to be Dictators like Hoffman, McCabe, and Smith (1994). However, in contrast to their results, where over half of subjects that earned the right to become Dictators give nothing, we find that only about a third of the 62 dictators gave nothing in the second stage. Only 6 out of the 62 Dictators gave half of the 120 ECUs to their matched partner. Figure 3.1 provides both the frequency distribution and the cumulative density function of the amount of 120 ECUs that were given by the Dictator to their matched partner in the second stage.

3.5.1 Effect of 'deservingness' on giving

Result 1 : In the absence of ex-ante inequality, when controlling for beliefs of receivers, there is no systematic relationship between signals of receiver 'deservingness' and prosocial behavior. However, there is a negative correlation between 'self-deservingness' and prosocial behavior.

Figure 3.1: Giving Levels



Our central hypothesis is that there is a monotonic relationship between signals of 'deservingness' and prosocial behavior by agents. Our theoretical model predicts that this relationship should exist independent of any beliefs that receivers may have regarding their own level of 'deservingness', and should exist even in the absence of ex-ante inequality. We consider three different possible signals of 'deservingness' in our analysis.

The first signal is the ratio of first-stage endowments of the matched partner that was 'earned' rather than a result of luck. While studies in the past have compared cases where income is either completely a result of effort or luck, we consider a scenario where endowments are a function of both. Subjects face matched partners with heterogeneous ratios of their first stage endowments resulting from effort, as opposed to luck. The experimental setup ensures that there is no inequality within subject pairs at the end of the first stage. We find that there is no effect of this particular signal of receiver 'deservingness' on giving in the following Dictator games. In the absence of ex-ante inequality, when receiver beliefs about their own level of 'deservingness' is controlled for, prosocial behavior is not monotonically correlated with a signal of receiver 'deservingness' comprising of ratios of endowments that are a result of effort rather than luck.

The next signal of 'deservingness' we considered is the difference in ratios of 'earned' first stage earnings between Dictators and their matched partners. Again, we find no systematic relationship on giving levels in the second round based on this signal of 'deservingness' from the first round. This lack of a monotonic relationship is somewhat surprising given that receiver endowments coming from effort vs luck have been shown to be predictive of different support for redistribution in previous studies. However, these previous studies all cover situations with some ex-ante inequality between subjects prior to the redistribution stage, and also rarely control for receiver's beliefs about their own 'deservingness'. Our experimental results suggest that the monotonic relationship breaks down in the absence of ex-ante inequality, or relies crucially on receivers having correct beliefs about their own level of 'deservingness'.

Our main finding is the negative relationship between the ratio of Dictator's own income that was a result of effort to giving in the second round. Dictators whose endowments in the first stage are more a result of effort than luck legitimize their claims to future winnings and give lower amounts to matched partners in the absence of ex-ante inequality.

Table 3.3 provides the results of regression analysis studying the determinants of the ratio of 120 ECUs that Dictators gave in the second round. Since this variable is truncated at the lower end at zero (and about a third of the subjects gave zero in the second round), we show the results of the OLS model as well as the Tobit model. In our analysis, we control for the gender and age of Dictator subjects, and cluster all standard errors at the session level. For every percentage point increase in the ratio of stage one Dictator earnings that results from effort (vs luck), there is a 1.8 percentage point drop in giving in the second stage.

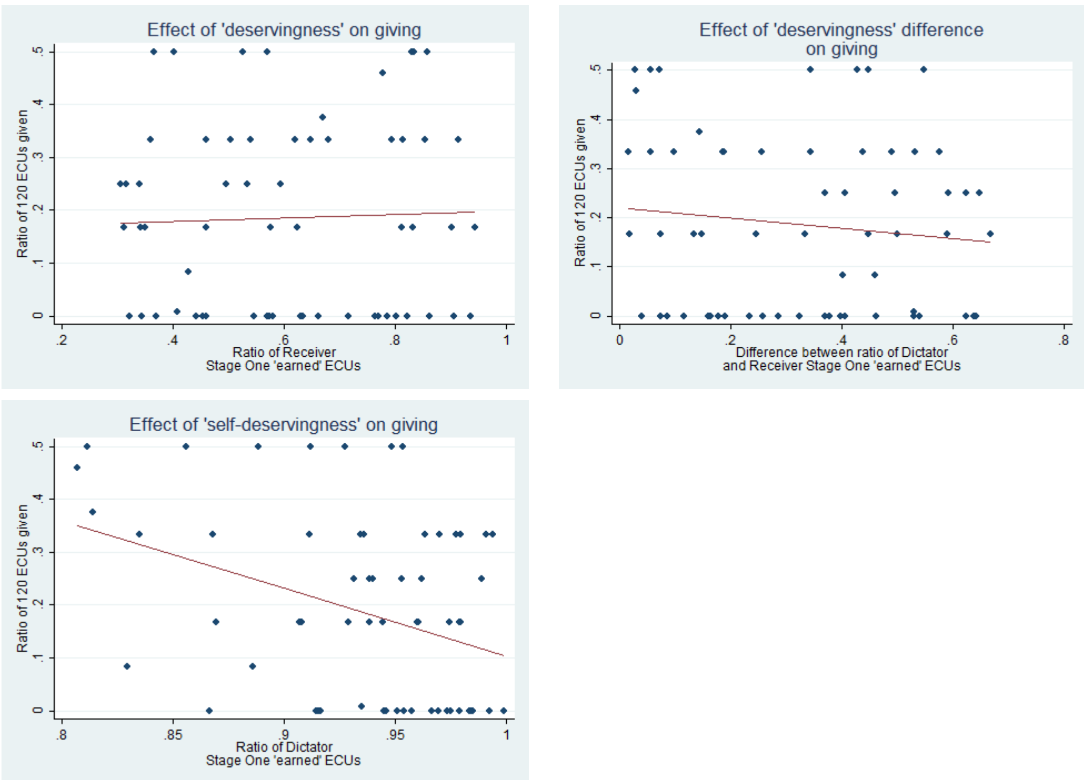
Table 3.3: The Effect of Deserving Signals on Ratio of Prosocial Behavior

DV: Ratio of 120 ECUs Given	(1)	(2)	(3)	(4)	(5)	(6)
Receiver's Earned Ratio	0.0300			-0.0191		
	(0.101)			(0.172)		
Dictator's Earned Ratio		-1.288**			-1.817***	
		(0.383)			(0.458)	
Difference in Earned Ratio			-0.0920			-0.0778
			(0.117)			(0.187)
Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes
Age Controls	Yes	Yes	Yes	Yes	Yes	Yes
Session Controls	Yes	Yes	Yes	Yes	Yes	Yes
Session Clusters	Yes	Yes	Yes	Yes	Yes	Yes
Model Type	OLS	OLS	OLS	Tobit	Tobit	Tobit
N	62	62	62	62	62	62
R^2	0.0224	0.129	0.0324			

Standard Errors Clustered at the Session level

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

Figure 3.2: What Affects Giving Ratio?



Result 2 : As the level of wealth (as represented by first stage endowment) increases across subjects, Dictators give a smaller share of second stage ECUs to Receivers.

One interesting result we find is the statistically significant negative relationship between wealth levels in the first stage with giving in the second round. Wealthier Dictators are less generous towards their matched partners in later rounds. This provides a complement to the results in Rutstrom and Williams (2000). They compare subjects in discrete income levels, and find that Dictators in higher income levels are less prosocial than those in lower income levels. Our results suggest that these results in discrete groups also extend in a monotonically continuous way. Table 3.4 provides the results of regression analysis studying the determinants of the amount of 120 ECUs that Dictators gave in the second round. Since this variable is truncated at the lower end at zero (and about a third of the subjects gave zero in the second round), we show the results of the OLS model as well as the Tobit model. In our analysis, we control for the gender and age of Dictator subjects, and cluster all standard errors at the session level. A one ECU increase in the level of overall earnings in the first stage is associated with a 0.15 ECU decrease in giving in the second stage.

We find that there is no statistical correlation between receiver's expectations in the second stage and the amounts they actually received. We also check to see if the amount (rather than the ratio) of 'earned' Dictator ECUs in the first stage is predictive of the amount (rather than the ratio) of 120 ECUs that were given in the second stage. We replicate the statistically significant negative relationship that we find between the ratio versions of these variables. This negative relationship exists even after controlling for the ratio of receiver earnings that are a result of effort, which serves as a proxy for receiver 'deservingness'.

We do find some evidence of a positive correlation between Dictator's beliefs about the ECUs that their partner was expecting in the second stage, and the ECUs they actually gave in the second stage. Since these beliefs are elicited **after** Dictators have already chosen the amount of ECUs to give in the second stage, we are unable to distinguish between some possible explanations for this trend in the data. It is possible that Dictators are simply exhibiting some wish-fulfilment by saying that receivers were expecting lower (higher) amounts if they actually gave lower (higher) amounts in the second stage. However, we cannot reject the possibiity that Dictators plausibly had the same beliefs ex-ante, and so the beliefs about their partner's expectations are what led to the giving levels in the second stage. We are skeptical of this latter explanation because there is no correlation between Dictator's beliefs about receiver expectations, and receiver's actual expectations. Furthermore, there is no correlation between receiver's expectations, and actual giving levels.

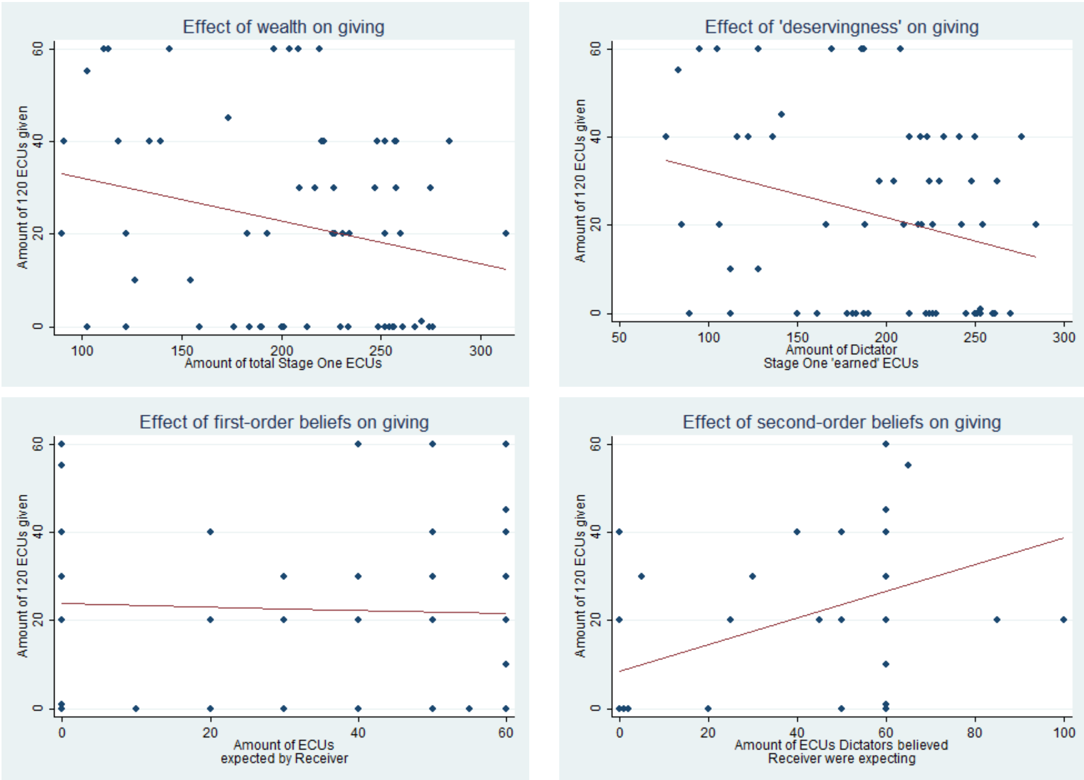
Table 3.4: The Effect of Deserving Signals on Amount of Prosocial Behavior

DV: Amount of 120 ECUs Given	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Stage One Total ECUs	-0.110*				-0.154**			
	(0.0480)				(0.0694)			
Receiver's Earned Ratio	-6.571	-6.803			-17.05	-17.42		
	(6.433)	(6.284)			(12.99)	(12.93)		
Dictator's Earned ECU Amount		-0.118*				-0.167**		
		(0.0509)				(0.0704)		
Dictator's Stage Two Expectations			0.313**				0.563***	
			(0.0894)				(0.158)	
Receiver's Stage Two Expectations				-0.0292				-0.0564
				(0.130)				(0.146)
Gender Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Session Clusters	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Model Type	OLS	OLS	OLS	OLS	Tobit	Tobit	Tobit	Tobit
N	62	62	62	62	62	62	62	62
R^2	0.0939	0.110	0.163	0.0222				

Standard Errors Clustered at the Session level

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

Figure 3.3: What Affects Giving Amounts?

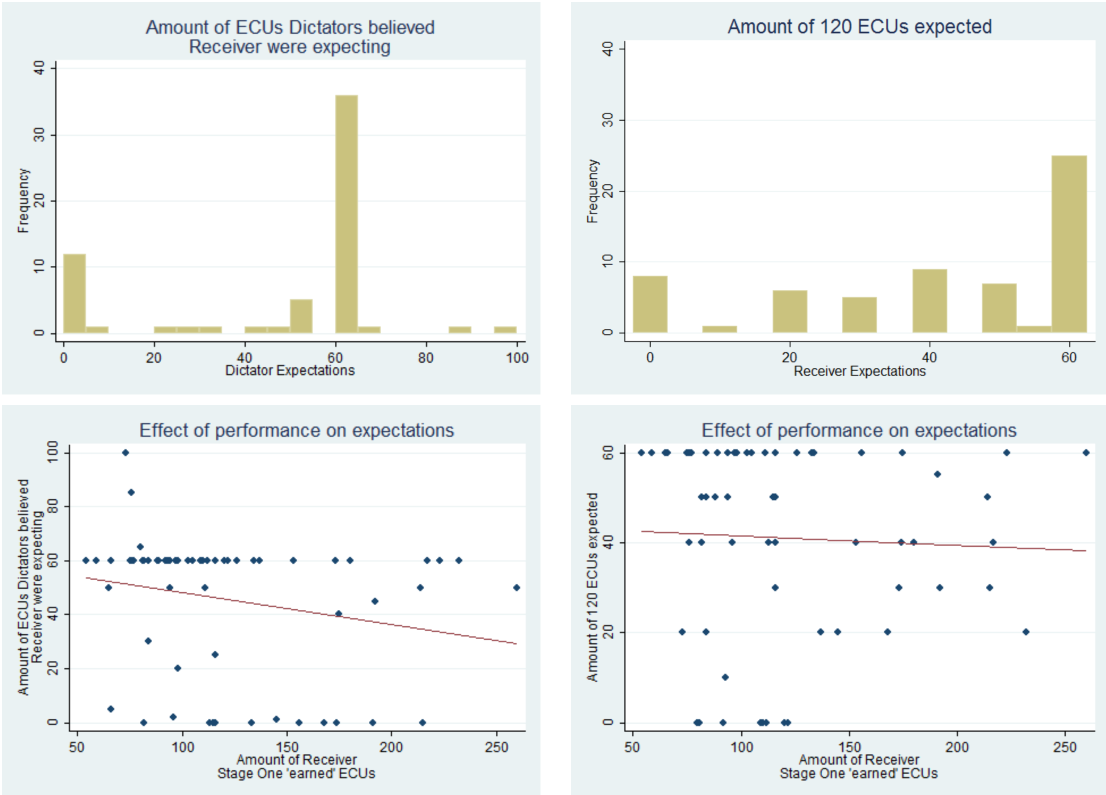


3.5.2 Effect of effort levels on expectations

Result 3 : There is no positive correlation between the level of receiver's performance in the first stage, and Dictators' beliefs about receiver expectations in the second stage.

Figure 3.4 shows the results for the amount of ECUs out of 120 that Receivers were expecting in the second stage, as well as the amount of ECUs that Dictators believed their matched partners were expecting in the second stage. Receivers were very optimistic, with slightly over a third of all Receivers expecting to receive half of the 120 ECUs in the second round. Half of all Dictators say that they believe their matched partner is expecting half the second round ECUs. However, both the Receiver and Dictator's expectations of the second round ECUs is completely uncorrelated with the actual effort levels of Receivers in the first round.

Figure 3.4: Effort and Expectations



Note that Dictators only provided their beliefs about their partner's expectations after they had already decided the actual amount of ECUs to give in the second stage. The experiment setup ensured that receivers were not aware of their exact first stage level earnings, so there should be no effect of the amount of ECUs that they 'earned' in the first round on their expectations in the second round. Similarly, Dictators were reminded that receivers would not know the breakdown of their earnings, so their beliefs about receiver expectations in the second round should also be uncorrelated with the amount of ECUs that receivers 'earned' in the first round. We find that the experimental setup to try and control first-order and second-order beliefs in the second stage was successful, with receiver's second stage expectations surprisingly having a slightly negative correlation to their level of 'earned' ECUs in the first stage (as shown in the first column of Table 3.5), and Dictator's second stage expectations completely uncorrelated with the receiver's level of 'earned' ECUs in the first stage (as shown in the second column of Table 3.5).

Table 3.5: The Effect of Effort Levels on Beliefs

	(1)	(2)
Receiver's Earned ECU Amount	-0.110*	-0.0158
	(0.0516)	(0.0405)
Gender Controls	Yes	Yes
Age Controls	Yes	Yes
Session Controls	Yes	Yes
Session Clusters	Yes	Yes
N	62	62
R^2	0.0721	0.0107

Standard Errors Clustered at the Session level
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.6 Conclusion

This paper presents a theoretical model that provides a basis for how antecedents can affect other-regarding preferences. We posit that signals of self and receiver 'deservingness' changes the behavior of economic agents by affecting their marginal rate of substitution between their concern for self-interested and other-regarding utility. This model can help explain behavior seen in past experimental studies which have used signals of receiver 'deservingness' by considering sources of past income, or legitimacy of past income. These studies have covered other-regarding preferences through the lens of proclivity towards redistribution in the face of ex-ante inequality, and contrast cases where ex-ante inequality is the result of unequal levels of effort, or random luck. Many of these results can also be explained by psychological game theoretical models of guilt aversion where receivers may have expectations based on their effort levels in the past, or different beliefs about their own 'deservingness', and redistribution is done to match these expectations. The theoretical model presented here suggests a channel outside of receiver expectations, and suggests a monotonic relationship between prosocial giving and 'deservingness' even in the absence of ex-ante inequality.

To this end, we conduct experimental sessions that test this predicted monotonic relationship. While earlier studies have looked at cases where income is either completely the result of effort, or completely a lucky outcome, we study a more continuous distribution where incomes are a combination of effort and luck. We conduct a specific randomization to ensure that there is no ex-ante inequality between subjects in the same pair before there is any avenue for redistribution. We find that, in contrast to some earlier studies, that there is no monotonic relationship between prosocial behavior and signals of receiver 'deservingness' in the absence of ex-ante inequality, when receiver beliefs about their own level of 'deservingness' are suppressed. We do however, find a statistically significant negative relationship between prosocial behavior and the level of Dictator's own 'deservingness'. This relationship is consistent even when controlling for factors like the age and gender of participants, and also after controlling the level of receiver 'deservingness'. Finally, we also find a statistically significant negative relationship between levels of income and prosocial behavior.

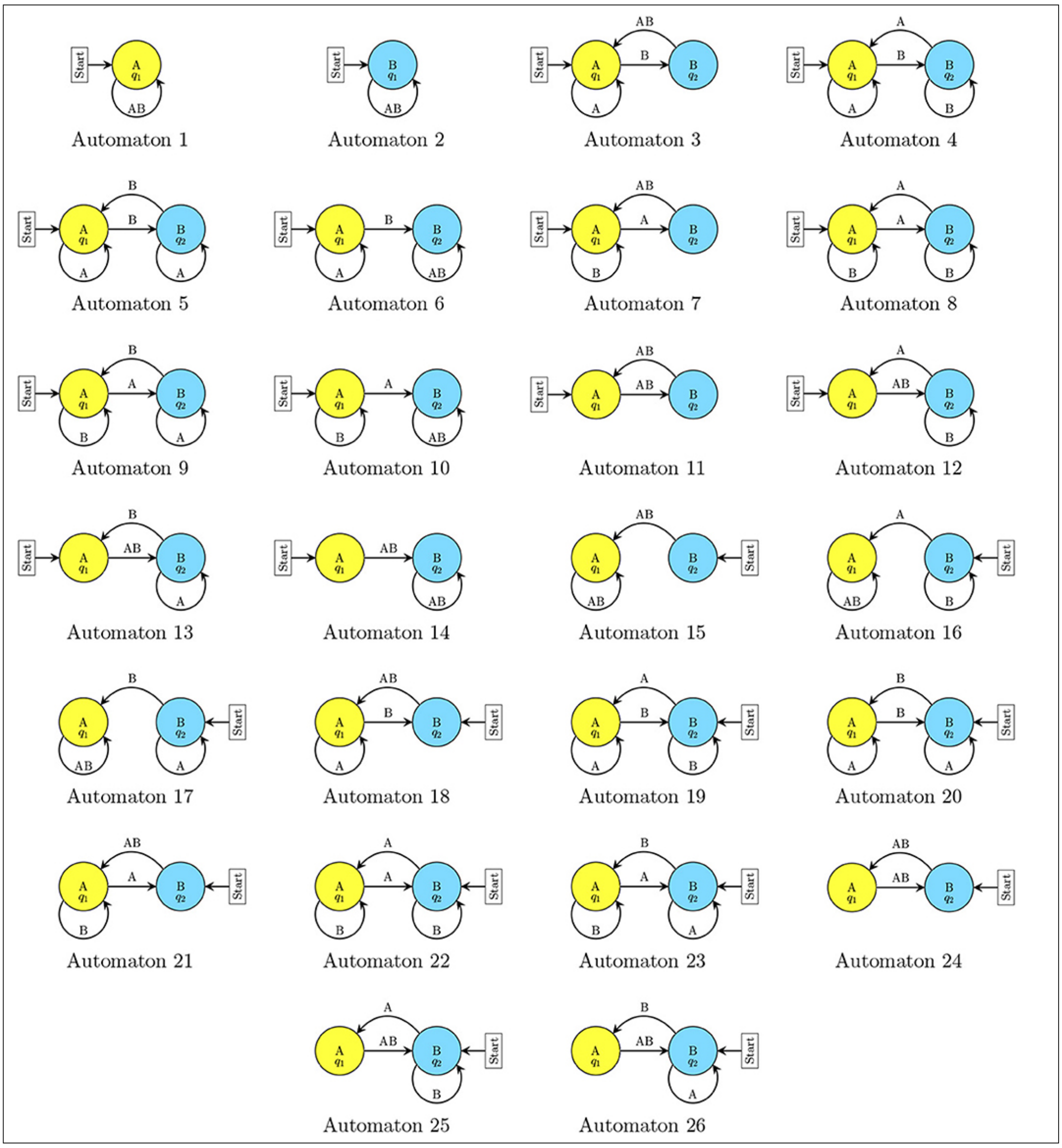
Our results weakly suggest that pay compression in the workforce need not necessarily lead to adverse outcomes or feelings of ill-will within groups who are at similar designation levels, and receive similar pay irrespective of productivity levels. However, any ill-will that is generated will be higher for group members who are at the top end of productivity, because they have the largest claim to legitimizing their income levels. Going forward, it would be interesting to test the source of the null results we find in our study regarding the connection between receiver 'deservingness' and level of prosocial behavior. The breakdown in the relationship between giving and the level of receiver 'deservingness' could either be attributed to the lack of ex-ante inequality, but could also be driven by the fact that receiver beliefs about their own level of 'deservingness' is suppressed, and hence guilt aversion channels are not being activated. We plan to study this in further detail in future projects. The key result we present here is that agents pay more attention to their own level of 'deservingness' rather than those of others when considering the optimal level of prosocial behavior.

Chapter 1 Appendix

A1 Repeated Game Strategy Automata

The following diagram shows the 26 strategies evaluated, represented as finite automata.

Figure A1: Two State Two Action Automata



A2 Efficiency & Fairness Metrics

The following diagrams show how the efficiency and fairness metrics are calculated for the Battle of the Sexes, Prisoner's Dilemma, Chicken, and Stag Hunt games.

Battle of the Sexes

Played repeatedly, the Battle of the Sexes is a simple, canonical game in which there is a tension between individual strategies (stage game Nash Equilbria) and coordination (alternating strategies that share the payoffs). The payoffs satisfy the following properties: $b^1 > c^1 > a^1 = d^1$ and $c^2 > b^2 > a^2 = d^2$. There are three Nash Equilibria in the stage game: Player 1 plays action A and Player 2 plays action B, Player 1 plays action B and Player 2 plays action A, and a mixed-strategy equilibrium. The Pareto set is $\{u = (u^1, u^2) | u^i = \lambda u^i(B, A) + (1 - \lambda) u^i(A, B)\}$, which is attained by alternating between the two pure-strategy Nash equilibria. For example, if the two players choose (B, A) for k rounds and then switch to (A, B) for j rounds and then switch back to (B, A) and so forth, the average payoffs would be $(\frac{c^1 k + b^1 j}{k+j}, \frac{c^2 k + b^2 j}{k+j})$ with $\lambda = \frac{k}{k+j}$ giving us the alternation ratio. By suitably choosing k and j, one can get many of the rational payoffs in the efficient set of payoffs, including a λ^*that equalizes the realized payoffs for the two players.

I define efficiency and fairness here using the following routine. Since (a^1, a^2) is the least efficient payoff outcome in the game, and the Pareto set generates the most efficient set of payoffs in the game, I draw a line l_1 from the sink point of the payoffs (a_1, a_2) to the payoffs on the Pareto frontier passing through the actual payoff point $(u^{*,1}, u^{*,2})$. Efficiency is given by the distance between the actual payoff point and the sink point, relative to the distance between the sink point and the Pareto frontier. Arifovic and Ledyard (2015) have a closed form solution for this measure that generalizes across all symmetric and asymmetric payoff games. The efficiency measure is:

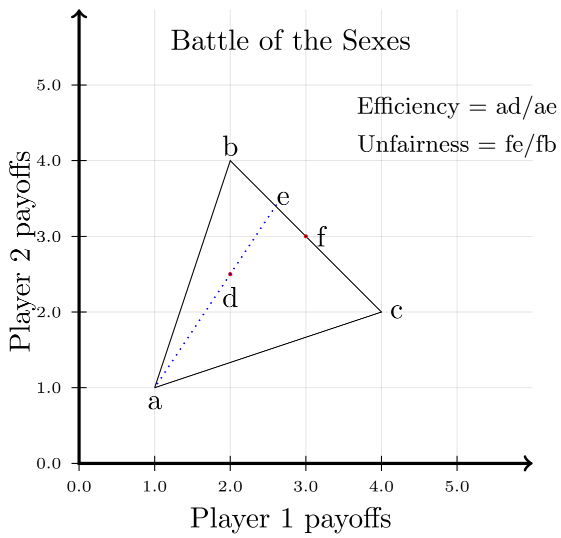
$$E = \frac{(u^{*,2} - d^2)(b^1 - c^1) - (u^{*,1} - d^1)(b^2 - c^2)}{(b^1 - c^1)(c^2 - d^2) - (b^2 - c^2)(c^1 - d^1)} \times 100$$

If payoffs are symmetric, the measure of unfairness is given by how far the actual payoff is from a 50-50 split. With asymmetric payoffs, the routine requires a few more steps. Suppose the players were able to coordinate on a 100% efficient strategy; that is, they alternate in some pattern between player 1's favorite stage game Nash equilibrium and that of player 2. Let λ be the % of time spent on player 1's favorite Nash equilibrium. Equal payoffs occur if $\lambda = \lambda^* = \frac{(c^2-c^1)}{b^1-b^2+c^2-c^1}$. Let the point (e^1, e^2) denote the point where the line from (a^1, a^2) to the line l_1 that passes through the average payoff pair $(u^{*,1}, u^{*,2})$. The point (e^1, e^2) can also be represented by $\bar{\lambda}$, where $\bar{\lambda}$ is the % of time spent on (b^1, b^2) and $(1-\bar{\lambda})$ is the % of time spent on (c^1, c^2). Unfairness can be measured by how far the value of $\bar{\lambda}$ is from the idealized value of λ^*. Arifovic and Ledyard (2015) present a closed form solution that generalizes to all symmetric and asymmetric payoffs.

$$U = \frac{\left|\frac{(u^1-d^1)(c^2-d^2)-(u^2-d^2)(c^1-d^1)}{-(u^1-d^1)(b^2-c^2)-(u^2-d^2)(b^1-c^1)} - \lambda^*\right|}{1-\lambda^*} \times 100$$

The fairness measure is given by subtracting the unfairness measure from 100. Thus, all efficiency and fairness measures are on a scale from 0 to 100.

Figure A2: Battle of the Sexes - Efficiency & Fairness



Prisoner's Dilemma

In a repeated setting, the Prisoner's Dilemma is a game with a tension between the dominant strategy of defection against the payoff-dominant situation of mutual co-operation. The payoffs satisfy the following properties: $b^1 > a^1 > d^1 > c^1$ and

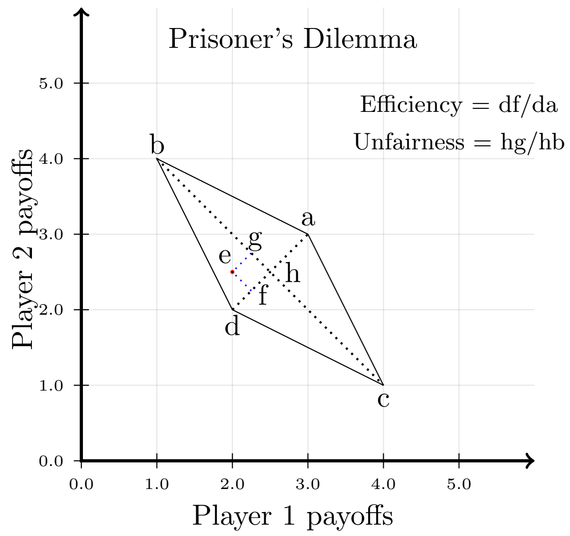
$c^2 > a^2 > d^2 > b^2$. There is a unique Nash Equilibria in the stage game: (B, B). The most efficient payoff is the mutual co-operation outcome: (A,A).

I define efficiency and fairness here using the following routine. On the co-ordinate system, I draw a line from the point (a^1, a^2) to (d^1, d^2). I denote this line as l_1. Then, I have lines l_2 and l_3 that are perpendicular to line l_1 and originate from points (b^1, b^2) and (c^1, c^2) respectively. For any point denoting the average payoff of the two players $(u^{*,1}, u^{*,2})$, I can project the point onto l_1. Denote this point (f^1, f^2). Efficiency is given by the ratio of the distance from (a^1, a^2) to (f^1, f^2), divided by the distance from (a^1, a^2) to (d^1, d^2).

To generate a fairness index, I calculate where the point $(u^{*,1}, u^{*,2})$ lies relative to the line l_1. If it lies above the line l_1, I project the point $(u^{*,1}, u^{*,2})$ to the line l_2 and call this the point (g^1, g^2). The unfairness index is given by the ratio of the distance from l_1 to (g^1, g^2), divided by the distance from l_1 to (b^1, b^2). If it lies below the line l_1, I project the point (u^1, u^2) to the line l_3 and call this the point (m^1, m^2). The unfairness index is given by the ratio of the distance from l_1 to (m^1, m^2), divided by the distance from l_1 to (c^1, c^2).

The fairness measure is given by subtracting the unfairness measure from 100. Thus, all efficiency and fairness measures are on a scale from 0 to 100.

Figure A3: Prisoner's Dilemma - Efficiency & Fairness



Chicken

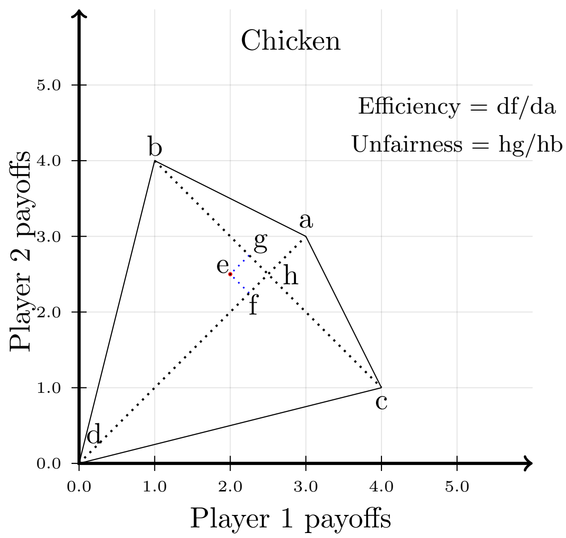
Played repeatedly, the Chicken game is a form of anti-coordination game, where it is mutually beneficial for the players to play different actions. The payoffs satisfy the following properties: $b^1 > a^1 > c^1 > d^1$ and $c^2 > a^2 > b^2 > d^2$. There are three Nash Equilibria in the stage game: Player 1 plays action A and Player 2 plays action B, Player 1 plays action B and Player 2 plays action A, and a mixed-strategy equilibrium. The most efficient payoff is the mutual co-operation outcome: (A,A). Graphically, the payoff structure in a Chicken game is not very different from that of a Prisoner's Dilemma game, and hence the efficiency and fairness indices are calculated using a similar routine.

On the co-ordinate system, I draw a line from the point (a^1, a^2) to (d^1, d^2). I denote this line as l_1. Then, I have lines l_2 and l_3 that are perpendicular to line l_1 and originate from points (b^1, b^2) and (c^1, c^2) respectively. For any point denoting the average payoff of the two players $(u^{*,1}, u^{*,2})$, I can project the point onto l_1. Denote this point (f^1, f^2). Efficiency is given by the ratio of the distance from (a^1, a^2) to (f^1, f^2), divided by the distance from (a^1, a^2) to (d^1, d^2).

To generate a fairness index, I calculate where the point $(u^{*,1}, u^{*,2})$ lies relative to the line l_1. If it lies above the line l_1, I project the point $(u^{*,1}, u^{*,2})$ to the line l_2 and call this the point (g^1, g^2). The unfairness index is given by the ratio of the distance from l_1 to (g^1, g^2), divided by the distance from l_1 to (b^1, b^2). If it lies below the line l_1, I project the point (u^1, u^2) to the line l_3 and call this the point (m^1, m^2). The unfairness index is given by the ratio of the distance from l_1 to (m^1, m^2), divided by the distance from l_1 to (c^1, c^2).

The fairness measure is given by subtracting the unfairness measure from 100. Thus, all efficiency and fairness measures are on a scale from 0 to 100.

Figure A4: Chicken - Efficiency & Fairness



Stag Hunt

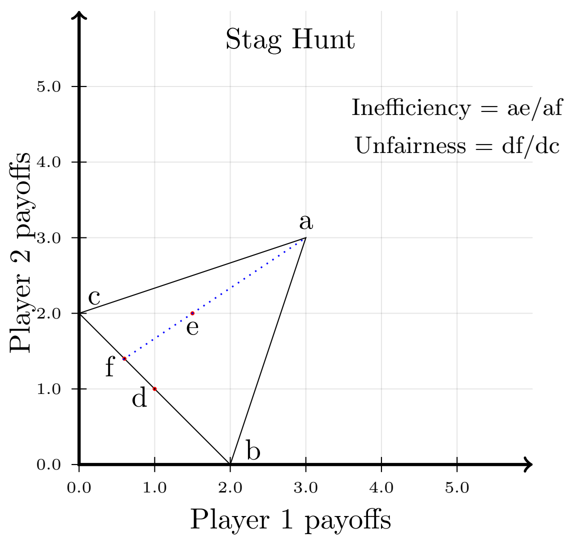
In a repeated setting, the Stag Hunt game represents a coordination game where there is a tension between the risk-dominant strategy of safety and the payoff-dominant situation of mutual cooperation. The payoffs satisfy the following properties, $a^1 > b^1 > d^1 > c^1$ and $a^2 > c^2 > d^2 > b^2$. There are three Nash Equilibria in the stage game: Both players mutually cooperating on action A, both players mutually playing action B, and a mixed-strategy equilibrium. The most efficient payoff is the mutual cooperation outcome: (A,A).

I define efficiency and fairness in Stag Hunt games using the following routine. On the co-ordinate system, I draw a line from the point (b^1, b^2) to (c^1, c^2). I denote this line as l_1. For any average payoff pair, $(u^{*,1}, u^{*,2})$, I can draw a line from (a^1, a^2) to the line l_1 that passes through the average payoff pair $(u^{*,1}, u^{*,2})$. The line l_1 represents the most inefficient range of payoffs in the game, while the point (a^1, a^2) represents the most efficient payoff. Inefficiency is given by the distance between the actual payoff point and the point (a^1, a^2) , relative to the distance between the point (a^1, a^2) and the line l_1. The efficiency measure is given by subtracting the inefficiency measure from 100.

As a measure of unfairness, if payoffs are symmetric, the measure is given by how far the actual payoff is from a 50-50 split. With asymmetric payoffs, I come up with a metric that generalizes to all symmetric and asymmetric games. Let the point (f^1, f^2) denote the point where the line from (a^1, a^2) to the line l_1 that passes through the average payoff pair $(u^{*,1}, u^{*,2})$. The point (f^1, f^2) can also be represented by $\bar{\lambda}$, where $\bar{\lambda}$ is the % of time spent on (b^1, b^2) and $(1-\bar{\lambda})$ is the % of time spent on (c^1, c^2). Equal payoffs occur if $\lambda = \lambda^* = \frac{(c^2-c^1)}{b^1-b^2+c^2-c^1}$. Unfairness of an actual outcome of repeated play is measured by how far away it is from this idealized result.

The fairness measure is given by subtracting the unfairness measure from 100. Thus, all efficiency and fairness measures are on a scale from 0 to 100.

Figure A5: Stag Hunt - Efficiency & Fairness



A3 Theoretical Proofs

I present here a proof that guarantees the existence of a fixed point for propensities towards strategies, given any starting point of initial propensities. While the games studied in the paper have two players and two actions, the proof technique is generalised for any finite number of players, with each having a finite number of actions. The proof is similar in spirit to the one used to show the existence of a Quantal Response Equilibrium in McKelvey and Palfrey (1995).

Let the game be played by n players, and let each player $i \in N = \{1, 2, ..., n\}$ have J_i strategies. Initial propensities towards each strategy for player i are given by $p_i = (p_{i,1}, p_{i,2}, ..., p_{i,J_i})$ such that $\sum_{k=1}^{J_i} p_{i,k} = 1$. Every element $p_i \in \Delta_i$ which is a simplex of size J_i. Taken across all players, I have $p = (p_1, p_2, ..., p_i, ..., p_n) \in \Delta$ where $\Delta = \Pi_{i \in N} \Delta_i$.

Let $s_{i,j} \in S_i$ denote strategy j for player i, an element of all strategies available to player i. The payoff function maps from the strategy space into $\mathbb{R}^+$. The algorithm defined in Section 2.3 and Section 2.4 then maps from the payoff space into a composite score $H = \alpha E + \beta AF + \delta RF + (1 - \alpha - \beta - \delta)S$. Thus, I have a function that maps from the strategy space into the composite score space, which is $\mathbb{R}^+$. For each player i, I have:

$g_i(s_i, s_{-i}) \to \mathbb{R}^+$ where $s_i \in S_i, s_{-i} \in S_{-i}$ and $S_i \times S_{-i} = S$

I can extend the domain of g_i to Δ which allows probability distributions over strategies by using the following rule:

$g_i : S \times \Delta \to \mathbb{R}^+$ where $g_i(s_i, s_{-i}; p) = \sum_{s \in S} p(s) g_i(s)$ where $p(s) = \Pi_{i \in N} p_i(s_i)$. Given the probability distribution over strategies being used by other players, player i's scaled expected composite score from using strategy j is given by
$\bar{g}_{i,j}(p) = exp(\lambda g_i(s_{i,j}, p_{-i}))$. Finally, I have a function that takes a set of probabilities for the n players, and returns an updated set of probabilities using the following algorithm.

$l(p_{i,j}) = \frac{\bar{g}_{i,j}(p)}{\sum_{k=1}^{J_i} \bar{g}_{i,k}(p)}$ for $p \in \Delta$

Thus, $l(.)$ is a function that maps Δ into Δ. Since $g_i(.)$ is continuous in p, $\bar{g}_{i,j}(.)$ is also continuous in p. Sums of continuous functions are continuous, and $\frac{f(x)}{g(x)}$ is continuous if $f(x)$ and $g(x)$ are continuous in x, and $g(x) \neq 0$. Given these properties, $l(.)$ is a continuous function in p, and maps Δ(a nonempty, compact, and convex subset of $\mathbb{R}^n$) into itself. By Brouwer's fixed point, the existence of a fixed point is guaranteed.

A4 Converged Fixed Points

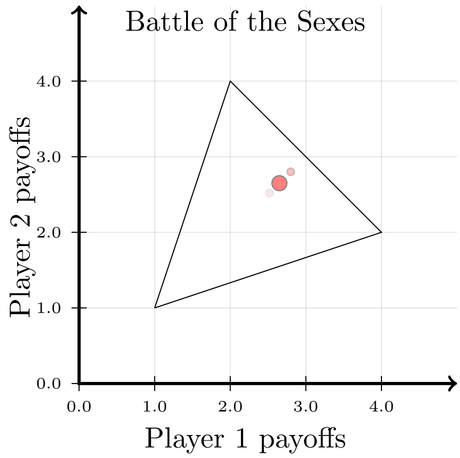
In the analysis, the 'pre-experimental' phase returns the final probability distributions for both players after a convergence algorithm. The starting point for the convergence algorithm is assumed to be a uniform distribution across all 26 strategies for both players. If the players follow this converged probability distribution, the average payoffs they would receive are highlighted in the following diagrams in large red circles. As a robustness check, I have run simulations using 100,000 random starting probability distributions to see if the selection of uniform distributions as a starting point is restrictive. The 100,000 random starting probability distributions converge to a small set of final probability distributions, and the average payoffs for players if they follow these probability distributions is showing in the following diagrams in small red circles.

Battle of the Sexes

Full Information, Symmetric Payoffs

In the Battle of the Sexes, the 100,000 starting probability distributions converge to one of four final fixed points. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by these four probability distributions. The payoffs are concentrated in a small region within a straight line, suggesting that the four points are likely mixed representations of the two probability distributions at the ends of this line. The large red circle representing the payoff when starting from the uniform distribution is almost at the centre of these two ends, suggesting that it is a good representation of all possible 'types' of players' 'pre-experimental' preferences.

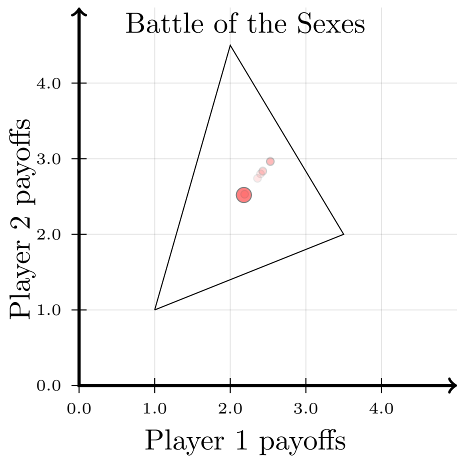
Figure A6: Battle of the Sexes - Full Information, Symmetric Payoffs Multiple Equilibria



Full Information, Asymmetric Payoffs

The 100,000 starting probability distributions converge to one of 8 final fixed points. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by these 8 probability distributions. The payoffs are concentrated in a small region within a straight line, suggesting that the 8 points are likely mixed representations of the two probability distributions at the ends of this line. The large red circle representing the payoff when starting from the uniform distribution is almost at one of the two ends, so it is not an ideal representation of all possible 'types' of players' 'pre-experimental' preferences, but comes reasonably close.

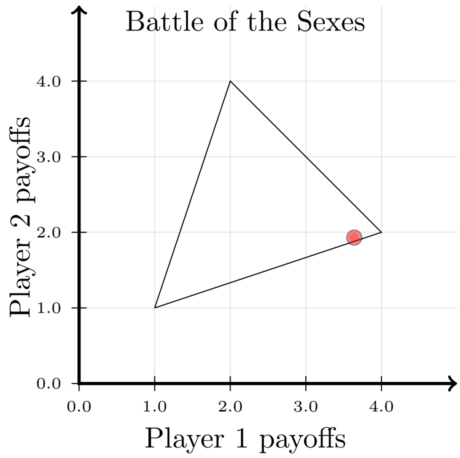
Figure A7: Battle of the Sexes - Full Information, Asymmetric Payoffs Multiple Equilibria



Dark Information, Symmetric Payoffs

The 100,000 starting probability distributions converge to one of two final fixed points. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by these two probability distributions. The payoffs are all concentrated in a very small region. The large red circle representing the payoff when starting from the uniform distribution is also very close to these two points, suggesting that it is a good representation of all possible 'types' of players' 'pre-experimental' preferences.

Figure A8: Battle of the Sexes - Dark Information, Symmetric Payoffs Multiple Equilibria

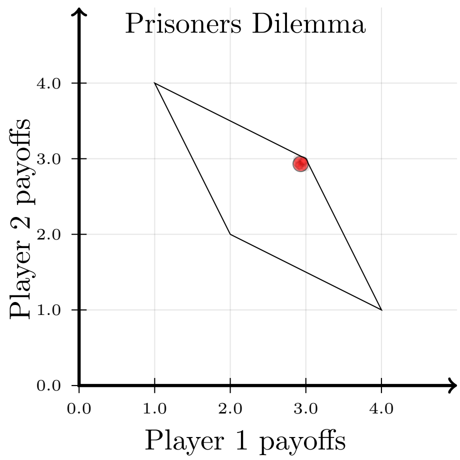


Prisoner's Dilemma

Full Information, Symmetric Payoffs

In the Prisoner's Dilemma, the 100,000 starting probability distributions converge to one of 20 final fixed points. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by these 20 probability distributions. The payoffs are all concentrated in a very small region close to the mutual cooperation outcome. The large red circle representing the payoff when starting from the uniform distribution is also very close to the mutual cooperation outcome, suggesting that it is a good representation of all possible 'types' of players' 'pre-experimental' preferences.

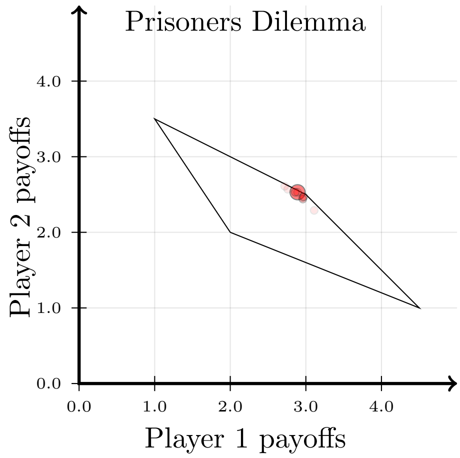
Figure A9: Prisoner's Dilemma - Full Information, Symmetric Payoffs Multiple Equilibria



Full Information, Asymmetric Payoffs

The 100,000 starting probability distributions converge to one of 16 final fixed points. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by these 19 probability distributions. The payoffs are concentrated in a very small region. The large red circle representing the payoff when starting from the uniform distribution is very close to this region, suggesting that it is a fairly reasonable representation of all possible 'types' of players' 'pre-experimental' preferences.

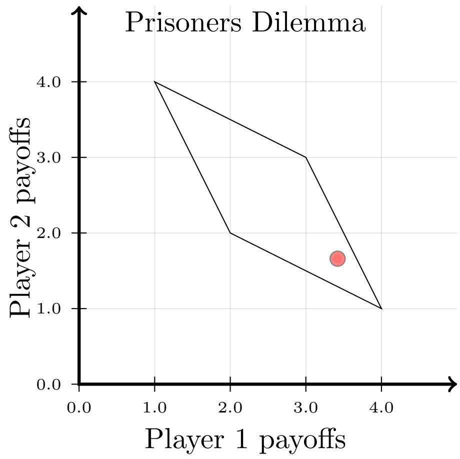
Figure A10: Prisoner's Dilemma - Full Information, Asymmetric Payoffs Multiple Equilibria



Dark Information, Symmetric Payoffs

The 100,000 starting probability distributions converge to a single final fixed point. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by this single probability distribution. The large red circle representing the payoff when starting from the uniform distribution is the only payoff outcome, suggesting that it is a good representation of all possible 'types' of players' 'pre-experimental' preferences.

Figure A11: Prisoner's Dilemma - Dark Information, Symmetric Payoffs Multiple Equilibria

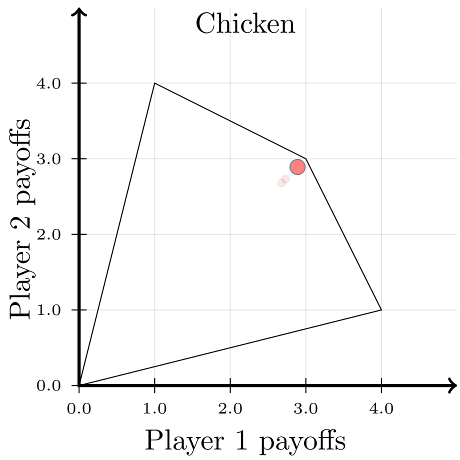


Chicken

Full Information, Symmetric Payoffs

In the Chicken game, the 100,000 starting probability distributions converge to one of two final fixed points. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by these two probability distributions. The payoffs are concentrated in a very small region. The large red circle representing the payoff when starting from the uniform distribution is very close to the two payoff points, suggesting that it is a fairly reasonable representation of all possible 'types' of players' 'pre-experimental' preferences.

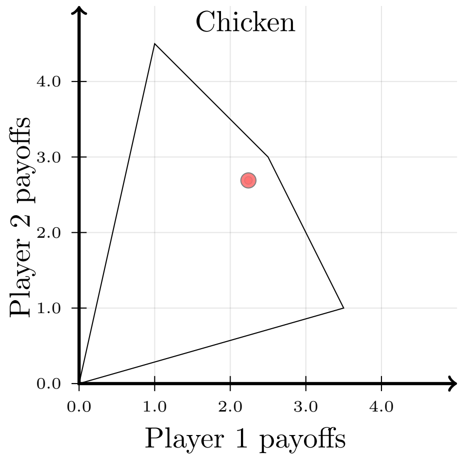
Figure A12: Chicken - Full Information, Symmetric Payoffs Multiple Equilibria



Full Information, Asymmetric Payoffs

The 100,000 starting probability distributions converge to a single final fixed point. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by this single probability distribution. The large red circle representing the payoff when starting from the uniform distribution is the only payoff outcome, suggesting that it is a good representation of all possible 'types' of players' 'pre-experimental' preferences.

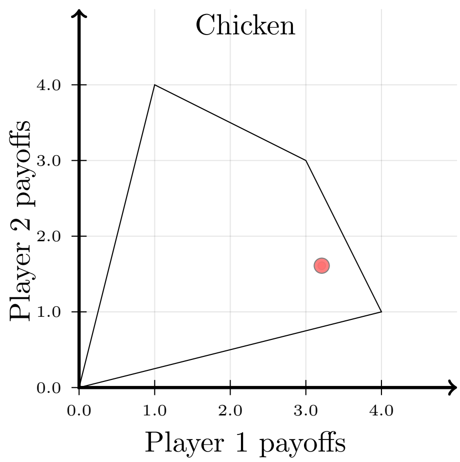
Figure A13: Chicken - Full Information, Asymmetric Payoffs Multiple Equilibria



Dark Information, Symmetric Payoffs

The 100,000 starting probability distributions converge to a single final fixed point. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by this single probability distribution. The large red circle representing the payoff when starting from the uniform distribution is the only payoff outcome, suggesting that it is a good representation of all possible 'types' of players' 'pre-experimental' preferences.

Figure A14: Chicken - Dark Information, Symmetric Payoffs Multiple Equilibria

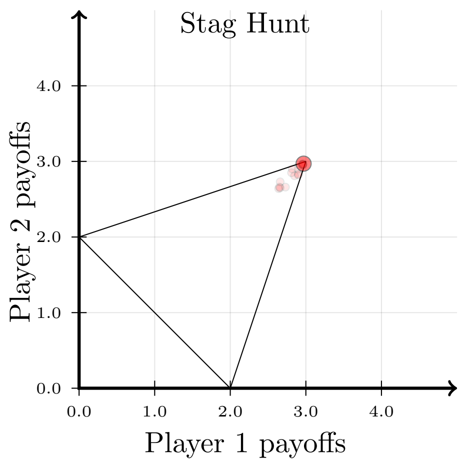


Stag Hunt

Full Information, Symmetric Payoffs

In the Stag Hunt game, the 100,000 starting probability distributions converge to one of 17 final fixed points. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by these 17 probability distributions. The payoffs are concentrated in a relatively small region close to the mutual cooperation outcome. The large red circle representing the payoff when starting from the uniform distribution is also very close to the mutual cooperation outcome, suggesting that it is a good representation of all possible 'types' of players' 'pre-experimental' preferences.

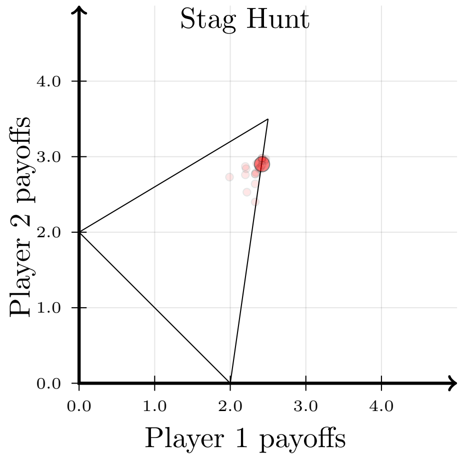
Figure A15: Stag Hunt - Full Information, Symmetric Payoffs Multiple Equilibria



Full Information, Asymmetric Payoffs

The 100,000 starting probability distributions converge to one of 19 final fixed points. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by these 19 probability distributions. The payoffs are concentrated in a very small region. The large red circle representing the payoff when starting from the uniform distribution is very close to this region, suggesting that it is a fairly reasonable representation of all possible 'types' of players' 'pre-experimental' preferences.

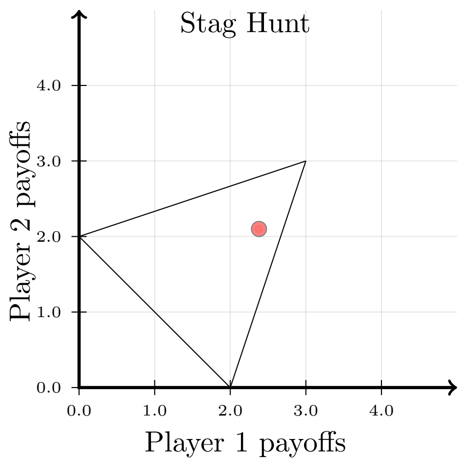
Figure A16: Stag Hunt - Full Information, Asymmetric Payoffs Multiple Equilibria



Dark Information, Symmetric Payoffs

The 100,000 starting probability distributions converge to a single final fixed point. The figure below presents the average payoff outcomes for both players, if they play strategies as prescribed by this single probability distribution. The large red circle representing the payoff when starting from the uniform distribution is the only payoff outcome, suggesting that it is a good representation of all possible 'types' of players' 'pre-experimental' preferences.

Figure A17: Stag Hunt - Dark Information, Symmetric Payoffs Multiple Equilibria

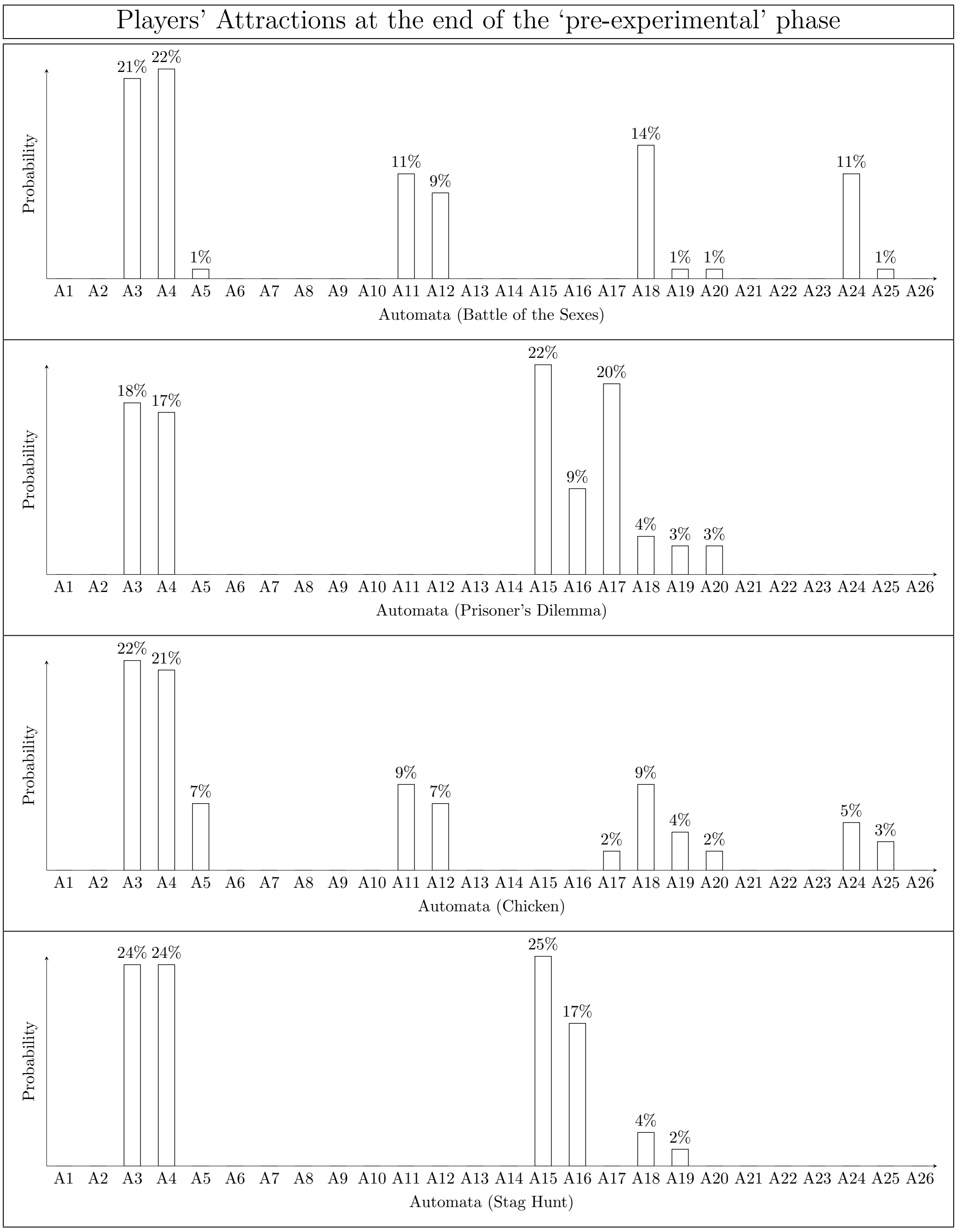


A5 Limiting Probability Distributions

Full information, asymmetric payoffs limiting distributions

Figure A18 shows the histogram of the probability distribution for the converged initial attractions for the row players across all four games in the full information, asymmetric payoffs treatment, when considering parameter values ($\alpha = 0.3, \beta = 0, \delta = 0.6, \lambda = 0.3, \gamma = 1.0$).

Figure A18: Converged initial attractions for the row players in each of the four games (full information, asymmetric payoffs treatment)



To simplify the discussion, I focus on strategies with limiting probability weights greater than or equal to 5%. These strategies are, in a sense, the principal strategies that players 'bring to the laboratory' for the 'experimental' phase.

Battle of the Sexes

Strategies A11 and A24 mechanically alternate between two states, but have different initial states. Strategy A12 or the 'co-ordinate once' strategy starts by playing action A and regardless of the opponent's action, switches to state B. It stays in this state if the opponent plays B; otherwise it returns to state A for one round before returning to state B. Strategy A18 or the 'punish once' strategy is qualitatively identical to strategy A12, with the role of the two actions reversed. Strategies A4 is the 'tit-for-tat' strategy, where players match the action taken by the opponent. Strategy A3 is a 'punish once' strategy, which starts by playing action A and remains in that state until the opponent plays B. At this point, the strategy plays B once before returning to play A irrespective of the opponents' action.

The highest weights on the 'tit-for-tat' and the 'punish once' strategies leads to very little coordination between subjects, and is likely to predict randomization between payoff outcomes.

Prisoner's Dilemma

Strategies obtaining high limiting probability weights are 'punish once,' 'tit-for-tat,' 'quick B, then A,' 'BA triggers A,' and 'BB triggers A', respectively. The 'quick B, then A' strategy starts at action B and and then mechanically moves to state A and stays there forever after. The 'BA triggers A' strategy starts at action B and stays at at that state as long as the opponent plays B. As soon as the opponent plays A once, the strategy dictates playing A forever after. The 'BB triggers A' strategy starts at action B and stays at at that state as long as the opponent plays A. As soon as the opponent plays B once, the strategy dictates playing A forever after.

The high weight on the 'tit-for-tat' strategy leads to a large number of mutual defection outcomes.

Chicken

Strategies A11 and A24 mechanically alternate between two states, but have different initial states. Strategy A12 or the 'co-ordinate once' strategy starts by playing action A and regardless of the opponent's action, switches to state B. It stays in this state if the opponent plays B; otherwise it returns to state A for one round before returning to state B. Strategy A18 or the 'punish once' strategy is qualitatively identical to strategy A12, with the role of the two actions reversed. Strategies A4 is the 'tit-for-tat' strategy, where players match the action taken by the opponent. Strategy A3 is a 'punish once' strategy, which starts by playing action A and remains in that state until the opponent plays B. At this point, the strategy plays B once before returning to play A irrespective of the opponents' action. The 'win-stay-lose-shift' strategy dictates that players will continue playing their current action if they receive a payoff higher than their opponent, and will shift to the other action if they receive a payoff lower than their opponent.

The highest weights on the 'tit-for-tat' and the 'punish once' strategies leads to very little coordination between subjects, and is likely to predict randomization between payoff outcomes.

Stag Hunt

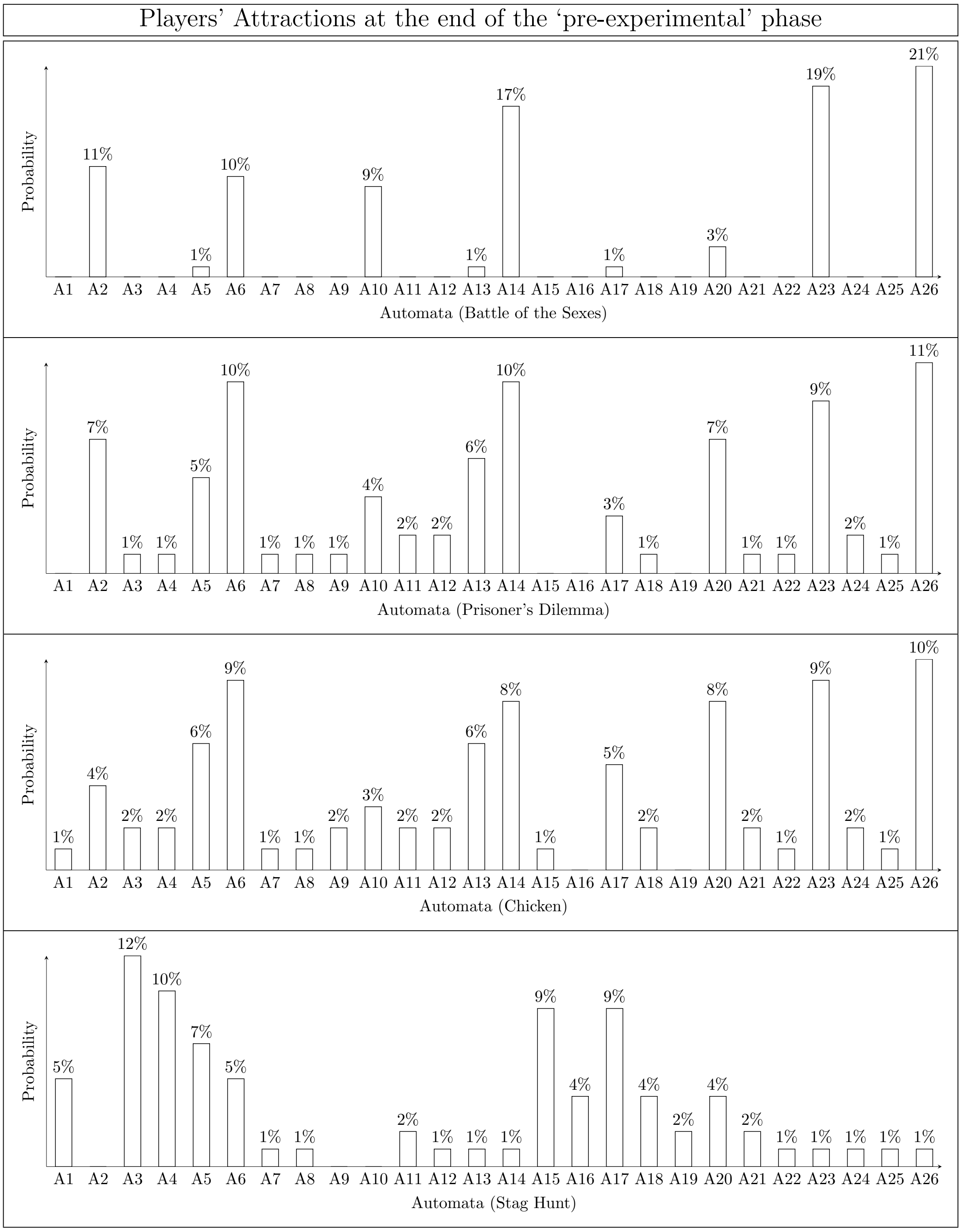
Strategies obtaining high limiting probability weights are 'punish once,' 'tit-for-tat,' 'quick B, then A,' and 'BA triggers A,' respectively. The 'BA triggers A' strategy starts at action B and stays at at that state as long as the opponent plays B. As soon as the opponent plays A once, the strategy dictates playing A forever after.

If these four strategies are played amongst themselves, the observed history of actions will likely involve mutual cooperation in all periods.

Dark information, symmetric payoffs limiting distributions

Figure A19 shows the histogram of the probability distribution for the converged initial attractions for the row players across all four games in the dark information, symmetric payoffs treatment, when considering parameter values ($\mu = 0.1, \lambda = 0.7, \gamma = 0.6$).

Figure A19: Converged initial attractions for the row players in each of the four games (dark information, symmetric payoffs treatment)



To simplify the discussion, I focus on strategies with limiting probability weights greater than or equal to 5%. These strategies are, in a sense, the principal strategies that players 'bring to the laboratory' for the 'experimental' phase.

Battle of the Sexes

Strategies obtaining high limiting probability weights are 'always play B,' 'grim trigger,' 'AA triggers B,' 'quick A, then B,' suspicious 'tit-for-tat,' and strategy A26, respectively. The 'always play B' strategy starts at action B and stays there forever. The 'grim trigger' strategy starts at action A and stays at at that state as long as the opponent plays A. As soon as the opponent plays B once, the strategy dictates playing B forever after. The 'AA triggers B' strategy starts at action A and stays at at that state as long as the opponent plays B. As soon as the opponent plays A once, the strategy dictates playing B forever after. The 'quick A, then B' strategy starts at action A and and then mechanically moves to state B and stays there forever after. The 'suspicious tit-for-tat' strategy starts at action B and has players match the action taken by the opponent. Strategy A26 starts at action B and stays at at that state as long as the opponent plays A. As soon as the opponent plays B once, the strategy dictates playing A once before mechanically returning to state B.

The highest weights on the suspicious 'tit-for-tat' and strategy A26 leads to players playing mis-coordinated actions, so players are likely to end up at the two pure Nash equilibria payoff outcomes.

Prisoner's Dilemma

Strategies obtaining high limiting probability weights are 'always play B,' 'win-stay-lose-shift,' 'grim trigger,' 'quick A, then B,' suspicious 'tit-for-tat,' suspicious 'win-stay-lose-shift,' strategy A13, and strategy A26. The 'win-stay-lose-shift' strategy dictates that players will continue playing their current action if they receive a payoff higher than their opponent, and will shift to the other action if they receive a payoff lower than their opponent. Strategy A13 starts at action A, mechanically moves to state B, and stays at at that state as long as the opponent plays A. As soon as the opponent plays B once, the strategy dictates playing A once before mechanically returning to state B.

The highest weights on the suspicious ‘tit-for-tat,’ ‘grim trigger,’ ‘quick A, then B,’ and strategy A26 leads to the mutual defection outcome.

Chicken

Strategies obtaining high limiting probability weights are ‘win-stay-lose-shift,’ ‘grim trigger,’ ‘quick A, then B,’ ‘BB triggers A,’ suspicious ‘win-stay-lose-shift,’ suspicious ‘tit-for-tat,’ strategy A13, and strategy A26. The ‘always play B’ strategy starts at action B and stays there forever. The ‘grim trigger’ strategy starts at action A and stays at at that state as long as the opponent plays A. As soon as the opponent plays B once, the strategy dictates playing B forever after. The ‘BB triggers A’ strategy starts at action B and stays at at that state as long as the opponent plays A. As soon as the opponent plays B once, the strategy dictates playing A forever after.

The highest weights on the suspicious ‘tit-for-tat’, and strategy A26 leads to players playing mis-coordinated actions, so players are likely to end up at the two pure Nash equilibria payoff outcomes.

Stag Hunt

Strategies obtaining high limiting probability weights are ‘always play A,’ ‘punish once,’ ‘tit-for-tat,’ ‘win-stay-lose-shift,’ ‘grim trigger,’ ‘quick B, then A,’ and ‘BB triggers A,’ . The ‘always play A’ strategy starts at action A and stays there forever. The ‘punish once’ strategy starts by playing action A and stays as the same state as long as the opponent plays A. If the opponent plays B, then the strategy dictates moving to state B for one round, before moving back to state A.

If these strategies are played amongst themselves, the observed history of actions will likely involve mutual cooperation in all periods.

A6 Experimental Instructions

Instructions

Welcome

You are about to participate in an experiment in the economics of decision-making. If you listen carefully, you could earn a large amount of money, that will be paid to you in cash, in private, at the end of the experiment.

It is important that you remain silent, and do not look at other people's work. If you have any questions, or need any assistance of any kind, please raise your hand and an experimenter will help you out. During the experiment, **do not talk, laugh or exclaim out loud** and be sure to keep your **eyes on your screen only**. In addition, please **turn off your cell phones, etc.** and put them away during the experiment. Anybody that violates these rules will be asked to leave and will **not** be paid. We expect, and appreciate your cooperation.

Agenda

1. Instructions.

2. Quiz (to make sure everyone understands the instructions). The quiz will consist of 12 questions. The experiment will not begin until everyone has correctly answered all questions.

3. Experiment.

4. Survey. After you have completed the experiment, you will be asked to answer several survey questions using the computer.

5. Payment. After everyone has finished the survey, you will be paid in cash.

Experiment Overview

In the experiment today you will be matched into pairs, and will be asked to make choices. The choices made by you may affect both your payoff, as well as the payoff of the subject that you are matched with. Similarly, the choices of the subject that you are matched with may affect their payoff and your payoff.

Experiment Details

- In the experiment today, you will be working with a fictitious currency called Francs. You will be paid in US Dollars at the end of the experiment.

Exchange rate: 1000 Francs = \$1.

- This experiment consists of **four matches**, each of which may have a different number of periods, and may have different payoffs.
- At the beginning of each match, you will be matched with one other subject that you have not been matched with before. The random pairing will ensure that if you have been paired with another participant before, then you will never be paired with this same participant in another match again.
- You will remain anonymous throughout the experiment. You will not know the identity of the subject that you are matched with, and the subject that you are matched with will not know your identity.
- The choices made by you, and the subject that you are matched with, have no effect on the payoffs of other pairs of subjects, and vice versa.
- You will remain matched with this same subject until the end of the match.
- After you and the subject that you are matched with have finished the match, please wait quietly for all pairs to finish their match.
- Once all pairs have finished their match, you will be randomly rematched with another subject that you have not been matched with yet, and the next match will begin.

Experimental Interface

- Next, we will go over the experimental interface. The experimental interface consists of five components:

  1. The payoff table is displayed at the top left of the screen. The payoff table has your choices (the rows, U, and D), and the choices of the subject that you are matched with (the columns, L, and R). For any combination of choices for you and the subject that you are matched with, there is a payoff. Your payoff is on the left, and the payoff for the subject that you are matched with is on the right. For example, if you choose D, and the subject that you are matched with chooses L, then your payoff is 6, and the payoff for the subject that you are matched with is 3. The payoff table in the experiment today will be different than the one displayed here. In addition, the payoffs will remain the same for all periods in a single match, but may change in different matches.
  2. The summary of the current period is shown in the top right.
  3. The summary of all periods is shown below that.
  4. The history of play is displayed in the bottom half of the screen.
  5. Finally, the current status is displayed across the middle of the screen. The current status tells you what to do at any point during the experiment.

Specific Instructions for Each Period

- At any point during a period, you will see one of three status messages:
- **Status #1:** Please select My Choice (U,D) and your guess for Other's Choice (L,R).
  - When you see this status message, you need to do two things.
    1. Select a row for your choice (in this example either U, or D). Once you have selected a row, it will be outlined, and the label "My Choice" will be added. My choice will also be updated on the period summary.

2. Select the column that you think the subject that you are matched with will select (in this example either L, or R). Once you have selected a column, it will be outlined, and the label "My Guess" will be added. My guess will also be updated on the period summary. If you correctly guess the choice of the subject that you are matched with in a given period, then you will earn a raffle ticket. At the end of each match, one raffle ticket will be randomly selected, and the winner will receive a bonus of $5.

- **Status #2:** Please wait for the other subject to finish making their choices.
  - When you see this status message, please wait patiently for the subject that you are matched with to finish making their choices.
  - If the other subject makes their choices before you in a specific period, you might not see this status.
- **Status #3:** Click on the payoffs for this period (in green) in the game table to move to next period.
  - When you see this status message, all choices for the period have been made, and you can view the final results.
  - The column selected by the subject that you are matched with is highlighted and labeled "Other's Choice". If your guess about their choice was correct, then the column will be highlighted in yellow.
  - The entry of the payoff table that is at the intersection of the row that you selected, and the column that the subject that you are matched with selected, is highlighted in green.
  - This entry displays the payoffs for the current period. Your payoff for the period is displayed on the left, and the payoff of the subject that you are matched with is displayed on the right.
  - At this point, the period summary has been updated to include the choice of the subject that you are matched with, and the current payoffs for this period. These choices have also been added to the history of play at the bottom of the screen.

- Finally, the overall summary has been updated to include the total payoffs including this period, as well as the number of correct guesses you have made about the choice of the subject that you are matched with.
- To move on to the next period, you need to click the payoffs for this period in the game table, which have been highlighted in green.

Number of Periods Per Match

- The number of periods in each match will be determined randomly using the following procedure.
  - At the end of each period, a number will be chosen randomly from the set of numbers $\{1, 2, 3, \dots, 98, 99, 100\}$, where **each number is equally likely**.
  - If the number is 1, then the match will end.
  - If the number is not 1, then the match will continue.
  - The number will always be placed back into the set after it is drawn.
  - Thus, in any period there is a 1% CHANCE that the match will end, and a 99% CHANCE that the match will have another period.
  - Therefore, the expected number of periods in each match will be 100.
  - This procedure has been performed on the computer before the experiment. Therefore, you will not see the number selected from $\{1, 2, 3, \dots, 98, 99, 100\}$.
  - To ensure that the length of the match is not dependent on your play, the number of periods for each match has been written on the board before the experiment, and will be uncovered at the end of the experiment.

Payoffs

- Your payment at the end of the experiment will contain the following:
  1. The $5 show up fee.
  2. Payment for all periods of all matches.
  3. A bonus payment of $5 if one of your raffle tickets is randomly selected at the end of the experiment.

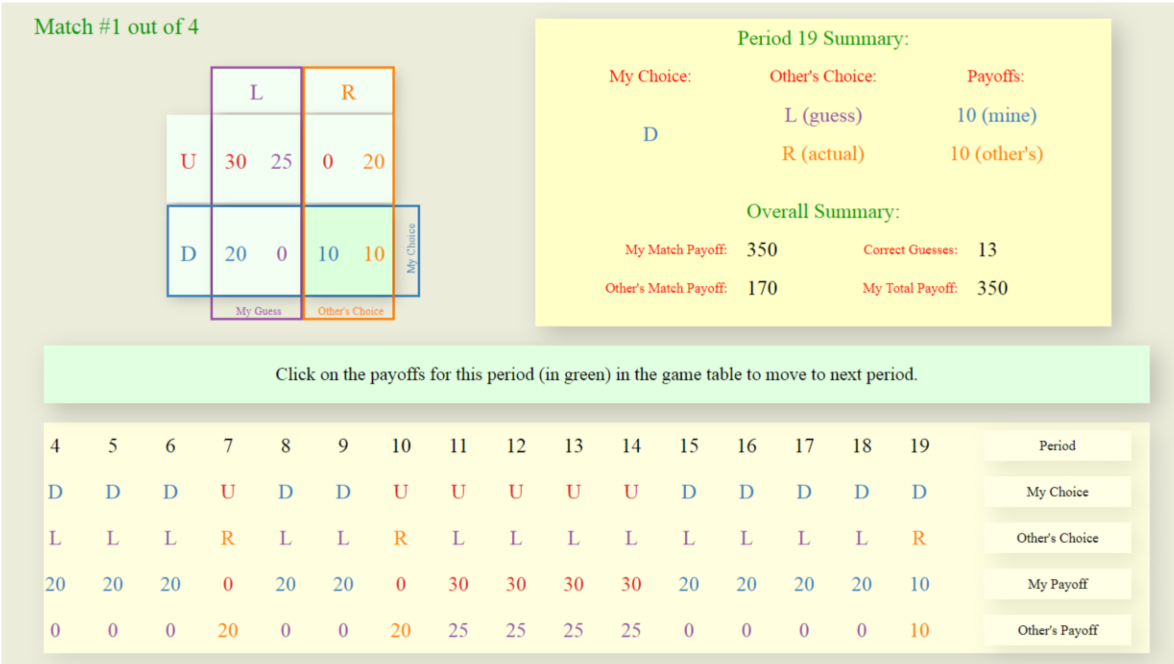
- At the end of the experiment, your earnings in Francs will be converted into US Dollars and you will be paid in cash, in private.

A7 Experimental Interface

The diagrams below show the experimental interface that subjects saw during the experiment.

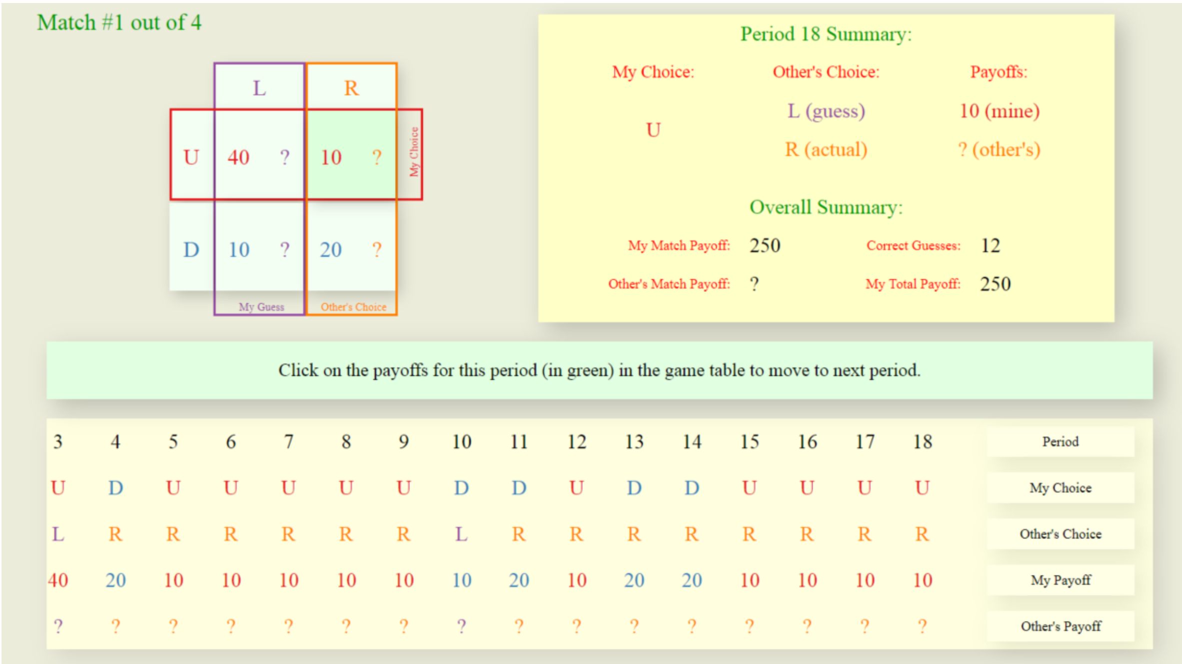
Full information, asymmetric payoffs experimental interface

Figure A20: Full Information, Asymmetric Payoffs Interface



Dark information, symmetric payoffs experimental interface

Figure A21: Dark Information, Symmetric Payoffs Interface

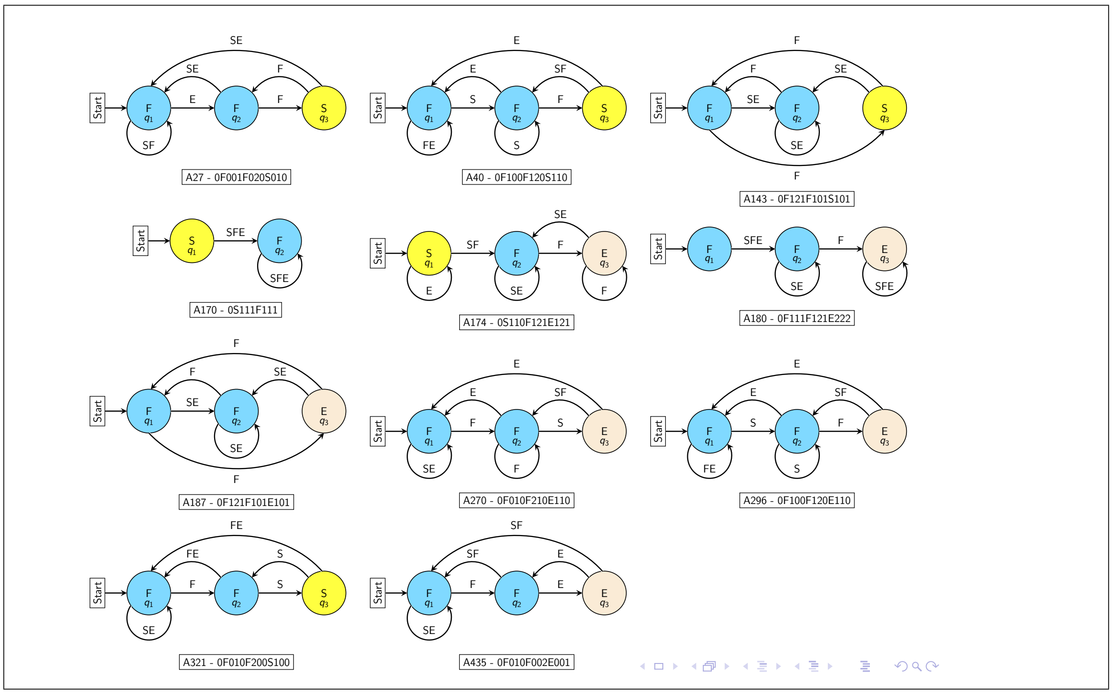


Chapter 2 Appendix

B1 Repeated Game Strategy Automata

The following diagram shows the 11 strategies represented as finite automata, that have a limiting probability distribution larger than 1% in the final fixed point for the parameter combination : $(\alpha = 0.3, \beta = 0.7, \delta = 0.0, \lambda = 1.0, \gamma = 0.8)$.

Figure B1: Three State Three Action Automata



B2 Experimental Instructions

Instructions

Welcome

You are about to participate in an experiment in the economics of decision-making. If you listen carefully, you could earn a large amount of money, that will be paid to you in cash, in private, at the end of the experiment.

It is important that you remain silent, and do not look at other people's work. If you have any questions, or need any assistance of any kind, please raise your hand and an experimenter will help you out. During the experiment, **do not talk, laugh or exclaim out loud** and be sure to keep your **eyes on your screen only**. In addition, please **turn off your cell phones, etc.** and put them away during the experiment. Anybody that violates these rules will be asked to leave and will **not** be paid. We expect, and appreciate your cooperation.

Agenda

1. Instructions.
2. Quiz (to make sure everyone understands the instructions). The quiz will consist of 12 questions. The experiment will not begin until everyone has correctly answered all questions.
3. Experiment.
4. Survey. After you have completed the experiment, you will be asked to answer several survey questions using the computer.
5. Payment. After everyone has finished the survey, you will be paid in cash.

Experiment Overview

In the experiment today you will be matched into pairs, and will be asked to make choices. The choices made by you may affect both your payoff, as well as the payoff of the subject that you are matched with. Similarly, the choices of the subject that you are matched with may affect their payoff and your payoff.

Experiment Details

- In the experiment today, you will be working with a fictitious currency called Francs. You will be paid in US Dollars at the end of the experiment.

  Exchange rate: 2000 Francs = \$1.

- This experiment consists of **ten matches**, each of which may have a different number of periods, but may or may not have the same payoffs.
- At the beginning of each match, you will be matched with one other subject that you have not been matched with before. The random pairing will ensure that if you have been paired with another participant before, then you will never be paired with this same participant in another match again.
- You will remain anonymous throughout the experiment. You will not know the identity of the subject that you are matched with, and the subject that you are matched with will not know your identity.
- The choices made by you, and the subject that you are matched with, have no effect on the payoffs of other pairs of subjects, and vice versa.
- You will remain matched with this same subject until the end of the match.
- After you and the subject that you are matched with have finished the match, please wait quietly for all pairs to finish their match.
- Once all pairs have finished their match, you will be randomly rematched with another subject that you have not been matched with yet, and the next match will begin.

Experimental Interface

- Next, we will go over the experimental interface. The experimental interface consists of five components:

  1. The payoff table is displayed at the top left of the screen. The payoff table has your choices (the rows, U, M, and D), and the choices of the subject that you are matched with (the columns, L, C, and R). For any combination of choices for you and the subject that you are matched with, there is a payoff. Your payoff is on the left, and the payoff for the subject that you are matched with is on the right. For example, if you choose D, and the subject that you are matched with chooses L, then your payoff is 6, and the payoff for the subject that you are matched with is 7. The payoff table in the experiment today will be different than the one displayed here. In addition, the payoffs will remain the same for all periods in a single match, but may change in different matches.
  2. The summary of the current period is shown in the top right.
  3. The summary of all periods is shown below that.
  4. The history of play is displayed in the bottom half of the screen.
  5. Finally, the current status is displayed across the middle of the screen. The current status tells you what to do at any point during the experiment.

Specific Instructions for Each Period

- At any point during a period, you will see one of three status messages:

- **Status #1:** Please select My Choice (U,M,D) and your guess for Other's Choice (L,C,R).

  – When you see this status message, you need to do two things.

    1. Select a row for your choice (in this example either U, M, or D). Once you have selected a row, it will be outlined, and the label "My Choice" will be added. My choice will also be updated on the period summary.

2. Select the column that you think the subject that you are matched with will select (in this example either L, C, or R). Once you have selected a column, it will be outlined, and the label “My Guess” will be added. My guess will also be updated on the period summary. If you correctly guess the choice of the subject that you are matched with in a given period, then you will earn a raffle ticket. At the end of the experiment, one raffle ticket will be randomly selected, and the winner will receive a bonus of $5. Therefore, the more correct guesses you make, the more likely you are to win the $5 bonus.

- **Status #2:** Please wait for the other subject to finish making their choices.
  - When you see this status message, please wait patiently for the subject that you are matched with to finish making their choices.
  - If the other subject makes their choices before you in a specific period, you might not see this status.
- **Status #3:** Click on the payoffs for this period (in green) in the game table to move to next period.
  - When you see this status message, all choices for the period have been made, and you can view the final results.
  - The column selected by the subject that you are matched with is highlighted and labeled “Other’s Choice”. If your guess about their choice was correct, then the column will be highlighted in yellow.
  - The entry of the payoff table that is at the intersection of the row that you selected, and the column that the subject that you are matched with selected, is highlighted in green.
  - This entry displays the payoffs for the current period. Your payoff for the period is displayed on the left, and the payoff of the subject that you are matched with is displayed on the right.
  - At this point, the period summary has been updated to include the choice of the subject that you are matched with, and the current payoffs for this

period. These choices have also been added to the history of play at the bottom of the screen.

  - Finally, the overall summary has been updated to include the total payoffs including this period, as well as the number of correct guesses you have made about the choice of the subject that you are matched with.
  - To move on to the next period, you need to click the payoffs for this period in the game table, which have been highlighted in green.

Number of Periods Per Match

- The number of periods in each match will be determined randomly using the following procedure.

  - At the end of each period, a number will be chosen randomly from the set of numbers $\{1, 2, 3, \ldots, 48, 49, 50\}$, where **each number is equally likely**.
  - If the number is 1, then the match will end.
  - If the number is not 1, then the match will continue.
  - The number will always be placed back into the set after it is drawn.
  - Thus, in any period there is a 2% CHANCE that the match will end, and a 98% CHANCE that the match will have another period.
  - Therefore, the expected number of periods in each match will be 50.
  - This procedure has been performed on the computer before the experiment. Therefore, you will not see the number selected from $\{1, 2, 3, \ldots, 48, 49, 50\}$.
  - To ensure that the length of the match is not dependent on your play, the number of periods for each match has been written on the board before the experiment, and will be uncovered at the end of the experiment.

Payoffs

- Your payment at the end of the experiment will contain the following:

  1. The \$5 show up fee.

2. Payment for all periods of all matches.
3. A bonus payment of \$5 if one of your raffle tickets is randomly selected at the end of the experiment.

- At the end of the experiment, your earnings in Francs will be converted into US Dollars and you will be paid in cash, in private.

B3 Theoretical Proofs

I present here a proof that guarantees the existence of a fixed point for propensities towards actions, given any starting point of initial propensities. While the games studied in the paper have two players and two actions, the proof technique is generalised for any finite number of players, with each having a finite number of actions. The proof is similar in spirit to the one used to show the existence of a Quantal Response Equilibrium in McKelvey and Palfrey (1995).

Let the game be played by n players, and let each player $i \in N = \{1, 2, ..., n\}$ have J_i strategies. Initial propensities towards each action for player i are given by $p_i = (p_{i,1}, p_{i,2}, ..., p_{i,J_i})$ such that $\sum_{k=1}^{J_i} p_{i,k} = 1$. Every element $p_i \in \Delta_i$ which is a simplex of size J_i. Taken across all players, I have $p = (p_1, p_2, ..., p_i, ..., p_n) \in \Delta$ where $\Delta = \Pi_{i \in N} \Delta_i$.

Let $s_{i,j} \in S_i$ denote strategy j for player i, an element of all strategies available to player i. The payoff function maps from the strategy space into $\mathbb{R}^+$. The algorithm defined in Section 2.3 and Section 2.4 then maps from the payoff space into a composite score $H = \alpha E + \beta AF + (1 - \alpha - \beta)S$. Thus, I have a function that maps from the strategy space into the composite score space, which is $\mathbb{R}^+$. For each player i, I have:

$g_i(s_i, s_{-i}) \rightarrow \mathbb{R}^+$ where $s_i \in S_i, s_{-i} \in S_{-i}$ and $S_i \times S_{-i} = S$

I can extend the domain of g_i to Δ which allows probability distributions over strategies by using the following rule:

$g_i : S \times \Delta \rightarrow \mathbb{R}^+$ where $g_i(s_i, s_{-i}; p) = \sum_{s \in S} p(s) g_i(s)$ where $p(s) = \Pi_{i \in N} p_i(s_i)$. Given the probability distribution over strategies being used by other players, player i's scaled expected composite score from using strategy j is given by $\bar{g}_{i,j}(p) = exp(\lambda g_i(s_{i,j}, p_{-i}))$. Finally, I have a function that takes a set of probabilities for the n players, and returns an updated set of probabilities using the following algorithm.

$l(p_{i,j}) = \frac{\bar{g}_{i,j}(p)}{\sum_{k=1}^{J_i} \bar{g}_{i,k}(p)}$ for $p \in \Delta$

Thus, $l(.)$ is a function that maps Δ into Δ. Since $g_i(.)$ is continuous in p, $\bar{g}_{i,j}(.)$ is also continuous in p. Sums of continuous functions are continuous, and $\frac{f(x)}{g(x)}$ is continuous if $f(x)$ and $g(x)$ are continuous in x, and $g(x) \neq 0$. Given these properties, $l(.)$ is a continuous function in p, and maps Δ(a nonempty, compact, and convex subset of $\mathbb{R}^n$) into itself. By Brouwer's fixed point, the existence of a fixed point is guaranteed.

Chapter 3 Appendix

C1 Theoretical Proofs

Normal distribution case for expected reference point

Considering an uniformative uniform prior distribution of μ could potentially lead the Dictator to place too much weight on the datapoint received x, and too little weight on the prior distribution. We now consider a more informative prior distribution for μ and find similar results.

$\theta \sim N(\mu, \sigma^2)$, $\mu \sim N(\mu_0, \tau^2)$

$f_\theta(\tilde{\theta}|\mu) = \frac{1}{\sqrt{2\Pi\sigma^2}} exp\left[\frac{-1}{2\sigma^2}\left(\tilde{\theta} - \mu\right)^2\right]$, $g(\mu) = \frac{1}{\sqrt{2\Pi\tau^2}} exp\left[\frac{-1}{2\tau^2}\left(\mu - \mu_0\right)^2\right]$

$s(\mu|\theta = x) = \frac{f_\theta(x|\mu)g(\mu)}{\int f_\theta(x|\mu)g(\mu)du}$

$s(\mu|\theta = x) \propto \frac{1}{\sigma} exp(\frac{-1}{2\sigma^2}(x - \mu)^2 \frac{1}{\tau} exp(\frac{-1}{2\tau^2}(\mu - \mu_0)^2$

$s(\mu|\theta = x) \propto exp\left[\frac{-1}{2}\left(\left(\frac{x^2}{\sigma^2} - \frac{2x\mu}{\sigma^2} + \frac{\mu^2}{\sigma^2} - \frac{2\mu\mu_0}{\tau^2} + \frac{\mu_0^2}{\tau^2}\right)\right)\right]$

$s(\mu|\theta = x) \propto exp\left[\frac{-1}{2}\left(\left(\mu^2 \frac{\sigma^2+\tau^2}{\tau^2\sigma^2} - \mu\frac{2x\tau^2+2\mu_0\sigma^2}{\tau^2\sigma^2}\right)\right)\right]$

$s(\mu|\theta = x) \propto exp\left[\frac{-1}{2(\frac{1}{\frac{1}{\tau^2}+\frac{1}{\sigma^2}})}\left(\left(\mu - \frac{x}{\sigma^2} - \frac{\frac{\mu_0}{\tau^2}}{\frac{1}{\tau^2}+\frac{1}{\sigma^2}}\right)^2\right)\right]$

Thus, the posterior distribution for μ after the Bayesian updating process is as follows:

$\mu \sim N(\frac{x}{\sigma^2} + \frac{\frac{\mu_0}{\tau^2}}{\frac{1}{\tau^2}+\frac{1}{\sigma^2}}, \frac{1}{\frac{1}{\tau^2}+\frac{1}{\sigma^2}})$

The posterior distribution has an expected mean that is directly proportional to the signal received x. Again, the theoretical prediction is that the amount allocated to the receiver is increasing in the realization of the signal x.

Gain-loss utility preferences

We further explore a modified version of the reference-dependent utility framework explored in Koszegi and Rabin (2006). In the original KR model, reference points are both stochastic and endogenously determined. In our model, we will consider stochastically determined reference points, but the reference points are not endogenous other than being a result of Bayesian updating. The utility form in Koszegi and Rabin (2006) is as follows:

$u(c|r) = m(c) + n(c|r)$ where $m(c)$ represents consumption utility, as covered in the earlier section. The second term $n(c|r)$ represents the gain-loss utility felt by receiving an allocation of c when the expectation was to receive an allocation of r, which represents the reference point. Thus, gain or loss is felt relative to the expected consumption point. The KR model also has the following form assumption on $n(c|r)$:

$n(c|r) = \psi(m(c) - m(r))$ where $\psi(.)$ follows a Kahneman and Tversky (1979) loss-utility form. The KT loss-utility form ensures that players feel more disutility from receiving allocations below the reference point, relative to the positive utility felt when receiving allocations above their reference point. Formally, we have $\psi(x) = \eta x$ for $x > 0$, and $\psi(x) = \eta\lambda x$ for $x < 0$. $\eta > 0$ represents the weight attached by the Dictator to the gain-loss utility, and $\lambda > 1$ represents the Dictator's coefficient of loss-aversion.

The original form of the Kahneman and Tversky loss-aversion expected utility model is as follows:

$U(F|r) = \int u(c|r)dF(c)$ where $F(c)$ represents the probability distribution of all the possible allocations c. $F(c)$ is the Dictator's belief over all the possible allocations that she could potentially end up with. Instead of leaving the reference point as a hanging parameter, or assuming the reference point to be equal to the status-quo point, the Koszegi and Rabin model argues that the reference point is based on expectations. Thus, given that the Dictator has a probability distribution $G(r)$ over the possible values of the reference point, and the overall utility is given by:

$U(F|G) = \iint u(c|r)dG(r)dF(c)$

The interpretation of this utility form is that players have an expectation about each outcome becoming the reference point, and then evaluate the gain loss of each other allocation relative to the reference point, multiplied by the likelihood of each allocation. Players choose an allocation that maximizes this final utility form. We note that in this form, the utility is additively separable over the consumption utility and the gain-loss utility.

In the ensuing analysis, we assume that the probability $G(r)$ over the possible values of the reference point corresponds to the distribution of the $c^* = (c_r^*, c_d^*)$ that maximizes the consumption utility covered in Proposition 1. Given this probability distribution over the possible reference points, we consider how the Dictator evaluates all other potential splits against the possible reference points. Consider a particular reference point $r = c^* = (c_r, c_d)$. For the Dictator to consider increasing the share of the receiver by Δt, it must be true that the Dictator receives a higher overall utility from the consumption level $c = (c_r + \Delta t, c_d - \Delta t)$ than from staying at the reference point. This is only possible if the Dictator expects that $\theta > \frac{(1+\eta\lambda)[f(c_d)-f(c_d-\Delta t)]}{(1+\eta)[f(c_r+\Delta t)-f(c_r)]}$.[8] Similarly, for the Dictator to consider decreasing the receiver's share by Δt, it must be true that the Dictator receives a higher overall utility from the consumption level $c = (c_r - \Delta t, c_d + \Delta t)$ than from staying at the reference point. This is only possible if the Dictator expects that $\theta < \frac{(1+\eta)[f(c_d+\Delta t)-f(c_d)]}{(1+\eta\lambda)[f(c_r)-f(c_r-\Delta t)]}$. Note that, as the signal x increases (as it does across subject pairs in the experiment), the expected distribution of θ shifts to the right, so $Pr(\theta > \frac{(1+\eta\lambda)[f(c_d)-f(c_d-\Delta t)]}{(1+\eta)[f(c_r+\Delta t)-f(c_r)]})$ is increasing as the signal x increases, and $Pr(\theta < \frac{(1+\eta)[f(c_d+\Delta t)-f(c_d)]}{(1+\eta\lambda)[f(c_r)-f(c_r-\Delta t)]})$ is decreasing as the signal x increases. Note that this is true for every possible value of c^*across the probability distribution $G(r)$ derived from the optimal consumption levels. Thus, when maximizing overall consumption and gain-loss utility, the Dictator will choose higher amounts of c_r as the signal $\hat{\theta}$ increases. We label this condition Proposition 2.

Proposition 2 : As the percentage of earned ECUs in the Receiver's endowment increases (or as the percentage of earned ECUs in the Dictator's endowment decreases),

[8]Calculations presented below

the distribution for the optimal choice of c_r^* *that maximizes the overall consumption and gain-loss utility shifts to the right, and Dictators are likely to give larger shares to receivers.*

Gain-loss utility calculations

For the Dictator to consider decreasing the share of the receiver by Δt, it must be true that the Dictator receives a higher overall utility from the consumption level $c = (c_r - \Delta t, c_d + \Delta t)$ than from the reference consumption level of $r = (c_r, c_d)$. Thus, we must have $U(c|r) > U(r|r)$.

$U(r|r) = m(r) = f(c_d) + \theta * f(c_r)$

$U(c|r) = m(c) + \psi(c|r) = m(c) + \psi(m(c) - m(r)) = f(c_d + \Delta t) + \theta * f(c_r - \Delta t) + \psi(f(c_d + \Delta t) + \theta * f(c_r - \Delta t) - f(c_d) - \theta * f(c_r))$

$U(c|r) > U(r|r) \Rightarrow \eta(f(c_d + \Delta t) - f(c_d)) - \theta * \eta\lambda(f(c_r - \Delta t) - f(c_r)) + f(c_d + \Delta t) + \theta * f(c_r - \Delta t) > f(c_d) + \theta * f(c_r)$

$\theta < \frac{(1+\eta)[f(c_d+\Delta t)-f(c_d)]}{(1+\eta\lambda)[f(c_r)-f(c_r-\Delta t)]}$

There is a similar calculation for the case where the Dictator considers the case for the consumption level $c = (c_r + \Delta t, c_d - \Delta t)$.

C2 Experimental Instructions

Welcome

You are about to participate in an experiment in the economics of decision-making. If you play carefully, you could earn a large amount of money, that will be paid to you in cash, in private, at the end of the experiment. It is important that you remain silent, and do not look at other people's work. If you have any questions, or need any assistance of any kind, please raise your hand and an experimenter will help you out. During the experiment, **do not talk, laugh or exclaim out loud** and be sure to keep **your eyes on your screen only**. In addition, please **turn off your cell phones, etc**, and put them away during the experiment. Anybody that violates these rules will be asked to leave and will **not** be paid. We expect, and appreciate your cooperation.

Agenda

1. Instructions
2. Quiz (to make sure everyone understands the instructions). The quiz will consist of 6 questions. The experiment will not begin until everyone has correctly answered all questions.
3. Experiment
4. Survey. After you have completed the experiment, you will be asked to answer several survey questions using the computer.
5. Payment. After everyone has finished the survey, you will be paid in cash.

Experiment

In the experiment today, you will be working with a fictitious currency called Experimental Currency Units (ECUs). You will be paid in US Dollars at the end of the experiment.

Exchange rate: 20 ECUs = $1.

The experiment consists of two stages.

Stage One Overview

In Stage One, you will undertake two types of tasks.

Task One is to count the correct number of zeros in a table. You will have 2 minutes to answer as many tables as you can correctly. You will receive 3 ECUs for every table that you answer correctly. You will have three attempts to answer a table correctly, before a new table is generated. The figure on the next page shows the work screen you will use later.

Task Two is to move a slider to its middle point, till the number next to slider reads 50. You will have 2 minutes to slide as many sliders as you can correctly. You will receive 1 ECU for every slider that you move correctly. The number next to each slider shows the current position of a slider, and you will receive 1 ECU for each slider where this number shows 50. The figure on the next page shows the work screen you will use later.

Figure C1: Task One

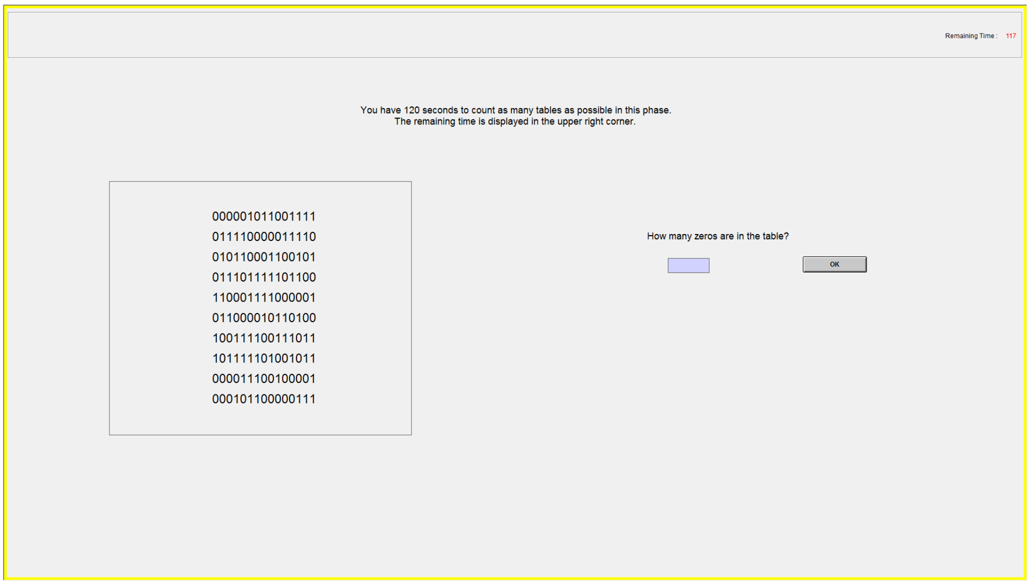
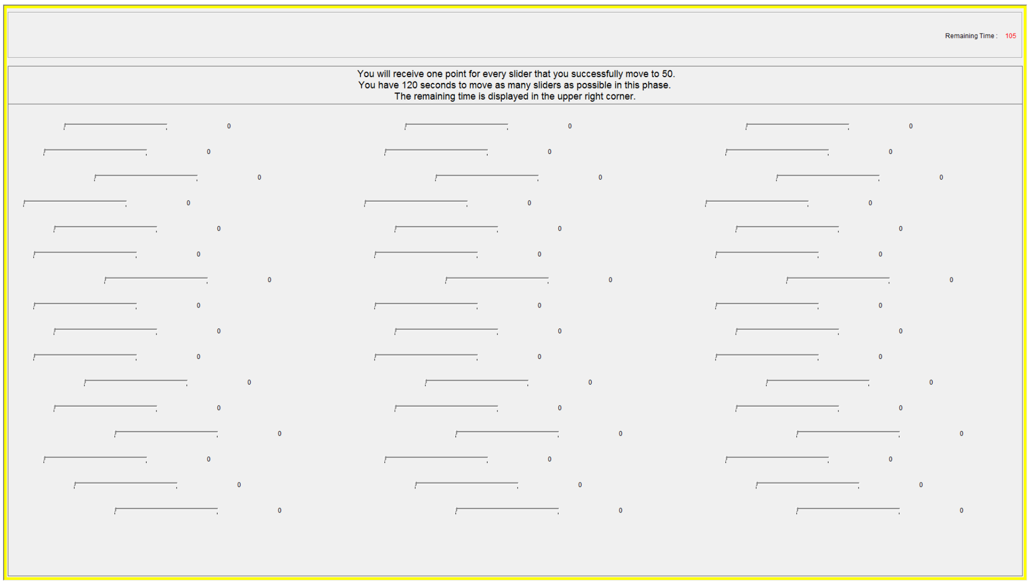


Figure C2: Task Two



You will alternate between Task One and Task Two for a total of 20 minutes. Thus, you will perform Task One five times, and Task Two five times. The total number of ECUs that you earn in the two tasks will reflect your total Stage One score.

Stage Two Overview

At the end of Stage One, you will be randomly matched with another player.

Between you and your partner, the subject with the higher Stage One score will be assigned to Group A, while the subject with the lower Stage One score will be assigned to Group B.

At this point, both you and your partner will also receive a Stage One Bonus.

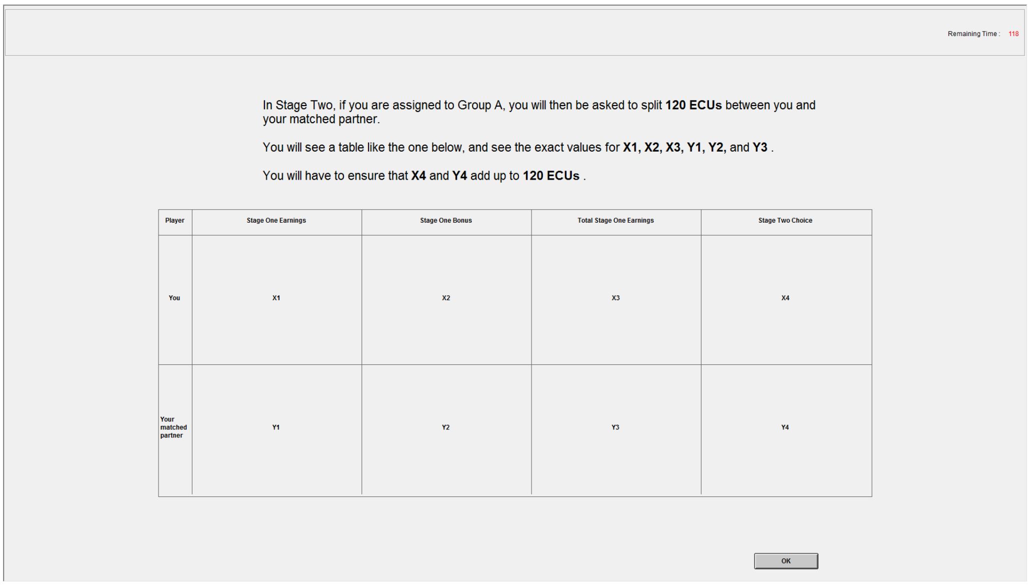
The Stage One Bonus has a completely random component, and a deterministic component that depends directly on whether you are in Group A or Group B.

Stage Two - Group A Subjects

In Stage Two, if you are assigned to Group A, you will be asked to split **120 ECUs** between you and your matched partner.

You will see a table like the one below, and see the exact values for **X1, X2, X3, Y1, Y2** and **Y3**. You will have to ensure that you choose **X4** and **Y4** in a way that they add up to **120**.

Figure C3: Group A Screen



Your matched partner will not know the **breakdown** of their earnings, but will only learn their **total** earnings after Stage Two.

You will also answer questions about what you believe about your matched partner's expectations of their Stage Two earnings.

At the end of the second stage, you will be notified of the total number of ECUs you have received during the experiment.

After completing a survey questionnaire, you will receive your payment, and after that, you can leave.

Stage Two - Group B Subjects

In Stage Two, if you are assigned to Group B, you will no longer be an active participant.

You will make no choices that can affect your matched partner's payoff.

The only chance you will have to affect your own payoffs is by answering questions about what actions you believe your matched partner will take in Stage Two.

In addition, you will not know the **breakdown** of your earnings, but will only learn your **total** earnings after Stage Two.

At the end of the second stage, you will be notified of the total number of ECUs you have received during the experiment.

After completing a survey questionnaire, you will receive your payment, and after that, you can leave.

Payoffs

Your payoff at the end of the experiment will be the sum of the ECUs in Stage One and Stage Two.

In addition, you will receive a fixed show-up fee of $5 for this experiment.

At the end of the experiment, your earnings in ECUs will be converted into US Dollars and you will be paid in cash, in private.

References

Abeler, J., A. Falk, L. Gotte, and D. Huffman (2011). Reference points and effort provision. The American Economic Review 101 (2), 470–492.

Akbas Merve, Dan Ariely, and Sevgi Yuksel. "When is inequality fair? An experiment on the effect of procedural justice and agency." (2016).

Alesina, A. and Angeletos, G.M. (2005). Fairness and redistribution: US versus Europe, American Economic Review, vol. 95(4), pp. 960–80.

Andreoni, J. (1989). Giving with impure altruism: applications to charity and Ricardian equivalence, Journal of Political Economy, vol. 97(6), pp. 1447–58.

Andreoni, J. and Bernheim, B. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. Econometrica, 77(5):1607–1636.

Aoyagi, Masaki, and Guillaume Frechette. "Collusion as public monitoring becomes noisy: Experimental evidence." Journal of Economic Theory 144.3 (2009): 1135-1165.

Arifovic, Jasmina, John Ledyard, and Lilong Shi. Repeated Battle of Sexes: Experimental evidence and individual evolutionary learning. HSS Working Paper, Caltech, 2015.

Axelrod, R., 1984, The evolution of cooperation, New York, Basic Books.

Balafoutas, L., M. G. Kocher, L. Putterman, and M. Sutter (2013). Equality, equity and incentives: An experiment. European Economic Review 60, 32-51.

Baron, James N., and David M. Kreps (1999). "Consistent human resource practices." California Management Review 41.3: 29-48.

Battigalli, Pierpaolo, and Martin Dufwenberg (2007). "Guilt in games." The American Economic Review 97.2: 170-176.

Battigalli, Pierpaolo, and Martin Dufwenberg (2009). "Dynamic psychological games. " Journal of Economic Theory 144.1: 1-35.

Becker, Gary S (1974). "A theory of social interactions." Journal of Political Economy 82.6: 1063-1093.

Benabou, R. and J. Tirole (2006). "Belief in a just world and redistributive politics." The Quarterly Journal of Economics 121(2): 699-746.

Bhaskar, V. "Egalitarianism and efficiency in repeated symmetric games." Games and Economic Behavior 32.2 (2000): 247-262.

Binmore, Ken. "Game theory and institutions." Journal of Comparative Economics 38.3 (2010): 245-252.

Bo, Pedro Dal, and Guillaume R. Frechette. "The evolution of cooperation in infinitely repeated games: Experimental evidence." The American Economic Review 101.1 (2011): 411-429.

Bolton, Gary, and Axel Ockenfels. (2000). “ERC: A Theory of Equity, Reciprocity, and Competition,” American Economic Review, 90(1): 166-93.

Bose, Devdeepta., Social Preferences in Infinitely Repeated Games. University of Arizona Working Paper, Tucson, 2018.

Brown, George W. "Iterative solution of games by fictitious play." Activity analysis of production and allocation 13.1 (1951): 374-376.

Camerer, Colin. 2003. Behavioral game theory: Experiments in strategic interaction. Princeton University Press.

Camerer, Colin, and Teck-Hua Ho. "Experience-weighted attraction learning in coordination games: Probability rules, heterogeneity, and time-variation." Journal of Mathematical Psychology 42.2 (1998): 305-326.

Camerer, Colin, and Teck Hua Ho. "Experience-weighted attraction learning in normal form games." Econometrica 67.4 (1999): 827-874.

Cappelen, A.W., Hole, A.D., Sorensen, E.O., & Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. American Economics Review, 97, 818-827.

Charness, Gary, and Matthew Rabin. "Understanding social preferences with simple tests." Quarterly Journal of Economics (2002): 817-869.

Charness, Gary, and Martin Dufwenberg. "Promises and partnership." Econometrica 74.6 (2006): 1579-1601.

Chen, Yan, and Fang-Fang Tang. "Learning and incentive-Compatible Mechanisms for Public Goods Provision: An Experimental Study." Journal of Political Economy 106.3 (1998): 633-662.

Cherry, Todd L., Peter Frykblom, and Jason F. Shogren (2002). "Hardnose the dictator." The American Economic Review 92.4: 1218-1221.

Cherry, Todd L. "Mental accounting and other-regarding behavior: Evidence from the lab." Journal of Economic Psychology 22.5 (2001): 605-615.

Cheung, Yin-Wong, and Daniel Friedman. "Individual learning in normal form games: Some laboratory results." Games and Economic Behavior 19.1 (1997): 46-76.

Cournot, Antoine-Augustin. Recherches sur les principes mathematiques de la theorie des richesses par Augustin Cournot. chez L. Hachette, 1838.

Cox, James C., Daniel Friedman, and Steven Gjerstad (2007). "A tractable model of reciprocity and fairness." Games and Economic Behavior 59.1: 17-45.

Dal Bo, Pedro, and Guillaume R. Frechette. "Strategy choice in the indefinitely repeated prisoners' dilemma." Available at SSRN 2292390 (2015).

Dreber, Anna, Drew Fudenberg, and David G. Rand. "Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics." Journal of Economic Behavior and Organization 98 (2014): 41-55.

Dreber, Anna, et al. "Winners don't punish." Nature 452.7185 (2008): 348-351.

Duffy, John, and Felix Munoz-Garcia. "Patience or fairness? Analyzing social preferences in repeated games." Games 3.1 (2012): 56-77.

Dufwenberg, Martin, and Georg Kirchsteiger. "A theory of sequential reciprocity." Games and economic behavior 47.2 (2004): 268-298.

Dufwenberg, Martin, Simon Gachter, and Heike Hennig-Schmidt (2011). "The framing of games and the psychology of play." Games and Economic Behavior 73.2: 459-478.

Durante, R., L. Putterman, and J. Weele (2014). Preferences for redistribution and perception of fairness: An experimental study. Journal of the European Economic Association 12 (4), 1059-1086.

Erev, Ido, and Alvin E. Roth. "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria." American Economic Review (1998): 848-881.

Erkal, N., Gangadharan, L., & Nikiforakis, N. (2011). Relative earnings and giving in a real-effort experiment. American Economic Review, 101, 3330-3348.

Falk, Armin, and Urs Fischbacher. "A theory of reciprocity." Games and economic behavior 54.2 (2006): 293-315.

Falk, A., Fehr, E., Fischbacher, U., (2003). On the nature of fair behavior. Economic Inquiry 41 (1), 20–26.

Fehr, Ernst, and Simon Gachter (2000). "Fairness and retaliation: The economics of reciprocity." The Journal of Economic Perspectives 14.3: 159-181.

Fehr, Ernst, and Klaus M. Schmidt. "A theory of fairness, competition, and cooperation." Quarterly Journal of Economics (1999): 817-868.

Fischbacher, Urs. "z-Tree: Zurich toolbox for ready-made economic experiments." Experimental Economics 10.2 (2007): 171-178.

Fong, C. M. (2001) “Social preferences, selfinterest, and the demand for redistribution”, Journal of Public Economics, vol. 82, pp. 225–46.

Fong, C.M., & Oberholzer-Gee, F. (2011). Truth in giving: Experimental evidence on the welfare effects of informed giving to the poor. Journal of Public Economics, 95, 436-444.

Frechette, Guillaume R., and Sevgi Yuksel. "Infinitely repeated games in the laboratory: Four perspectives on discounting and random termination." Experimental Economics 20.2 (2017): 279-308.

Frey, Brendan J., and Delbert Dueck. "Clustering by passing messages between data points." Science 315.5814 (2007): 972-976.

Fudenberg, Drew, David G. Rand, and Anna Dreber. "Slow to anger and fast to forgive: Cooperation in an uncertain world." The American Economic Review 102.2 (2012): 720-749.

Geanakoplos, John, David Pearce, and Ennio Stacchetti (1989). "Psychological games and sequential rationality." Games and Economic Behavior 1.1: 60-79.

Gee, Laura K., Marco Migueis, and Sahar Parsa (2017). "Redistributive choices and increasing income inequality: experimental evidence for income as a signal of deservingness." Experimental Economics 20.4: 894-923.

Gill, David, and Victoria L. Prowse (2011). "A novel computerized real effort task based on sliders." Available at SSRN 1732324.

Haile, Philip A., Ali Hortacsu, and Grigory Kosenok. "On the empirical content of quantal response equilibrium." The American Economic Review 98.1 (2008): 180-200.

Hanaki, N., Sethi, R., Erev, I., and Peterhansl, A. (2005). Learning strategies. Journal of Economic Behavior & Organization, 56(4), 523-542.

Hoffman, E., McCabe, K., Smith, V.L., (1994). Preferences, property rights, and anonymity in bargaining games. Games and Economic Behavior 7, 346–380.

Hoffman, E., Spitzer, M., (1985). Entitlements, rights, and fairness: an experimental examination of subjects' concepts of distributive justice. Journal of Legal Studies 14, 259–297.

Ioannou, Christos A., and Julian Romero. "A generalized approach to belief learning in repeated games." Games and Economic Behavior 87 (2014): 178-203.

Kahneman, Daniel and Amos Tversky, (1979), Prospect theory, an analysis of decision under risk, Econometrica 47-69.

Koszegi, Botond, and Matthew Rabin (2006). "A Model of Reference-Dependent Preferences" Quarterly Journal of Economics, 121(4): 1133-1166.

Krawczyk, M. (2010). A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. Journal of Public Economics 94 (1), 131-141.

Ku, H. and T. C. Salmon (2013). Procedural fairness and the tolerance for income inequality. European Economic Review 64, 111-128.

Kuhn, Harold W. "The Hungarian method for the assignment problem." Naval research logistics quarterly 2.1-2 (1955): 83-97.

Kuzmics, Christoph, Thomas Palfrey, and Brian W. Rogers. "Symmetric play in repeated allocation games." Journal of Economic Theory 154 (2014): 25-67.

Lazear, Edward P (1989). "Pay equality and industrial politics." Journal of Political Economy 97.3: 561-580.

Lefgren, Lars J., David P. Sims, and Olga B. Stoddard (2016). "Effort, luck, and voting for redistribution." Journal of Public Economics 143: 89-97.

List, John (2007): "On the Interpretation of Giving in Dictator Games," Journal of Political Economy, 115, 482–493.

Mathevet, Laurent, and Julian Romero. "Predictive repeated game theory: Measures and experiments." (2012).

McKelvey, Richard D., and Thomas R. Palfrey. "Quantal response equilibria for normal form games." Games and Economic Behavior 10.1 (1995): 6-38.

Norton, Michael I., and Dan Ariely. "Building a better America—One wealth quintile at a time." Perspectives on Psychological Science 6.1 (2011): 9-12.

Oxoby, R., & Spraggon, J. (2008). Mine and yours: Property rights in dictator games. Journal of Economic Behavior and Organization, 65, 703–713.

Rabin, Matthew (1993). "Incorporating Fairness into Game Theory and Economics." American Economic Review, 83(5): 1281-1302.

Rabin, Matthew. "Fairness in repeated games." Department of Economics, UCB (1997).

Rapoport, Anatol, and Albert M. Chammah. Prisoner's dilemma: A study in conflict and cooperation. Vol. 165. University of Michigan press, 1965.

Rapoport, Anatol, Guyer, Melvin J., Gordon, David G., The 2 X 2 Game. University of Michigan Press, Ann Arbor, 1976.

Rey-Biel, Pedro, Roman M. Sheremeta, and Neslihan Uler (2018). "When income depends on performance and luck: The effects of culture and information on giving."

Reinstein, D. and G. Riener (2012). "Decomposing desert and tangibility effects in a charitable giving experiment." Experimental Economics 15(1): 229-240.

Rosenthal, Robert W. 1989. "A Bounded-Rationality Approach to the Study of Noncooperative Games." International Journal of Game Theory, 18(3): 273—292.

Rubin, J., & Sheremeta, R.M. (2016). Principal-agent settings with random shocks. Management Science, 62, 985-999.

Ruffle, B. J. (1998). More is better, but fair is fair: Tipping in dictator and ultimatum games. Games and Economic Behavior 22, 247–265.

Rutstrom, E. E. and M. B. Williams (2000). Entitlements and fairness:: an experimental study of distributive preferences. Journal of Economic Behavior & Organization 43 (1), 75-89.

Schokkaert, Erik. 2006. "The empirical analysis of transfer motives." Handbook of the economics of giving, altruism and reciprocity 1: 127-181.

Shubik, Martin. Game theory in the social sciences. Cambridge: MIT press, 1984.

Starmans, Christina, Mark Sheskin, and Paul Bloom. "Why people prefer unequal societies." Nature Human Behaviour 1 (2017): 0082.

Tang, Fang-Fang. Anticipatory Learning in Two-Person Games: An Experimental Study, Part II. Learning. University of Bonn, Germany, 1996.