Essays in Behavioral Economics

Thesis by
Jeffrey Roy Zeidel

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2023
Defended May 15, 2023

Jeffrey Roy Zeidel
ORCID: 0009-0004-0407-2391

ACKNOWLEDGEMENTS

I am deeply indebted to Professor Thomas R. Palfrey for his exceptional and tireless advising. I am also extremely grateful to Matthew Shum for his generous and patient assistance. Many thanks also to Marina Agranov and Luciano Pomatto for their wise and helpful advice.

ABSTRACT

This dissertation contains three essays in three chapters. Chapter 1 contributes to the literature on reference dependent preferences, chapter 2 introduces a new solution concept for games played by teams of players, and chapter 3 analyzes a model of biased beliefs in law enforcement.

In Chapter 1, I study the role of reference dependent preferences in motivating effort in online chess. In online chess, players are assigned ratings that measure chess skill and update after every game. I find evidence of bunching above round numbers in the distribution of ratings, suggesting that players care about their rating and that round numbers serve as reference points. I estimate a dynamic discrete choice model of the decision to end a playing session that nests both loss aversion and an alternative ‘aspiration’ specification involving a discrete jump in utility at reference points. I reject loss aversion in favor of aspirational preferences. I show that higher skilled players are significantly more aspirational, and that aspiration does not diminish with experience.

In Chapter 2, coauthored with Jeongbin Kim and Thomas R. Palfrey, we develop a general framework for the analysis of games where each player is a team and members of the same team all receive the same payoff. The framework combines standard non-cooperative game theory with collective choice theory, and is developed for both strategic form and extensive form games. We introduce the concept of *team equilibrium* and identify conditions under which it converges to Nash equilibrium with large teams. We identify conditions on the collective choice rules such that team decisions are stochastically optimal: the probability the team chooses an action is increasing in its equilibrium expected payoff. The theory is illustrated with some binary action games.

In Chapter 3, I model a social welfare maximizing law enforcement agency that does not know the supply of crime, that may have incorrect beliefs about its ability to detect crime, and that only observes the quantity of crime that it detects. An equilibrium is defined in which the enforcement agency is not surprised by the crime data it observes, and believes itself to be maximizing social welfare. Sufficient conditions for existence are provided. The model is shown to capture the intuition of crime-policing “feedback loops" in which inefficient overpolicing or underpolicing is supported in equilibrium.

PUBLISHED CONTENT AND CONTRIBUTIONS

Jeongbin Kim, Thomas R Palfrey, and Jeffrey R Zeidel. Games played by teams of players. *American Economic Journal: Microeconomics*, 14(4):122–57, 2022. doi: 10.1257/mic.20200391.

Jeffrey R Zeidel contributed to derivation of results, formulation of illustrative examples, and preparation of the manuscript

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

INTRODUCTION

Individual and group decision makers are the smallest unit of analysis, the fundamental building blocks, of modern economic theory. All economic models studied by economists therefore rest on a foundation of some theory of economic decision making and behavior. Typically, decision makers are assumed to possess well-behaved preferences over material outcomes, to have accurate beliefs about how their behavior interacts with their environment to produces these outcomes, and to be capable of optimally responding to their environment in order to maximally satisfy their preferences. In recent decades, behavioral economists have critiqued this description of purely rational, materially oriented behavior as unrealistic, and have worked to introduce a higher degree of psychological realism into economic models.

This dissertation contains three essays that contribute to this effort. One (Chapter 1) is an empirical study on non-standard preferences, another (Chapter 2) is a theoretical study, co-authored with Jeongbin Kim and Thomas R. Palfrey, of how collective decision making procedures effect group behavior in strategic contexts, and the third (Chapter 3) is an analysis of a model of monitoring and enforcement that captures how a decision maker's incorrect, biased beliefs about their environment can be sustained even in the presence of informational feedback from which the decision maker could learn and update their beliefs toward the truth. These three chapters contribute to an understanding of how decision makers' preferences, beliefs, and decision making processes, respectively, may depart from the neo-classical vision of a rational economic actor in meaningful ways.

In Chapter 1, I study reference dependence, the dependence of a decision maker's evaluation of an outcome on a comparison to some reference outcome, and the intrinsic motivation of effort. To date, the popular loss aversion model, in which agents are more sensitive to changes in outcomes below the reference outcome than above it, has predominated in the literature on reference dependence. Loss averse agents may suffer from the endowment effect in their trading behavior if giving up their endowment is perceived as a loss, they may be willing to bear higher costs of effort to improve outcomes when they are averse to shortfalls below the reference point, and they may be unreasonably risk averse over small stakes gambles, because even small losses loom larger than small gains.

However, in some contexts, decision makers may only care about whether or not their outcome compares favorably to the reference, and may not attend to the magnitude of shortfalls. When reference points are goals that decision makers aspire to achieve, all failures to meet the goal may be perceived to be equally undesirable. This distinction between loss aversion and goal aspirations is relevant to understanding the motivation of effort in settings in which agents can choose to expend effort to improve their outcome, as in labor supply decisions and decisions over effort provision in contests, and has received little attention in the behavioral economic literature.

I leverage an ideal data set for the study of aspirational goals as reference points, a set of approximately 20 million online 'blitz' chess games played on the website lichess.org, to distinguish between two different models of reference dependence, and to study the motivational role of goals in this non-pecuniary, intrinsically motivated activity. On the website, players are assessed an individual 'glicko rating', a measure of latent chess skill, which updates after every game, adjusting upward after a win and downward after a loss. I model the decision of when to stop a playing session as a dynamic discrete choice problem. The model nests two different models of reference dependence, the common loss aversion model, and an alternative 'aspirational' model. I reject loss aversion as an explanation of player behavior, and find that the aspirational model provides a good fit to observed choices. I also investigate heterogeneity in these preferences, and find that higher skill players are more reference dependent than low skill players, and that these preferences do not diminish with experience. The result on experience runs contrary to evidence from many other economic contexts of decreasing reference dependence or other 'behavioral' patterns in behavior with experience, which I hypothesize to be due to there being no pecuniary benefit to 'unlearning' reference dependence in this activity.

In Chapter 2, coauthors Thomas Palfrey, Jeongbin Kim and I study collective decision making in teams in strategic settings. When applying game theoretic models to institutions composed of large numbers of individuals (firms, political parties, unions, governments), many researchers adopt the unitary actor assumption, modeling strategic units as individual decision makers. This assumption is unreasonable in many contexts, because the internal decision making process of those institutions influences the strategic choices of the group, and actors in strategic situations may find it necessary to account for the decision making process of other actors in order to make optimal choices themselves. For example, democracies may behave systematically differently from dictatorships in international diplomacy and war, and therefore members of democracies and dictators must rationally take into account the constitutions of their opponents when making their own decisions.

We define a new model and solution concept for games played by teams of players, in which team members share common payoffs, may have heterogeneous beliefs about the expected payoffs received from taking actions available to their team, and are endowed with collective choice rules that aggregate their heterogeneous beliefs. In a team equilibrium, team members must have on average accurate beliefs about payoffs given the behavior of all other teams, and team choices are determined by the team's collective choice rule. We identify conditions on collective choice rules that guarantee that team equilibria converge to nash equilibria as team sizes grow to infinity, and show how with small teams, a teams behavior is sensitive both to its own collective choice rule, and to the collective choice rules adopted by other teams.

In Chapter 3, I study a theoretical behavioral model of belief formation, monitoring and enforcement, in the presence of information feedback. A robust finding of behavioral economists is that agents often hold biased beliefs about the determinants of outcomes. For example agents may be persistently overconfident in their own abilities, believing that their effort is more effective at producing desirable outcomes than it is in truth. When decision makers act on their biased beliefs, they may pursue sub-optimal plans or make harmful decisions. However, if agents receive feedback from their decision that contradict their incorrect beliefs, we might expect them to learn from their prediction errors, and update their beliefs toward the truth with experience.

I contribute to the nascent theoretical literature on misspecified learning, by developing a model of a biased decision maker facing a monitoring, enforcement, or policing problem. Decision makers are endowed with a model, which may be misspecified, of the relationship between their monitoring investment and the probability of detecting a harmful action taken by another agent. Agents are assumed to make optimal investments into monitoring given their beliefs, and must not be surprised by the feedback they receive, i.e. the realized frequency of detection, given their investment decision. Decision makers with misspecified models may persistently make suboptimal investments, and never update because the feedback they observe confirms their incorrect beliefs. Thus agents in the model are capable of systematically under- or over- investing in monitoring.

If enforcement agents suffer from these biases, they may cause harm by under or over policing, and escape corrective measures because their model of crime explains the data they collect and justifies their behavior. In future work I hope to apply this model to questions about racial statistical discrimination and bias in law enforcement, to help detect these biases, and hope to use it to discover better ways to evaluate the performance of law enforcement agencies.

Chapter 1

REFERENCE POINTS AS MOTIVATING GOALS: EVIDENCE FROM ONLINE CHESS

1.1 Introduction

Over the past 30 years, an accumulation of experimental and field evidence in economics and psychology verifies the principle that an agent's evaluation of an outcome may depend in part on a comparison of that outcome to some reference outcome, or reference point, a phenomenon known as reference dependent preference (DellaVigna, 2009, Kahneman, 1992). Several classes of reference points have accrued significant empirical observation and theoretical attention. Principle among these are status quo reference points, as in prospect theory introduced by Kahneman (1979), and forward looking expectations based references, introduced by Kőszegi and Rabin (2006). A third type of reference outcome, the aspirational goal, has received considerably less attention.

In this paper, using data from the chess website lichess.org, I provide novel evidence of reference dependence and effort motivating goals in the decisions of online chess players of when to end a playing session. To facilitate matching of similarly skilled players, chess players are assessed a 'glicko rating', a measure of latent chess skill. This rating is updated after every game, increasing after a win and decreasing after a loss. I find that on average, players are discontinuously more likely to end a playing session (an uninterrupted run of short games) when their rating lies just above a multiple of 100, than when it lies just below, indicating that at least some players care about the evolution of their rating, and focus on particularly salient round numbers when evaluating ratings.

An observable consequence of this stopping behavior is so-called 'bunching' in the distribution of player ratings, ratings are more likely to be just above a round number reference point than just below. Bunching in the distribution of a variable generally arises when some related incentive or preference is discontinuous or kinked in the bunching variable.[1] In reference dependent choice, evaluations of the outcome

[1] For an example of bunching arising from rational responses to external incentives, in the field of public finance, Saez (2010) finds evidence of bunching at the EITC kink, Mortenson and Whitten (2020) finds bunching at seven different kinks in the income tax schedule, and Patel et al. (2017)in the corporate tax schedule.

variable are typically discontinuous at the reference point, the boundary between perceived 'losses' and 'gains'. Bunching of this kind has been previously observed by Allen et al. (2017), who find that the distribution of marathon run times bunches just below hour marks, and show that if evaluations of finishing times are modeled by a utility function, three different discontinuities could, separately or jointly, explain this pattern: a 'jump' in the utility function at reference points, which I will call the 'aspirational' model, a discontinuity in the first derivative at reference points, commonly referred to as 'loss aversion' when the utility function is steeper below the reference point than above, and a discontinuity in the second derivative of utility at reference points, characteristic of the original prospect theory value function. They point out, however, that which of these three types of discontinuity best explains the observed bunching cannot be identified simply from the bunched distribution itself.[2]

To separate the loss aversion and aspirational models using choice data, I adopt a structural econometric approach and model the decision to continue or end a playing session as a dynamic discrete choice problem. Upon stopping a session, the player realizes a payoff that depends upon his current rating, and how that rating compares to round number reference ratings, while choosing to play a game yields a payoff that does not depend on reference points, and continues the playing session. I estimate this model with a stopping utility that nests both piece-wise linear loss aversion, and the aspirational specification with discrete jumps in stopping utility at reference points. I also allow for a disntinction between reference dependent preferences associated with reference points above and below the status quo, taken to be the player's rating at the beginning of their session. In principle, players may perceive reference points that compare favorably to the status quo differently than they perceive reference points that are less desirable than the status quo.

The aspirational model maximum likelihood estimates fit the data better than the loss aversion specification at ratings near reference points, and in the the combined specification the estimated loss aversion parameters are significantly below 1, indicating loss tolerance rather than loss aversion. I also derive comparative statics results for stopping probability with respect to rating, and show that the loss aversion specification predicts increasing stopping probabilities in rating below reference points,

[2] Markle et al. (2018) provide further evidence on hour reference points in marathons by surveying runners and eliciting directly their subjective satisfaction with different finishing times (unincentivized). They find evidence for both loss aversion and aspiration, however the question is still open whether these stated preferences on the part of runners accurately reflects their true preferences, and whether it is reflected in their choices of effort provision throughout actual races.

contrary to observed stopping probabilities, while the aspirational specification predicts stopping probabilities decreasing in rating below reference points, matching the data. I therefore reject the loss aversion specification in favor of the aspirational specification.

I next turn to the study of the relationship between reference dependent preferences for rating and the intrinsic motivation to play and improve at chess. Reference ratings may serve as goals for players to aspire to, and players who attend more to these goals may be motivated to play more games than those who do not attend to goals, thereby accumulating more games played and improving their abilities through practice. If this were the case, we should observe a positive relationship between session length, skill, and reference dependent behavior.

A related question concerns how experience relates to reference dependence. If reference ratings motivate effort, experience may be positively associated with the intensity of reference dependent preferences. Conversely, if players it may be that as players gain experience, they learn to disregard round number comparisons to their rating. Player beliefs about their own chess ability may stabilize over time, rendering short-run rating fluctuations less important to the player, or alternatively players may come to understand round numbers are arbitrary. If so we should expect to find that experienced players are less reference dependent. This second possibility has received some supporting evidence in market contexts, in which many participants are found to become less 'behavioral' with experience. For example, List (2003) finds that the magnitude of the endowment effect among sports card traders decreases with experience, De Sousa and Munro (2012) similarly find that the endowment effect diminishes with experience in trading virtual items in an online multiplayer game, and in an experimental market, Mayhew and Vitalis (2014) found that experience mitigates myopic loss aversion. In each example, participants may have learned to maximize value through repeated market interactions. To the extent that reference dependence impedes the maximization of portfolio value, it tends to be unlearned over time. In the chess setting, as in other settings that don't involve pecuniary incentives, it is an open question whether similar experience effects exist, or whether, conversely, experience serves an important motivating function in the absence of material rewards.

To study the relationship between experience, skill and reference dependence, I estimate a heterogeneous version of the aspirational model in which all utility parameters are modeled as linear functions of two observable characteristics of the player-session observation, experience, measured by the number of sessions played previously to the current playing session, and skill level, measured by rating at the beginning of a session. The estimated coefficients for skill in the aspirational bonus parameters are both positive and statistically significant, that is, more skilled players are estimated to be more reference dependent than low skill players. Experience is estimated to be positively associated with the reference point below player's starting rating, and negatively associated with reference points that lie above session starting rating. Players of all experience and skill groups are estimated to be reference dependent to some degree. The results suggest that reference dependence is not unlearned with experience, but neither is it strongly positively associated with experience. On the other hand, skill is strongly positively associated with reference dependence.

Finally, in order to study the motivational function of reference dependent preferences, I compute the average partial effects of the utility parameters on stopping behavior. If, holding all other utility parameters constant, an increase in the aspirational bonus causes stopping probabilities to decrease on average, then we can say that aspiration is motivational and at least partly contributes to the observed heterogeneity of stopping behavior with respect to skill and experience. To the contrary, I find increasing the aspirational bonus is predicted to have little effect on stopping probability and session length. Aspiration can actually be de-motivational, and cause sessions to be shorter, because players may tend to stop playing earlier in order to avoid their rating falling below a reference point. This de-motivational effect balances out the motivational effect aspirations in this setting.

An understanding of the motivational effect of aspirational goals induced by the glicko rating system may aid in the design of non-pecuniary incentive systems that increase user demand of online chess services, or other similar services.[3] If ranking systems or explicit goals tend to increase participation, these systems might be implemented and designed in order to increase demand. Conversely, if ranking systems discourage use by inducing anxiety or fear of a falling ranking, demand might be increased by not using rankings or hiding rankings from users. Indeed, lichess has recently taken action to help players avoid rating related anxieties by implementing an optional 'zen mode' which, when activated, hides glicko ratings

[3]non-pecuniary incentive structures have also been applied in non-game contexts, as in 'gamification', in which game elements such as scores or points, rankings, badges, or levels, are incorporated into activities that traditionally do not involve these elements, for the purpose of improving motivation or productivity, or otherwise altering behavior. See Blohm and Leimeister (2013) for a review.

on the user interface.

Online chess data provide a unique opportunity for the study of effort provision and strategic decision making, and a number of other studies have recently made use of this data. Two papers have previously addressed session stopping behavior in online chess. Anderson and Green (2018) show that online chess players are more likely to quit a playing session just after achieving a personal best rating. Avoyan et al. (2020) find that some online chess players are more likely to end a playing session after a loss, and some are more likely to after a win, and apply this finding to the design of a matching algorithm that may increase session lengths. On the topic of rational inattention, Howard (2021) studies the trade-off players must make between better moves and more time for future moves, and finds that while skilled players make this trade-off optimally on average, unskilled players tend to leave too little time to future moves, they are sub-optimally too attentive to each move. Salant and Spenkuch (2021) develop a model of decision-making under complexity and tests its predictions using chess move data, finding that errors are more likely in more complex positions.

The paper is organized as follows: In section 2 I introduce the institutional setting and data, define playing sessions and measures of skill and experience; in section 3 I present a dynamic discrete choice model of when to end a playing session, and prove several results on the comparative statics of stopping probability with respect to rating; in section 4 I present the main results and compare the aspirational model with the loss aversion model, the estimates for the heterogeneous preferences model, and show how the observed heterogeneity in stopping probability and session length is not explained by heterogeneity in reference dependence; and finally conclude in section 5.

1.2 Data

Data on online chess games played from the beginning of 2013 to the end of 2015 was collected from the chess website lichess.org, a free to use, open source chess website run by volunteers and funded primarily through donations. The site offer match making services, matching players from around the world against other players of similar skill who are searching for chess games of the same type. The site additionally hosts tournaments and provides computer analysis, streaming video and other services. It makes information on all rated games publicly available, including: the account names of the opponents, their ratings, the time control, the start time, the end condition, all moves played, and computer analysis for a subset of games.

Most competitive games on the site are timed. When it is a player's turn to move, her clock counts down until she makes a move. If a player's clock hits zero, that player either immediately loses the game, or the game is drawn if the board state is such that it is impossible for the opponent to give checkmate on a future move. Typically players are given a fixed amount of time with which to make all moves (the starting time on each players clock), and in some time controls an additional amount of time (the increment) is added to the clock for each move. Games are categorized by amount of time given to each player, from longest game to shortest these are correspondence, classical, blitz and bullet.[4] Correspondence games are untimed, each of the other categories consist of multiple different controls. The longest starting time for each player for classical games in the sample is 180 minutes, but the most popular classical time control is 10 minutes for each player with no increment, so that while most classical games last for about half an hour to an hour, a portion of these games can last for longer than 6 hours. The longest blitz starting time is 7 minutes for each player, and the most popular blitz time control is 5 minutes with no increment, followed by 3 minutes with no increment. The upper bound for bullet games is 2 minutes for each player and the single most popular time control on the site is one minute for each player with no increment.

Table 1.1 displays summary statistics for the four game types. Statistics are shows for non-tournament, 'standard', games and tournament games separately. The 'Games' and 'Players' columns show the total number of rated games played and the number of unique accounts playing at least one game, in thousands, of each game type. Of the four game types, blitz is the most popular by number of games played, with a total of about 15.4 million non-tournament games, while classical is the most popular by number of unique players and amount of time spent playing. The 'Games-per-player' column shows the average number of games played by each unique account name participating in at least one game of that type. The 'Games-per-day' column shows the average number of games players initiate on days in which they play at least one game.

For the remainder of the analysis, I focus on non-tournament blitz games, because this is the most popular category by number of games and because these games are short enough to allow players the opportunity to play longer sessions.

[4] The rapid category, for game shorter than classical and longer than blitz games, was first added in 2017, after the sample period.

Table 1.1: Rated Game Summary Statistics

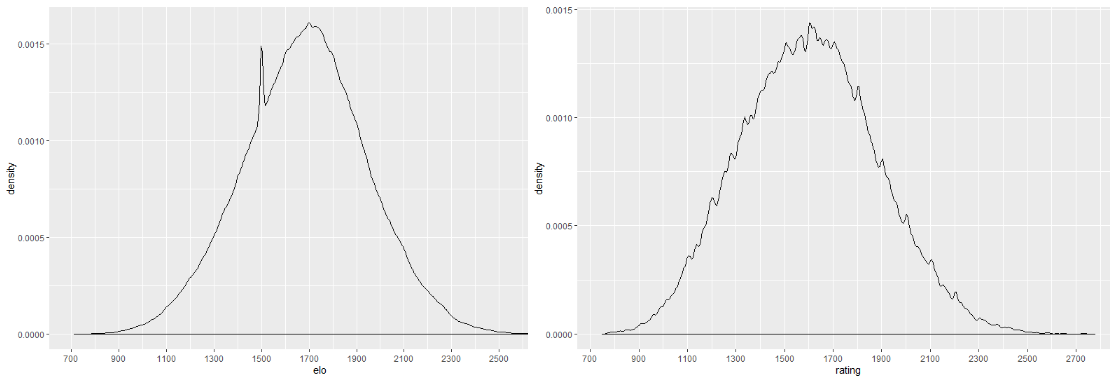
		Games (000s)	Players (000s)	Games-per-player	Games-per-day
Correspondence		157	107	2.9	2.0
Classical	standard	11309	125	180.3	4.8
	tournament	287	105	5.5	3.7
Blitz	standard	15386	121	255.2	6.8
	tournament	3021	103	58.6	7.1
Bullet	standard	10929	113	193.8	10.0
	tournament	3477	107	65.1	9.1

Glicko Ratings

In order to facilitate the matching of similarly skilled players, each player is given a rating, a measure of latent chess skill, for each category of game that user participates in. Player ratings are calculated using the Glicko-2 rating system, introduced by Glickman (2012),[5] which models player performance as a random variable with a distribution with player-specific unknown mean. When two players play a game, they each receive a realization from their distribution, if the realizations lie within some distance of each other, the game is a draw, and otherwise the player with the higher realization wins the game. A player's glicko rating is a point estimate of the mean of that player's distribution of performance, which is updated according to principles of bayesian statistics after every individual game or sample of games. Ratings increase after wins and decrease after losses, with the magnitude of the adjustment depending upon the difference in rating between the players, and in the degree of uncertainty about the player's rating, which in the Glicko-2 system is a function of the number of games recently played and the volatility in the player's recent results.

On lichess.org, ratings are updated immediately after every game, and players are able to observe their new rating upon the conclusion of each game, before choosing to play a new game or not. Furthermore, ratings are prominently displayed in multiple locations on the user interface, always next to user names. The association of ratings with the identity of the player, and the fact that ratings measure chess

[5] This rating system is a refinement of the Glicko system, detailed in Glickman (1995), which was itself an improvement on the popular Elo rating system. For details see Elo (1978). Variations and improvements on the Elo rating system are standard in many chess federations, and have also come into use in many other sports.



(a) All Ratings (b) Ratings Averaged Over Time

Figure 1.1: Glicko Rating Distribution

ability, a characteristic that players are typically interested in, make it likely that many players take a personal interest in their rating and possibly take action to improve it.

Figure 1.1 displays density plots of the distribution of blitz glicko ratings. The left panel shows the density of ratings for every player for every game in the sample. The distribution is a smooth gaussian, centered around 1700 and ranging from about 700 to 2700. 1500 is the starting rating for all new players, resulting in extra mass at that rating. The right panel shows the density of ratings at the end of playing sessions in the sample. The distribution of ratings at the beginning of playing sessions is similar, because ratings do not change until a game is played. Bunching can be observed just above round number ratings in this distribution, with corresponding missing mass just below. This bunching is not a feature of the rating system itself; if it were, bunching would be observed in the full distribution of ratings in the left panel of 1.1, rather it must be the result of decisions made by players about when to end playing sessions.

Playing Sessions

The average number of non-tournament blitz games played per day per player in the sample is 6.8. These games are often played in succession, forming sequences of games which I call playing sessions. Player decisions of when to initiate and terminate playing sessions that are sensitive to round number reference ratings can generate the bunching phenomenon discussed above.

I formally define a session as a run of games played by one user, in which there is no hour long gap between two game start times. To generate my main data set, I first

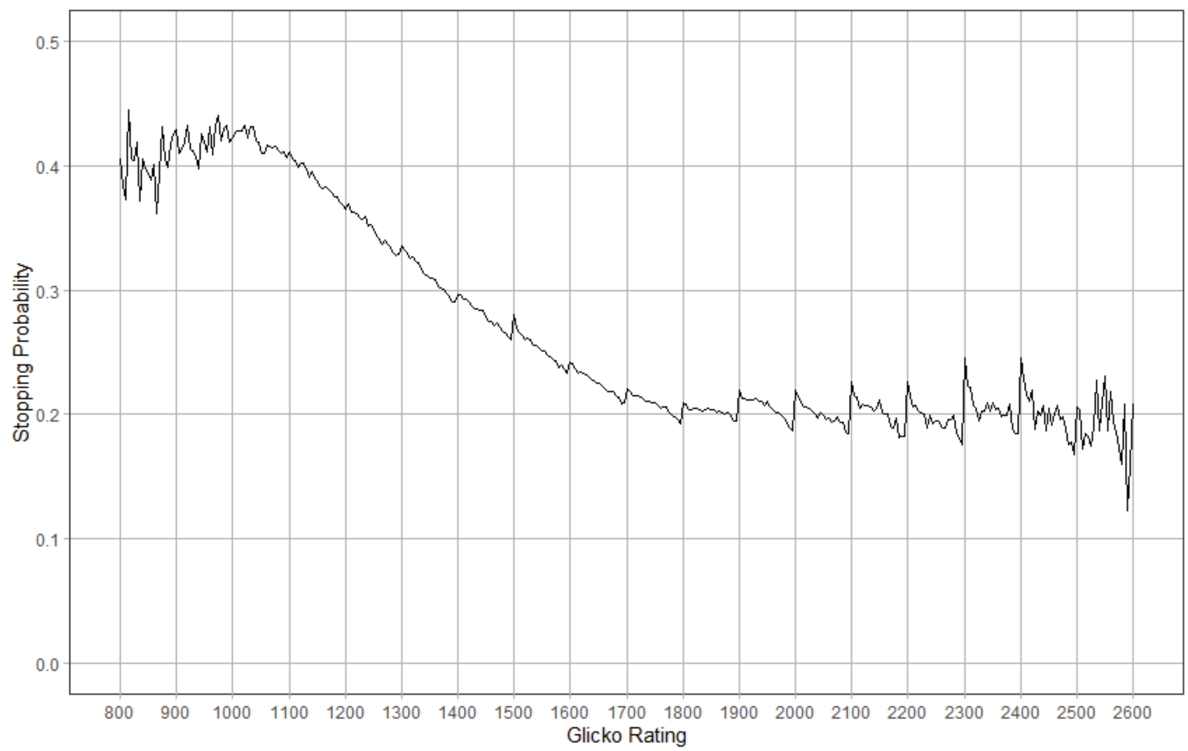


Figure 1.2: Session Stopping Probabilities By Rating

drop all non-blitz games, then separate blitz games for each player into sessions, then drop all sessions in which the player either has a provisional rating for at least one game, or plays at least one tournament game. It is necessary to drop sessions that include tournament games because tournaments typically run on a fixed schedule. After completing this process, there are 23.8 million games in 6.3 million sessions played by 176 thousand unique players in the sample. The mean session length is 3.9 games and the standard deviation is 5.5 games. The maximum number of games played in a single session was 299. Each additional game played in a session adds on average 7.9 additional minutes to the time length of the session.

In Figure 1.2, session stopping probabilities by rating are shown. This is the probability that a game is the final game in a session. There is a large scale trend of decreasing stopping probability rating, as more skilled players play longer sessions on average than lower skilled players. Stopping probabilities increase discontinuously at multiples of 100 and sharply decrease in rating just below reference points, matching the bunching pattern observed in the distribution of ratings.

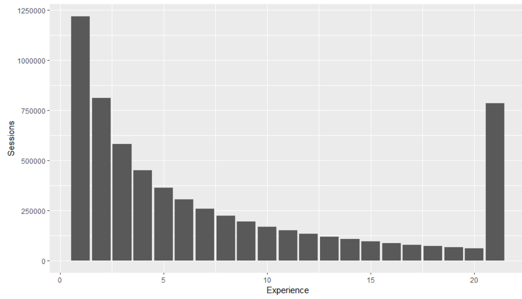
This sensitivity of stopping behavior to reference ratings clearly suggests reference dependent preferences for ratings. Participants are likely to care about their own ability, prefer to improve measures of their own ability, and act in ways that tend to accomplish this. Since current rating is the most relevant indicator of ability, it is reasonable to hypothesize that players experience psychological rewards or punishments based on their current rating. Furthermore, since ratings only update after a game, players can 'lock in' their current rating for a period of time by choosing not to play and ending their current playing session. Players should accrue a larger reward for their end of session rating, than for transitory, mid-session ratings, simply because they will enjoy this rating for a longer period of time. This may explain why players are more likely to end a session when their rating yields a relatively large psychological reward, in this case, when it compares favorably to a reference rating.

Skill and Experience

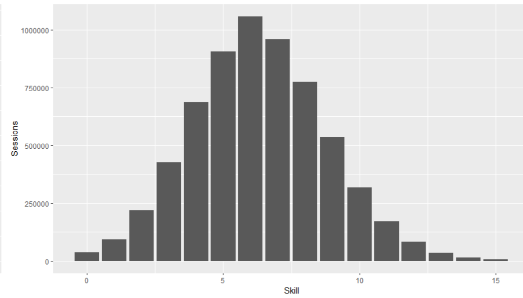
Figure 1.2 is suggestive of heterogeneity in reference dependent preferences, as the magnitude of the jumps in stopping probability at round numbers increase with rating, for example there is an increase in stopping probability from 2295 to 2300 of about 7.0%, but further down the rating distribution, the jump from 1695 to 1700 is only about 1.1%. It may be that higher skilled or more experienced players are more reference dependent than lower skill players. To facilitate the study of this heterogeneity, I introduce measures of skill and experience.

In this subsection, I present summary statistics on the relationship between skill, experience, session stopping behavior, and average session length. For each session observation, I adopt as a measure of experience the number of sessions completed before the beginning of the current session, and as a measure of skill the player rating at the beginning of the session.

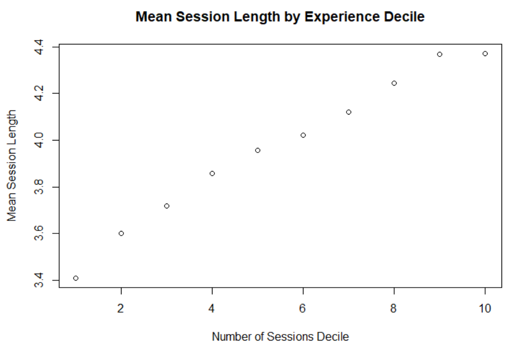
To solve the dynamic discrete choice model presented in section 3, a fixed point must be computed for every value of the parameters. In the heterogeneous version of the model, utility parameters will be modeled as functions of skill and experience. Solving for the fixed points is computationally expensive, so to render the computation estimation procedure computationally feasible, the number of possible values of the skill and experience variables must be constrained through binning. I therefore define player skill at the time of a given session, by the 100 rating point-wide bin in which the sessions starting rating lies. For example, every session in which the starting rating lies between 1200 and 1299 is assigned the same value for the skill variable, sessions with starting rating between 1300 and 1399 have a skill value that



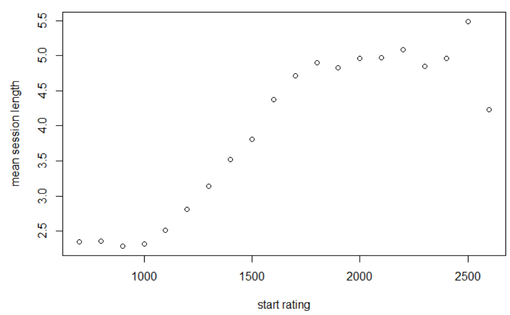
(a) Experience Distribution



(b) Skill Distribution



(c) Average Session Length by Experience



(d) Average Session Length by Skill

Figure 1.3: Experience and Skill Statistics

is one greater, and so on. All sessions with starting rating below 1000 are binned together and assigned the lowest value for skill of 0. Similarly, experience is defined by the number of sessions played, in bins of 25. Every player-session observation in which the player has previously played in 0-24 sessions is assigned an experience of 1, 25-49 and experience value of 1, and so forth, up to a value of 21 for players who have participated in at least 500 previous sessions.

Figure 1.3 shows the distributions of these variables and their relationship with average session length. Both variables are positively associated with session length, with session length increasing approximately linearly in number of sessions played, and increasing in skill in the middle of the skill distribution, then plateauing near 5 games per session above ratings of about 1800. The relationship between skill and experience, which is not shown in the figure, is non-monotonic, with players in the middle of the skill distribution having more sessions played on average than both low skill and high skill players.

1.3 Model

I model the decision of players to end or continue a playing session. Decisions in each session are made independently of every other session played by the same player. That is, I assume that when a player decides to end a session or not, he is not taking into account the effect of this decision on the payoffs he will receive from future sessions. Within a session, periods are indexed by $t \in \{0, 1, ...\}$, and the decision in period $t > 0$ occurs after t games have been played. The decision in each period starting at period 1 is $d^t \in \{0, 1\}$, where 0 denotes continue and 1 denotes stop.

To reduce the size of the state space, ratings are binned together in groups of 5, so for example every player-period observation with a glicko rating between 1795 and 1799 are binned together, 1800 to 1804 are binned together, and so on. Denote the rating bin in period t by $r^t \in \mathcal{R} = \{1, 2, ..., R\}$, with r^0 denoting the rating at the beginning of the session, before any games are played. Let $\overline{\mathcal{R}} \subset \mathcal{R}$ be a fixed set of reference rating bins, corresponding to ratings bins at multiples of 100[6].

For each player-session observation, there are two key reference ratings that figure into the stopping utility function of the player, the highest reference rating at or below r^0, denoted $\overline{r}_1$, and the lowest reference rating above r^0, denoted $\overline{r}_2$. These references differ in how they compare to the status quo at the beginning of the session. Players may perceive these reference ratings differently, weight them differently when comparing them to their current rating, or have different internal experiences when their rating approaches them. In comparing their results to the lower reference point $\underline{r}$, players are evaluating how their rating is situated relative to a bad outcome. We can interpret stopping utility shortfalls below $\underline{r}$ as punishing losses, being unfavorable in comparison to the focal rating below the status quo. In contrast, the upper reference point $\overline{r}$ may function as an aspirational, motivating goal for the player during the playing session. Ratings that exceed this point are likely to be perceived as rewarding gains.

The full set of payoff relevant, observable state variables are $z^t = (r^t, s^t, e^t)$. $s^t \in \mathbb{R}_+$ denotes the skill of the player at time t, and $e^t \in \mathbb{R}_+$ denotes the experience level at time t[7] Both of these variables are fixed within session, so the t superscript for

[6]Glicko ratings in which the two trailing digits are between 00 and 05 are binned into reference rating bins.

[7]The measures s^t and e^t are presented in section 2.4. s^t is calculated from the rating at the beginning of the session r^0, in particular sessions are binned by starting rating into bins of width 100, so that every player-session observation that has the same leading two digits are assigned the

these variables is dropped going forward. The unobserved state variable for period t is $\epsilon^t = (\epsilon_0^t, \epsilon_1^t)$, which are assumed to be two iid type 1 extreme value distributed random variables for each of the two actions. These variables capture payoff shocks flowing from opportunity costs, cognitive costs, or other costs or benefits of playing an additional game that vary stochastically between periods.

When the choice in a period is to continue playing, the flow utility received for that period is $-(c_0 + c_1(r^t - \overline{r}_1)) + \epsilon_0^t$. I call $c_0 + c_1(r^t - \overline{r}_1)$ the 'cost' of playing a game at rating r^t, $c_0(s, e)$ is the cost intercept, which gives the average cost of playing a game at the lower reference rating $\overline{r}_1$, and c_1 is a linear coefficient on rating in the playing cost function. This cost of playing a game is really the difference between the costs and benefits of playing a game at a given state. Costs may include the cost of cognitive effort or attention, the pain of losing or potentially losing a game, as well as the opportunity costs of forgoing other activities, while benefits may include the enjoyment of playing a game, winning a game, or temporarily attaining a rating between games of a session. The cost of playing is allowed to vary with rating through the parameter c_1 because all of these costs and benefits may vary with rating. For example higher rated players are on average matched with higher rated opponents, who play higher quality moves on average that necessitate more careful thought for the identification of adequate responses, so that games played at higher ratings may be more cognitively taxing on average than games played at lower ratings.

After a choice of continue, a game is played and state variables are updated following the transition probabilities $p(z^{t+1}, \epsilon^{t+1}|z^t, \epsilon^t) = p(r^{t+1}, \epsilon^{t+1}|r^t, \epsilon^t) = p(r^{t+1}|r^t)q(\epsilon^{t+1})$, which denote the probability that the next period rating is x given that the current period rating is y.

If the choice in a period is to stop, then the payoff received in that period is $v(z^t|\overline{\mathcal{R}}) + \epsilon_1^t$, the session ends and no further payoffs are received.[8] The stopping payoff $v(z^t|\overline{\mathcal{R}})$ is a reference dependent utility function, with two reference points, of the following form.

same value of s^t. e^t is calculated from the number of sessions played prior to the current session, with bins of width 15 sessions.

[8]This follows the assumption that players are myopic with respect to future playing sessions when making decisions in the present playing session.

$$v(z^t|\overline{\mathcal{R}}) = \begin{cases} \lambda_1\lambda_2\eta(r^t - \overline{r}_1) & \text{if } r^t < \overline{r}_1 \\ \alpha_1 + \lambda_2\eta(r^t - \overline{r}_1) & \text{if } \overline{r}_1 \le r^t < \overline{r}_2 \\ \alpha_1 + \alpha_2 + \eta(r^t - \overline{r}_2) + \lambda_2\eta(\overline{r}_2 - \overline{r}_1) & \text{if } \overline{r}_2 < r^t \end{cases}$$

Figure 1.4 illustrates the stopping utility in the case when $\alpha_1, \alpha_2, \eta > 0$ and $\lambda_1, \lambda_2 > 1$. The two reference points $\overline{r}_1$ and $\overline{r}_2$ divide the domain of rating bins into three regions, the region below the lower reference point, which can be thought of as a region of losses, a gain region above the upper reference point, and a status quo region between the two reference points where the first two digits of ratings are equal to the first two digits of the session starting rating. Associated with each reference point are piece-wise linear loss aversion parameters λ_1 and λ_2, and aspirational bonus parameters α_1 and α_2. The parameter η is a 'gain' utility parameter, giving the slope of the utility function with respect to rating bin in the gain domain. The slope of the utility function in the status quo region is $\lambda_2\eta$ and the slope in the loss domain is $\lambda_1\lambda_2\eta$. The loss aversion parameters are therefore the ratio of the marginal utilities of rating above and below reference points, as in other piecewise-linear loss aversion models common in the literature. When $\eta > 0$ and $\lambda_1, \lambda_2 > 1$, players are loss averse, they are more sensitive to changes in rating below reference points than above.

When $\alpha_1, \alpha_2 > 0$, the player gains a 'utility bonus' from ending their session above reference points. These bonuses parameters differ from loss aversion parameters in that they do not influence the slope of the stopping utility function with respect to rating. The specification nests a pure loss aversion model, when $\alpha_1 = \alpha_2 = 0$, and also nests a pure bonus specification when $\lambda_1 = \lambda_2 = 1$. In addition, single reference point specifications are nested, when for example $\lambda_1 = 1$ and $\alpha_1 = 0$, only the upper reference point is relevant to session stopping decisions, and the lower reference point acts like any non-reference rating bin.

To study the relationship between experience, skill, and preferences, a heterogeneous version of the model is estimated, in which all utility parameters discussed above are modeled as linear functions of skill and experience. For example, the value of the utility gain parameter η for a given session is a linear function of skill and experience, parametrized by (η_0, η_e, η_s), the intercept, the coefficient on experience and the coefficient on skill, respectively, and for any given session with observable state variable $z^t = (r^t, s, e)$, we have $\eta(z^t) = \eta_0 + \eta_e e + \eta_s s$. For each utility parameter X,

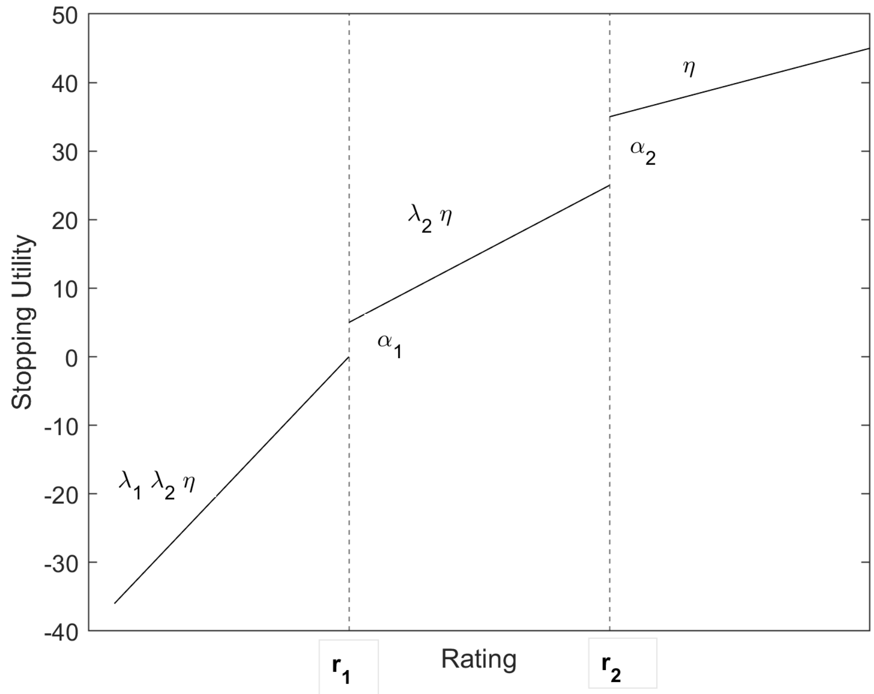


Figure 1.4: Example of Stopping Utility

the parameters X_e, X_s give the estimated relationship between preference parameter X and experience or skill respectively, holding the other constant. Since the value of e and s are fixed within session, the value of each utility parameter is also fixed within session, and players maximize their sum of discounted payoffs given their fixed utility functions over rating.

Given the utilities of continuing and of stopping the playing session for each value of the state variable, and the transition probabilities for ratings, players maximize their discounted expected utility, in every period choosing the maximum of $u(d^t = 1, z^t, \epsilon^t) = v(z^t|\overline{\mathcal{R}}) + \epsilon_1^t$ and $u(z^t = 0, r^t, \epsilon^t) = -c_0(z^t) - c_1(z^t)r^t + \epsilon_0^t + \beta\mathbb{E}[V(z^{t+1}, \epsilon^{t+1})|d^t = 0, z^t, \epsilon^t]$, where the value function is given by the recursive formula $V(z^{t+1}, \epsilon^{t+1}) = \max\{u(d^t = 1, z^t, \epsilon^t), u(z^t = 0, r^t, \epsilon^t)\}$. Given the distributional assumptions on ϵ^t, choice probabilities are given by the familiar logit formula below.

$$P(z^t) = \mathbb{P}[d^t = 1|z^t] = \frac{\exp\{v(z^t|\overline{\mathcal{R}})\}}{\exp\{v(z^t|\overline{\mathcal{R}})\} + \exp\{c(z^t) + \delta \mathrm{EV}(z^t)\}}$$

where the expected value function $EV(z^t)$ is the integral of the value function with respect to rating and structural errors at time $t+1$. Since the expected value of the maximum of two type 1 extreme value variables with means x, y is $log[\exp\{x\} + \exp\{y\}]$, we have the $EV(z^t)$ is the unique fixed point of the contraction mapping $T(\cdot)$ written below

$$\mathrm{EV}(z^t) = T(\mathrm{EV})(z^t) = \sum \log[\exp\{v(z^{t+1}|\overline{\mathcal{R}})\} + \exp\{c(z^{t+1}) + \delta \mathrm{EV}(z^{t+1})\}] p(r^{t+1}|z^t)$$

Reference Dependence and Stopping Probabilities

The two models, loss aversion and aspiration, both nested in the stopping utility function introduced above, generate sharply divergent predictions of stopping probabilities, and how stopping probabilities change with rating. Generally, in this dynamic setting, the stopping probability of loss averse players will be increasing in rating just below reference ratings, while aspirational players will have decreasing stopping probabilities in rating just below reference ratings. Below, I prove that this property holds for a subset of the parameter space of the model, in which there is only one reference rating $\overline{r}_1$ relevant to stopping decisions, there is no skill or experience heterogeneity, and the cost of playing is constant over current rating. This occurs when $\lambda_2 = 1$, $\alpha_2 = 0$, and $c_1 = 0$. The value of stopping is then given by

$$v(r^t|\mathcal{R}) = \mathbb{1}\{r^t < \overline{r}_1\}[\lambda_1 \eta (r^t - \overline{r}_1)] + \mathbb{1}\{r^t \geq \overline{r}_1\}[\alpha_1 + \eta(r^t - \overline{r}_1)]$$

First note that, for any distribution of ϵ^t, we have that the probability of stopping at rating r^t is

$$P(r^t) = \mathbb{P}[\epsilon_1^t - \epsilon_0^t > \delta \mathrm{EV}(r^t) - v(r^t|\mathcal{R})]$$

We can see that $P(r^t + 1) \geq P(r^t)$, the stopping probability increases from rating r^t to rating $r^t + 1$, if and only if $\delta \mathrm{EV}(r^t + 1) - \delta \mathrm{EV}(r^t) \leq (v(r^t + 1) - v(r^t))$, i.e.

if the value of stopping increases faster than the delta discounted expected value of continuing. The following proposition uses the fact that $T(\cdot)$ is a contraction mapping to establish a useful upper bound for $\delta\text{EV}(r^t + 1) - \delta\text{EV}(r^t)$.

Proposition 1: Suppose $c_1 = 0$. For any value function v, $\delta\text{EV}(r^t+1) - \delta\text{EV}(r^t) \leq \max_{r^t}\{v(r^t + 1) - v(r^t)\}$

Proof. Denote $m = \max_{r^t}\{v(r^t + 1) - v(r^t)\}$.

Suppose EV^0 is such that $\text{EV}^0(r^t + 1) - \text{EV}^0(r^t) \leq \frac{1}{\delta}m$ for all r^t. We can write

$$
\begin{aligned}
T(EV^0)(r^t + 1) - T(EV^0)(r^t) = & \\
&= \sum_{r^{t+1}}\{\log(\frac{\exp(v(r^{t+1} + 1)) + \exp(\delta EV^0(r^{t+1} + 1))}{\exp(v(r^{t+1})) + \exp(\delta EV^0(r^{t+1}))})\}p(r^{t+1}|r^t) \\
&\leq \sum_{r^{t+1}}\{\log(\frac{\exp(v(r^{t+1}) + m) + \exp(\delta EV^0(r^{t+1}) + m)}{\exp(v(r^{t+1})) + \exp(\delta EV^0(r^{t+1}))})\}p(r^{t+1}|r^t) \\
&= \sum_{r^{t+1}}\{\log(\exp(m))\}p(r^{t+1}|r^t) \\
&= m \leq \frac{1}{\delta}m
\end{aligned}
$$

The result follows from the fact that T is a contraction mapping.

□

The proposition says that the continuation value cannot increase in rating faster than the fastest rate at which the stopping utility increases. This means that, at those ratings where the stopping value is increasing at its fastest possible rate, the stopping probability must be increasing, leading to the following two corollaries corresponding to the pure loss aversion case and the pure aspirational case respectively.

Corollary 1: When $c_1 = 0$, $\lambda_2 = 1$ and $\alpha_1 = \alpha_2 = 0$, $\eta > 0$, and $\lambda_1 > 1$, we have $P(r) - P(r - 1) \geq 0$ whenever $r \leq \bar{r}_1$.

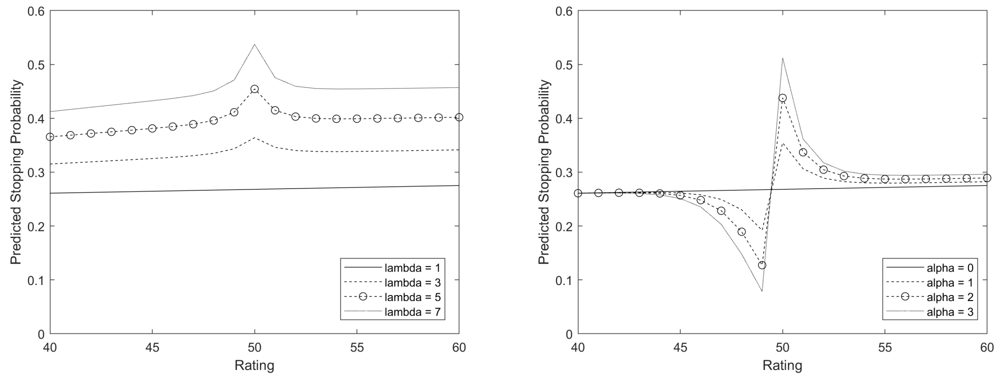
Corollary 2: When $c_1 = 0$, $\lambda_1 = \lambda_2 = 1$, $\alpha_2 = 0$, $\eta > 0$, and $\alpha_1 > 0$, we have $P(\bar{r}_1) - P(\bar{r}_1 - 1) \geq 0$.

Corollary 1 says that under loss aversion with a single reference point, the stopping probability must be increasing in rating at all ratings below the reference point. Corollary 2 says that under the aspirational model with a single reference point, there must be a jump in stopping probability at the reference point.

The differences in the stopping behavior of players under these two models can be further illustrated computationally in the following simplified example. Let the set of ratings be $\mathcal{R} = \{1, ..., 99$ and the only reference rating $\bar{r}_1 = 50$. Assume that the transition probabilities are $p(r^t + 1|r^t) = p(r^t - 1|r^t) = 0.5$, for all $r^t \neq 1, 99$, and $p(2|1) = p(1|1) = p(98|99) = p(99|99) = 0.5$. That is, at every rating the probability the rating increases or decreases by 1 (upon choosing to continue) is 0.5, except when the rating is already at the minimum or maximum possible rating, in which case there is 0.5 chance the rating stays the same. Fix the cost of playing to $c_0 = 0.5$, the gain parameter $\eta = 0.1$, and the variance of the unobservable error terms to 1.

In Figure 1.5 I plot the predicted stopping probabilities by rating, under different values for the parameters α_1, λ_1. In the left panel, the value of α_1 is fixed to 0, and the value of λ_1 varies from 1 (no loss aversion) to 7. Increasing the value of λ_1, the degree of loss aversion, increases the stopping probability at every rating, and also increases the prominence of a 'peak' in stopping probability centered at the reference rating. In the right panel, the value of λ_1 is fixed at 1 and the value of α_1 is varied from 0 (no aspirational preference) to 3. As α_1 increases, stopping probabilities above the reference rating increase, and decrease below the reference rating. The closer the rating is to the reference rating, the larger the partial derivative of stopping probability with respect to α. As a result, increasing α_1 causes the magnitude of the discontinuous jump in stopping probability at the reference rating to increase. In all cases, at ratings far away from the reference point, stopping probabilities increase with rating due to the positive value of the gain parameter η.

Note that, following the result of Corollary 1, pure loss aversion implies that stopping probabilities increase in rating at ratings below the reference point. In this numerical example, pure aspirational preferences predict the opposite, decreasing probabilities in rating below the reference point. These properties are generally robust to changes in the values of the other parameters, and to the addition of more reference points. Comparing these predictions to the empirical choice probabilities displayed in Figure 1.2, it seems likely that the aspirational model outperforms loss aversion in fitting the data. In the next section I confirm this conjecture by estimating the model



(a) Loss Aversion Only (b) Aspiration Only

Figure 1.5: Simplified Model Stopping Probabilities

parameters and evaluating model.

1.4 Estimation and Results

To estimate maximum likelihood model parameters, I implement the nested fixed point algorithm introduced by Rust (1987), which follows a two stage procedure. The data consists of tuples $x_i^t = (d_i^t, r_i^t, s_i, e_i)$, where d_i^t is the decision of the player whether to end session i in period t, r_i^t is the rating in session i, period t, and s_i, e_i are the skill and experience of the player in session i. Denote by T_i the total number of periods in session i, and denote by I the total number of sessions. Under the independence assumptions on the unobserved state variables, the full likelihood and log-likelihood functions are

$$\ell(\{x_i^t\}) = \prod_{i=1}^{I} \prod_{t=1}^{T_i} P(r_i^t, s_i, e_i)^{d_i^t} (1 - P(r_i^t, s_i, e_i))^{1-d_i^t} p(r_i^t | r_i^{t-1})$$

$$\log(\ell(\{x_i^t\})) = \sum_{i=1}^{I} \sum_{t=1}^{T_i} P(r_i^t, s_i, e_i)^{d_i^t} (1 - P(r_i^t, s_i, e_i))^{1-d_i^t} + \sum_{i=1}^{I} \sum_{t=1}^{T_i} p(r_i^t | r_i^{t-1})$$

In the first stage of the estimation procedure, maximum likelihood estimates of state transition probabilities are obtained from sample averages of the data. These estimates maximize the second sum in the log-likelihood formula above. In the second stage, utility parameters are estimated via maximization of the first sum of the log-likelihood function, using the transition probability estimates from the first stage. This maximization involves an 'outer loop' that searches over values of the utility parameters, in every run of the loop computing the choice probabilities for a given set of parameter values. To compute choice probabilities, the expected value function must be computed, necessitating an 'inner loop' that solves the contraction mapping $T(\cdot)$ at every iteration of the outer loop.[9]

Homogeneous Stopping Utility

In this section, I present estimates of a version of the model in which all players are assumed to have identical stopping utilities, while the cost of playing is allowed to vary with skill and experience, in other words $\lambda_1(z^t), \lambda_2(z^t), \eta(z^t), \alpha_1(z^t), \alpha_2(z^t)$ are all constants in z^t, and $c(z^t) = c_0$. Heterogeneity in playing costs is necessary for the model to adequately fit the macro-scale trend of decreasing average stopping probabilities in skill and experience present in the data.[10]

The first column of table 1.2 reports the estimated stopping utility parameters of the full model specification including loss aversion and aspirational preferences. Columns 2 and 3 present estimates for the restricted, loss aversion only and aspiration only models. Standard errors are reported in the parentheses. The scale of the utilities are fixed by the assumption that the structural error terms have variance $\frac{\pi}{6}$. At every state, the standard deviations of the stage payoffs of both playing a game and stopping are roughly 0.72. The magnitude of the parameter estimates should be compared to this value.

In the full model, the marginal benefit of increasing rating above the upper reference point, η is estimated to be about 0.3, or roughly three sevenths of a standard deviation of the error terms. Both loss aversion parameters are estimated to be below one, indicating that players are on average loss tolerant, they are less sensitive to changes in rating below the upper reference point than above. These estimates contrast sharply with the typical finding in the behavioral literature, recently summarized in a meta analysis by (Brown et al., 2021), which found that the mean reported loss aversion estimate from 150 articles in economics and psychology was between 1.8 and 2.1. The products of the estimated loss aversion parameters with η give the estimated slope of the stopping utility between the two reference points and

[9] To solve for the fixed point of $T(\cdot)$, I adopt the 'polyalgorithm' described in Rust (2000). Successive iterations of $T(\cdot)$ are first performed, then Newton-Kantorovich iterations are used to convergence.

[10] I find that when playing costs are assumed to be homogeneous, the marginal benefit of rating η is estimated to be negative, which allows the model to fit this decreasing stopping probability trend but renders the stopping utility estimates unintelligible because a negative marginal benefit of rating would imply players could maximize their stopping utility by intentionally losing their games.

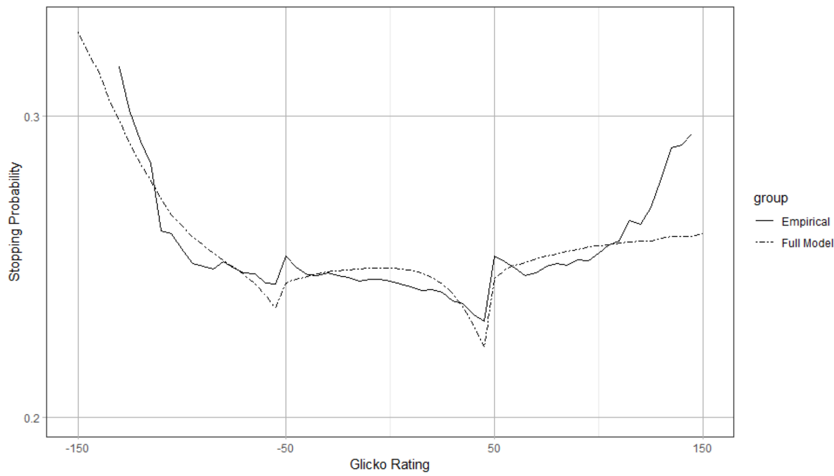
Table 1.2: Homogeneous Stopping Utility Estimates

Parameter	Full Model	Loss Aversion	Aspiration
Upper Bonus (α_2)	0.1743	0^*	0.1589
	(0.0013)		(0.0014)
Lower Bonus (α_1)	0.0692	0^*	0.0432
	(0.0011)		(0.0009)
Gain (η)	0.2948	0.2953	0.2454
	(0.0020)	(0.0020)	(0.0019)
Upper Loss (λ_2)	0.7651	0.7656	1^*
	(0.0021)	(0.019)	
Lower Loss (λ_1)	0.8312	0.8739	1^*
	(0.0023)	(0.0018)	
log-likelihood	-1.3212e7	-1.3214e7	-1.3214e7

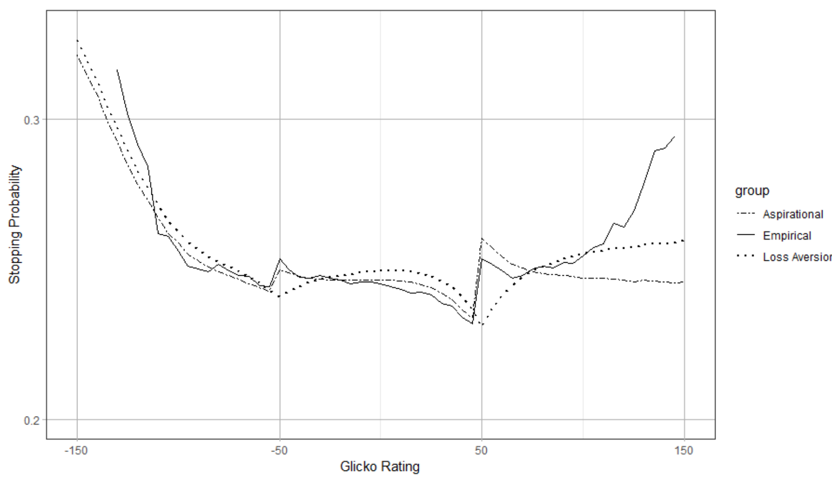
below the lower reference point, which are about 0.226 and 0.187 respectively. The estimated slopes in the restricted model with $\alpha_1 = \alpha_2 = 0$ are close to those in the full model, at 0.226 and 0.188.

The aspirational estimates are positive, with the upper bonus more than twice the size of the lower bonus. Both bonuses are smaller than the estimated slope of the stopping utility everywhere, implying that a player's stopping utility is always increased more by increasing his rating by two bins, than by increasing his rating by one bin from just below to just above a reference point. These results also hold in the restricted model with loss aversion parameters fixed at 1.

In Figure 1.6, observed stopping probabilities and predicted stopping probabilities from the fitted models are plotted. In each panel of the Figure, the x-axis is the difference between the current rating and the center of the starting 100 point-wide rating bin. All players begin their session with rating between -50 and +50, the lower reference rating is at -50 and the upper reference rating is at +50, and zero on the x-axis corresponds to a rating with trailing digits between 50 and 55. Data for all skill and experience groups is thereby pooled together based on how the current rating compares to the two reference points. The solid black line shows the observed sample averages of stopping decision by bin. Stopping probability initially decreases in rating, plateaus at close to 0.25 just below the lower reference point, then increases in rating above the upper reference point. At the lower reference point, the stopping probability jumps by about 0.009 in one rating bin, and at the upper reference point it increases by 0.022. Between these two points, the stopping probability steadily decreases by a total of 0.022.



(a) Full Model Fit



(b) Restricted Model Fit

Figure 1.6: Homogeneous Model Fit

In the top panel, the dashed line shows the predicted stopping probabilities from the fitted full model. The predicted stopping probabilities generally co-moves with rating in the same direction as the data, with the exception of the region between the lower reference point and the mid-point between the reference points, and the region just above the upper reference point, in which the predicted stopping probabilities increase in rating. The predicted troughs and just below, and peaks at reference points are too low, though the predicted size of the jumps from -55 to -50 and from 45 to 50 approximately match the data, at .008 and .023 respectively.

In the bottom panel, the dashed line shows the predicted probabilities for the restricted aspiration only model, and the dotted line for the loss only model. This plot illustrates how the two restricted models decompose the predicted behavior from the full model fit into a portion attributable to loss tolerance, and a portion attributable to aspirations. The loss only model predictions form inverted tent shapes around reference points, with probabilities decreasing below the reference rating, then increasing above it, with no discontinuous jump at the reference point. As rating approaches a reference point, the player become more risk loving for gambles over rating, and therefore more likely to play an additional game. In contrast, the aspirational model behavior more closely matches the observed behavior, with jumps in stopping probability at reference points, and stopping probability decreasing in rating everywhere else. The aspirational model predicts jumps of 0.007 and 0.027 at the lower and upper reference points respectively, and predicts a gradual decrease in stopping probability of 0.016 between reference points.

The fit of the aspirational model is superior to that of the loss aversion model at every bin from -50 to +70, except for one where the predictions of the two models cross, with the best over-performance at ratings just above reference points, where the loss aversion model fails to predict any discontinuous jump. The fit of the loss aversion model is superior at all ratings outside of that range. The loss aversion parameters control the slope of the stopping utility over the entire range of ratings, so that predicted behavior at all ratings is influenced by small changes in the loss aversion parameters, whereas the aspirational bonuses can only influence behavior when a player's rating is close to a reference point, and it is more likely for the player's rating to cross the reference point in the near future. This may explain why the loss aversion model has an advantage in fitting data far away from reference points, whereas the aspirational model has an advantage near reference points, where observed behavior qualitatively matches the predictions of that model. It is clear from these observations how the two sets of parameters contribute to the fit of the full model. The full model predictions coincide with the loss aversion model predictions far from reference points, the discontinuous behavior at reference points is entirely the contribution of the aspirational model, and between reference points the full model predictions are a mix of the two restricted models.

Heterogeneous Stopping Utility

In this subsection I present estimates for the version of the aspirational model with heterogeneity in utility parameters. The specification of playing costs remains unchanged from the models estimated in the previous section, and every stopping utility parameter is modeled as a linear function of skill and experience.

Table 1.3 shows the coefficient estimates for each of the stopping utility parameters. First, the estimated aspirational bonus for the upper reference point is larger than the estimated bonus for the lower reference point, for players of all skill and experience. Both α_1 and α_2 increase in skill, with α_2 increasing faster, so that the gap between α_1 and α_2 grows with skill. The estimated coefficient on experience is closer to 0 for both, with a negative coefficient for the upper bonus and positive for the lower bonus, implying that the difference shrinks with experience. Reference dependence does not generally diminish in this setting, in contrast to various market settings in which it has been shown that loss aversion diminishes with experience. The gain parameter η has a negative intercept, but is strongly positively associated with skill, so that by the second lowest skill group it is positive. The fact that η is estimated to be less than zero for some players is likely the result of the linear specification.

The fixed playing costs c is estimated to be decreasing in both skill and experience, but with a positive second derivative with respect to both, the interpretation being that higher skill and experience is associated with lower costs, or high average payoff, from playing an additional game. This is unsurprising, as we should expect players who enjoy playing games more, or who have lower opportunity costs for their time, to play more games and improve their skill over time, and conversely those who have played more games and have greater ability are likely to have acquired stronger tastes for playing. The linear in rating playing cost parameter k, is estimated to be negative for all players, indicating that the average payoff from playing increase in rating. One potential explanation might be that at higher ratings players are matched with higher ability opponents, and games played against stronger competition may be more enjoyable.

In Figure 1.7, average stopping probabilities by rating are plotted along with the predicted stopping probabilities from the fitted model. The quadratic specification of c allows the model to neatly fit both the decreasing stopping probability in rating in the lower to middle portion of the rating distribution, and the plateau around 0.2 above a rating of 1800. The model also fits the increasing sensitivity to reference ratings with increasing skill.

The empirical and predicted stopping probabilities for two different skill groups are shown in Figure 1.8, players with session starting ratings between 1400 and 1500 in the top panel, starting ratings between 2000 and 2100 in the bottom panel. These figures illustrate the key observed patterns in the relationship between skill and stopping behavior. First, stopping behavior for both skill groups is more responsive

Table 1.3: Heterogeneous Stopping Utility Estimates

	Coefficient Estimates				
	Intercept	Experience	Skill	Experience^2	Skill^2
Upper Bonus (α_2)	0.0331	-0.0025	0.0287		
	(0.0035)	(0.0002)	(0.0005)		
Lower Bonus (α_1)	-0.1190	0.0054	0.0197		
	(0.0031)	(0.0002)	(0.0006)		
Gain (η)	-0.1651	-0.0207	0.1061		
	(0.0031)	(0.0002)	(0.0009)		
Cost (c)	0.5968	-0.0038	-0.0736	0.0001	0.0032
	(0.0002)	(3.3e-5)	(9.5e-5)	(1.6e-6)	(8.6e-6)
Linear Cost (k)	-0.0026	0.0002	-0.0004		
	(3.2e-5)	(2.4e-6)	(9.4e-6)		
log-likelihood	-1.3212				

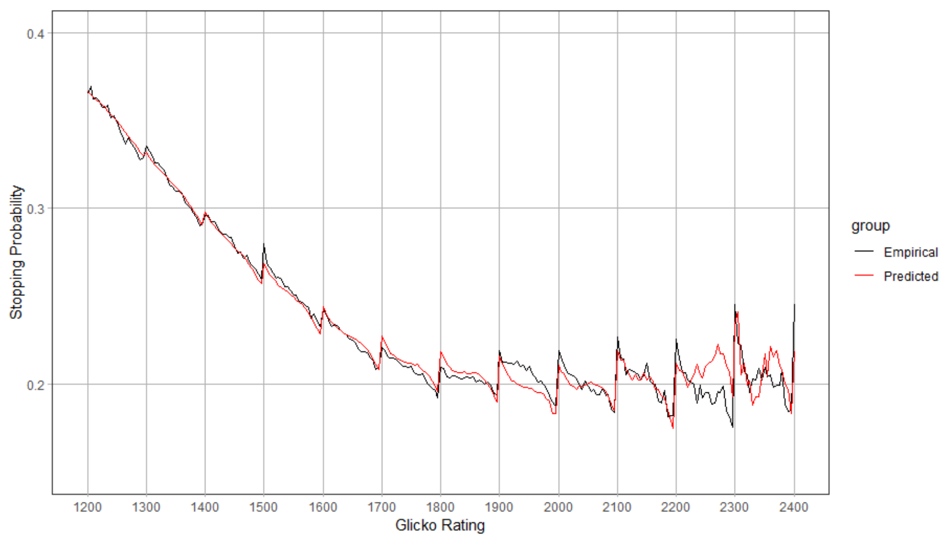


Figure 1.7: Heterogeneous Model Fit

to the reference point that lies above their starting rating, in fact the low skill group has no apparent jump in stopping probabilities at the lower reference point, and only a small jump of about 0.02 at the upper reference point. The higher skill players react more strongly to both reference points than do lower skill players, with jumps in stopping probability of about 0.031 and 0.065 at the lower and upper reference points respectively.

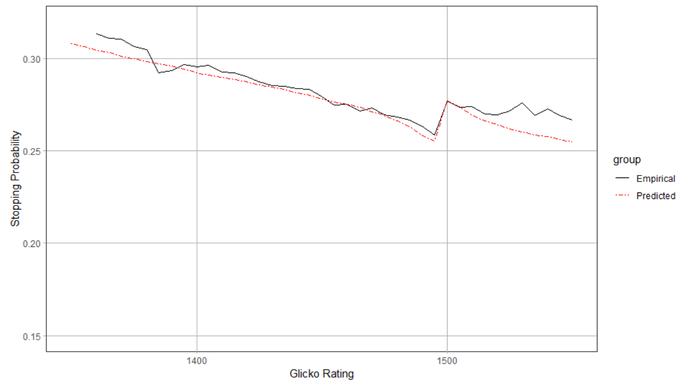
Separating the data by skill group makes it clear that nearly the entirety of the stopping probability jumps and bunching at reference points below the median rating of about 1500, is the result of the behavior of players who have improved their rating from the beginning of their playing session. For example, the jump at a rating of 1300 apparent in Figure 1.7 is almost entirely caused by players who started their session between 1200 and 1300 and played until their rating rose above 1300 and stopping, and almost none of it is caused by players who started between 1300 and 1400 and stopped their session prematurely before their rating fell below 1300. At higher ratings, players with falling ratings and rising ratings both contribute to bunching at a given reference rating, however the asymmetry in aspirational preferences between $\overline{r}_1$ and $\overline{r}_2$ means that a majority of the players with end of session rating in the bunching regions improved their rating over their session.

Other notable patterns in the data include that the stopping probability of the low skill group is above that of the high skill group at every rating bin, and decreases in rating, while the high skill players become more likely to stop as rating increases. The model fits all of these patterns, with the higher skill group estimated to have a lower playing cost, higher marginal benefit of increasing rating, and greater aspirational bonuses, resulting in lower stopping probabilities, a positive slope in rating, and larger jumps at reference points.

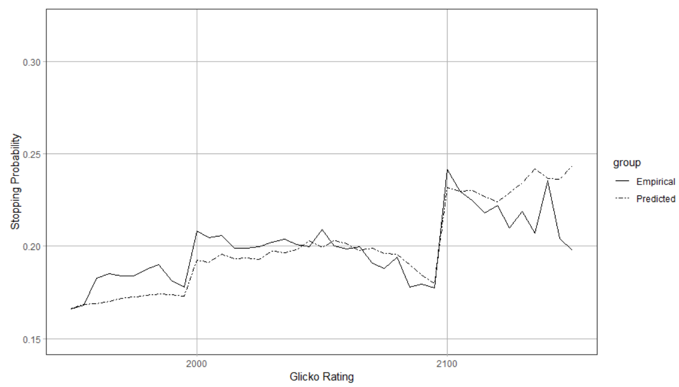
Aspiration and Session Length

Reference dependent, aspirational preferences may motivate players to prolong their playing sessions in an effort to improve to their goal rating, thereby playing more games per session on average, accumulating more games played and improving their abilities at a faster rate if they are able to learn from their experience. If this is the case, the heterogeneity in aspiration found in the previous section may provide a unified explanation for the relationship between responsiveness in stopping behavior to reference points, average session lengths, and skill levels of the players in the sample. That is, higher skilled players may play more games and improve their abilities at least in part because they are more motivated to achieve goal ratings.

In this section I use the estimated heterogeneous aspirational model to study this question. First, I decompose the average partial effects of skill and experience on stopping probability into the various utility parameter channels. We can write the expected partial derivatives of stopping probability with respect to the skill and experience as



(a) Skill Group 1400-1500



(b) Skill Group 2000 - 2100

Figure 1.8: Heterogeneity of Reference Dependence in Skill

$$\mathbb{E}[\frac{\partial P}{\partial s}\Big|_z] = \sum_{\theta\in\Theta}\mathbb{E}[\frac{\partial P}{\partial \theta}\Big|_z \frac{\partial \theta}{\partial s}\Big|_z]$$
$$\mathbb{E}[\frac{\partial P}{\partial e}\Big|_z] = \mathbb{E}[\sum_{\theta\in\Theta}\frac{\partial P}{\partial \theta}\Big|_z \frac{\partial \theta}{\partial e}\Big|_z]$$

where Θ is the set of utility parameters in the model, here α_1, α_2, η, c, and k, and the expectations are taken with respect to the distribution of the observable state variable. Each summand above is one component of the effect of skill or experience on average stopping probability. They can be interpreted as the change in average stopping probability when a player is given the estimated utility parameter of an otherwise identical player, but in the next higher skill or experience bin. By the consistency of the maximum likelihood estimates and the continuous mapping theorem, consistent estimates for these average partial effects can be computed

Table 1.4: Partial Effects

Average Partial Effects			
θ	$\mathbb{E}[\frac{\partial P}{\partial \theta}]$	$\mathbb{E}[\frac{\partial P}{\partial \theta}\frac{\partial \theta}{\partial e}]$	$\mathbb{E}[\frac{\partial P}{\partial \theta}\frac{\partial \theta}{\partial s}]$
α_1	0.0108	5.9e-5	0.0002
α_2	-0.0038	9.7e-6	-0.0001
η	0.0655	-0.0013	0.0070
c	0.7286	-0.0012	-0.0264
k	7.2313	0.0012	-0.0032
		$\mathbb{E}[\frac{\partial P}{\partial e}] = -0.0012$	$\mathbb{E}[\frac{\partial P}{\partial s}] = -0.0225$

by differentiating the stopping probabilities at the estimated parameter values, and taking sample averages.

Table 1.4 shows the estimated average partial effects of experience and skill, acting through the various utility parameters. The average partial effects of experience and skill, shown at the bottom of the table, are both negative, confirming that higher skill and experience are associated with lower stopping probabilities and longer sessions. These are simply the sums of the terms in the column above them in the table.

The partial effect of α_1 is positive, the more players care about their rating not dropping below the lower reference point, the sooner they end their session on average. The other bonus parameter α_2 on the other hand has a negative effect on stopping probability, because players that care more about reaching the next higher reference point will prolong their session in order to reach that point, they are willing to bear a larger total playing cost in order to have a better chance of reaching their goal. These effects are small however, because stopping probability moves in opposite directions with α above and below reference points. Since both parameters increase with skill, the effect of increasing skill on stopping probability, acting through the α_1 and α_2 channels, is positive and negative, respectively. On average, when a player is given the α_1 parameter of an otherwise identical player, but in one skill bin higher, their average stopping probability is predicted to increase by 0.0002, and when given the α_2 parameter of a player in the next highest skill bin, their stopping probability decreases by -0.0001. The combined effect of changing both reference dependence utility parameters is simply the sum of these two effects, meaning that on balance, reference dependence is estimated to be very slightly de-motivational for players in this sample.

The model attributes most of the heterogeneity in stopping probabilities to difference in fixed playing costs. Higher playing cost induces larger stopping probabilities, and both skill and experience are estimated to be negatively related to playing cost. Notably, the effect of the gain parameter is positive, so that the higher the marginal benefit of increasing the rating is, the more often players end their session, and this effect is much stronger than the effect of reference dependence.

1.5 Conclusion

Users of an online chess website are discontinuously more likely to end their sessions above reference points than below them. Using structural, reference dependent, dynamic discrete choice model of the decision to end a playing session that nests loss aversion and an alternative aspirational utility specification, I show that loss aversion fails to explain stopping behavior near reference points, but that aspiration explains it well, matching the qualitative features of observed behavior and providing a superior fit to the loss aversion model near reference points.

I also show that players have heterogeneous reference dependent preferences, with the magnitude of the estimated aspirational bonuses growing significantly with skill, but only a small relationship between reference dependence and experience. The result that reference dependence does not diminish with experience is notable, as many recent studies have found that decision-makers become less loss averse with experience. One potential explanation for this discrepancy is that, in this setting, there are no pecuniary rewards at stake, and so 'behavioral' decisions are not as materially costly as they are in market settings.

Indeed, in the absence of financial incentives, intrinsic psychological incentives, like the goal oriented ones studied in this paper, may serve an important motivational function, and decision makers who are not motivated by goals may be more likely to stop participating, leaving only goal motivated players behind to become more experience and skilled. It could also be that reference dependence in session stopping decisions directly motivates effort, by causing players to prolong their sessions. However, an analysis of the partial effects of the utility parameters reveals that aspirational preferences are neither motivational nor de-motivational, in the sense that increasing the magnitude of player aspirations neither increases nor decreases average stopping probability, and therefore the average number of games played per session. To confirm the validity of these causal claims of the aspirational model, it would be necessary to gather experimental evidence in future work on the effect of goals on the provision of effort and risk taking in a similar environment. If exogenous manipulations of the location and importance of goals or reference points could be performed without monetary incentives, evidence of their causal effects on behavior could be gathered and compared to the predictions of the aspirational model.

Beyond developing a better understanding of the psychology of decision involving effort and risk, the results of this study have the potential to inform the design of non-pecuniary incentives in other settings, such as in the video game industry, or in the gamification of other activities.

Chapter 2

GAMES PLAYED BY TEAMS OF PLAYERS

2.1 Introduction

For most applications of game theory, each "player" of the game is actually a team of players. For reasons of analytical convenience and longstanding tradition, these teams are modeled as if they are unitary actors - i.e., single individuals. Examples abound. In spectrum auctions, the players are giant corporations such as Verizon, AT&T, and Sprint. The same is true in virtually any model used to study problems in industrial organization: oligopoly, limit pricing and entry deterrence, R&D races, and so forth. In the crisis bargaining literature aimed at understanding international conflict, the players are nation states. In the political arena key players include parties, civic organizations, campaign committees, large donor groups, commissions, panels of judges, advisory committees, etc. These “teams" range not only in size and scope but also in their organizational structure and procedures for reaching decisions.

A basic premise of the theoretical framework developed in this paper is that the unitary actor approach misses a critical component of these strategic environments, namely the collective choice problem within each competing team. This premise is not merely conjectural but is supported by a growing body of experimental work that has begun to uncover inconvenient facts pointing to important behavioral differences between games played by teams of players and games played by individual decision makers.

Many of the studies that compare group and individual behavior in games find that team play more closely resembles the standard predictions of game theory. To quote from Charness and Sutter (2012): “In a nutshell, the bottom line emerging from economic research on group decision-making is that groups are more likely to make choices that follow standard game-theoretic predictions...” Similarly, Kugler et al. (2012) summarize the main finding of their survey in the following way: “Our review suggests that results are quite consistent in revealing that group decisions are closer to the game-theoretic assumption of rationality than individual decisions.” A similar conclusion has been reached in many individual choice experiments as well. For example a variety of judgment biases that are commonly observed in individual decision making under uncertainty are significantly reduced by group decision-making.[1]

Given these extensive findings about group vs. individual choice in games, which cross multiple disciplines, it is perhaps surprising that these observations remain in the category of "anomalies" for which there is no existing general theoretical model that can unify these anomalies under a single umbrella. In particular, one hopes such a model might apply not only to interactive games, but also to non-interactive environments such as those experiments that have documented similar team effects with respect to judgment biases and choice under uncertainty. This paper takes a step in that direction.[2]

Our theoretical framework of team games combines two general approaches to modeling strategic behavior and team behavior: non-cooperative game theory and collective choice theory.[3] Non-cooperative game theory provides the basic structure of a strategic form game, formalized as a set of players, action sets, and payoff functions, or more generally a game in extensive form, which includes additional features including moves by nature, order of play, and information sets. The focus here is exclusively on games played by teams in a pure common value setting, i.e., all players on the same team share the same payoff function.[4]

[1] Such biases include probability matching (Schulze and Newell, 2016), hindsight bias (Stahlberg et al., 1995), overconfidence (Sniezek and Henry, 1989), the conjunction fallacy (Charness et al., 2010), forecasting errors (Blinder and Morgan, 2005), and inefficient portfolio selection (Rockenbach et al., 2007).

[2] Charness and Sutter (2012) and others have offered some qualitative conjectures about factors that might play a role in the differences between group and individual decision making. For example, perhaps group dynamics lead to more competitive attitudes among the members, due to a sense of group membership. Or perhaps groups are better at assessing the incentives of their opponents; or groups follow the lead of the most rational member ("truth wins"). Of these conjectures, our model is closest to the last one, in the sense that the aggregation process of the diverse opinions can produce better decisions if there is a grain of truth underlying those opinions. However, this would depend on the collective decision-making procedures.

[3] There is also a more distant connection with the economic theory of teams. See Marschak and Radner (1972), although the focus there is on other issues, such as communication costs, with no strategic interaction between different teams.

[4] The assumption of common values is motivated to a large extent by the many experimental studies of games played by teams of players, where all players on the same team receive exactly the same payoff. Given the extensive empirical findings in these pure common value settings, it seems like the natural starting point for developing a team theory of games. In principle, this could be extended to allow for heterogeneous preferences among the members of the same team, for example diverse social preferences. It should be clear that the focus here is *not* on applications where individual members of a group engage in costly private investments of different amounts for the benefit of some common outcome for the group, as for example in partnership games, voter turnout games, or more generally in public good contribution games where free riding plays a key role. Those settings already have their own extensive theoretical and empirical/experimental literatures.

Collective choice theory provides an established theoretical structure to model the effect of different procedures or rules according to which a group of individuals produces a group decision. If all members of the team have perfectly rational expectations about the equilibrium expected payoffs in the game, then they could all agree unanimously on an optimal action choice, and the collective choice problem would be trivial. For this reason, our approach relaxes the usual assumption of perfectly rational expectations about the expected payoffs of actions. Instead, individual members' expectations about the payoff of each available action to the team are correct *on average*, but subject to unbiased errors, so that members of the same team will generally have different expectations about the payoff of each available action, which one can view as opinions, but on average these expectations are the same for all members and equal the true (equilibrium) expected payoffs of each action.[5]

Thus, the aggregation problem within a team arises because different members of the team have different opinions about the expected payoffs of the available actions, where these different opinions take the form of individual estimates of the expected payoff of each possible action. The collective choice rule is modeled abstractly as a function mapping a profile of team members' opinions (i.e., estimates) into a team action choice. Because the individual estimates are stochastic, this means that the action choices by a team will not be deterministic, but will be "as if" mixed strategies, with the distribution of a team's effective mixed strategy a product of both the error distribution of the individual team members' estimates and the collective choice rule that transforms these estimates into a team action choice. The equilibrium restriction is that individuals have rational expectations on average, given the mixed strategy profile of all the other teams, which results from aggregation of their members' diverse estimates via some collective choice rule for each of the other teams.

Even though the collective choice rules are modeled abstractly, many of these collective choice rules correspond to voting rules or social choice procedures that are familiar. For example, if a team has exactly two possible actions, then the *majority rule* would correspond to a collective choice rule in which the team's action choice is the one for which a majority of members estimate to have the higher expected payoff (with some tie-breaking rule in case of an even-number of

[5]These errors could alternatively be interpreted as idiosyncratic additive payoff disturbances, as in quantal response equilibrium. In fact, if each team has only one member, the team equilibrium of the game will be a quantal response equilibrium (McKelvey and Palfrey, 1995), (McKelvey and Palfrey, 1996), (McKelvey and Palfrey, 1998), because there is no collective choice problem. That correspondence generally breaks down for teams with more than one member.

members). With more than two actions, this could be extended naturally to *plurality rule*. *Weighted voting* would give certain member estimates more weight than others. At the extreme, a *dictatorial rule* would specify a particular team member, and the team action choice would be the one that is best in the opinion of only that team member. A *Borda rule* would add up the individual opinion ordinal ranks of each action and choose the one with the highest average ranking.

Many other collective choice rules that are included in our formulation do not have an obvious natural analog in social choice or voting theory. Various choice procedures involving direct communication between team members could be encoded in a collective choice rule. For example, an *average* rule would average the members' announced estimates of each action's expected payoff and choose the action with the highest average opinion. Thus, the notion of collective choice rules includes all familiar ordinal-based rules, but is broader in the sense that it includes rules that can depend on the cardinal values of the estimates as well.

The existing literature on games played by teams of players is extensive and growing, and essentially all focused on experimental investigations of differences between the choice behavior of teams and individuals, where - as in the theory presented here - team choices are determined by an exogenously specified collective choice rule and all members of the same team receive identical payoffs. There are two identifiable strands depending on whether the experimental task was a multi-player game (such as the prisoners dilemma), or a single-player decision problem (such as a lottery choice task or the dictator game). There are far too many papers to describe them all here, and the interested reader should consult the surveys of experimental studies of groups vs. individuals by Charness and Sutter (2012) and Kugler et al. (2012) mentioned earlier.

The focus here is on the experimental studies of games rather than single-agent decision tasks, although we note that the broad finding in both classes of studies is that group decision making conforms more closely to economically rational behavior than individual decision making. The range of games studied to date is quite broad. The earliest studies were conducted by social psychologists who were interested in examining alternative hypotheses about social dynamics, based on psychological concepts such as social identity, shared self-interest, greed, and schema-based distrust (fear that the other team will defect). The consistent findings in those studies is that teams defect more frequently than individuals.[6] Bornstein

[6]The experimental social psychology literature on the subject is extensive. See, for example,

and Yaniv (1998) find that teams are more rational than individuals in the ultimatum game, in the sense that proposers offer less and responders accept less. Elbittar et al. (2011) study several different voting rules in ultimatum bargaining between groups, with less clear results, but also report that proposers learn with experience to offer less. Bornstein et al. (2004) find that teams "take" earlier than individuals in centipede games. In trust games, Kugler et al. (2007), Cox (2002), and Song (2008) find that trustors give less and trustees return less. Cooper and Kagel (2005) find that teams play more strategically than individuals in a limit-pricing signaling game. Charness and Jackson (2007) compare two different voting rules for team choice in a network-formation game that is similar to the stag-hunt coordination game. They report a highly significant effect of the voting rule. Sheremeta and Zhang (2010) observe 25% lower bids by teams than individuals in Tullock contests, where individuals bid significantly above the Nash equilibrium. A similar finding is reported in Morone et al. (2019) for all pay auctions. Group bidding behavior has also been investigated in auctions (Cox and Hayne (2006), Sutter et al. (2009)). Most studies compare the behavior of individuals with the behavior of teams of 2 or 3 members. Variations in team size are not usually considered. An exception is Sutter (2005) in which an individual, a team with two members, and a team with four members play a beauty-contest game for four rounds. He finds that the behavior of 4-member teams is closer to the Nash equilibrium action, while there is no significant difference between individuals and 2-member teams.

We are aware of only two other comparable theoretical models of team behavior in games. Duggan (2001) takes the opposite approach to the present paper, by assuming that members of the team share common and correct beliefs about the distribution of actions of the other teams, but have different fixed (i.e., non-stochastic) payoff functions. The team action is assumed to be the core of a voting rule. With this approach, existence of team equilibrium typically fails because of non-existence of a core for many voting rules in many environments. Cason et al. (2019) proposes a model specifically for the prisoner's dilemma game that incorporates homogeneous group-contingent inequity-averse preferences and common/correct beliefs. The team decision is determined by a symmetric quantal response equilibrium of the within-team majority-rule voting game, assuming all members have identical inequity averse preferences. In their model, voting behavior in the team decision

Insko et al. (1988) and several other studies by Insko and various coauthors. This literature refers to this difference between teams and individuals as a “discontinuity effect". Wildschut and Insko (2007) provide a survey of much of this literature in the context of various explanations that have been proposed.

process becomes more random as team size increases which can lead to behavior further from Nash equilibrium, in contrast to the experimental findings cited above, and in contrast to the results in this paper.

We first develop the formal theoretical structure of finite team games in strategic form, and provide a proof of the general existence of team equilibrium. In Section 3, the effects of changing team sizes on team equilibrium are illustrated with three examples with majority rule in 2×2 games. These effects can be rather unintuitive: while these examples illustrate how majority rule converges to Nash equilibrium with large teams, they also show that convergence is not necessarily monotone in team size; i.e., larger teams can lead to team equilibria further from Nash equilibrium. Furthermore, in mixed strategy equilibrium, individual voting probabilities within a team can be very different from the team mixed strategy equilibrium; in fact, if the Nash equilibrium is mixed, then individual voting probabilities converge to one-half, while the team equilibrium converges to the Nash equilibrium. Section 4 generalizes the finding in the examples in Section 3 that team equilibrium with large teams converges to Nash equilibrium under majority rule in 2×2 games. We prove that this *Nash convergence property* holds in *all* finite n-person games if the collective choice rules used by each team is a scoring rule. Section 5 shows that under general conditions on collective choice rules, team choice behavior will satisfy two different kinds of stochastic rationality for all finite n-person games: *payoff monotonicity*, where the probability a team chooses a particular action is increasing in its equilibrium expected payoff; and *rank dependence*, where a team's choice probabilities will always be ordered by the expected payoffs of the actions. Section 6 generalizes the framework to extensive form games, establishes existence, and proves that the results about stochastic rationality and Nash convergence extend to arbitrary extensive form team games. In fact, every convergent sequence of team equilibrium as teams grow large necessarily converges to a *sequential equilibrium* of the game. Team equilibrium in extensive form games are illustrated in Section 7, with a sequential prisoner's dilemma game and the four-move centipede game. Section 8 discusses the results of the paper and points to some possible generalizations and extensions of the framework.

2.2 Team Games in Strategic Form

A team game is defined as follows. Let $\mathbf{T} = \{1, ..., t, ..., T\}$ be a collection of *teams*, where $t = \{i_1^t, ..., i_j^t, ..., i_{n^t}^t\}$, where i_j^t denotes member j of team t, and denote the *team size profile* by $n = (n^1, ..., n^T)$. Each team has a set of available actions,

$A^t = \{a^t_1, ..., a^t_{K^t}\}$ and the set of action profiles is denoted $A = A^1 \times ... \times A^T$. The *payoff function* of the game for team t is given by $u^t : A \to \mathfrak{R}$. Given an action profile a, all members of team t receive the payoff $u^t(a)$. A mixed strategy for team t, α^t, is a probability distribution over A^t, and a mixed strategy profile is denoted by α. We denote the expected payoff to team t from using action a^t_k, given a mixed strategy profile of the other teams, by $U^t_k(\alpha) = \sum_{a^{-t}\in A^{-t}} \left[\prod_{t'\neq t} \alpha^{t'}(a^{t'})\right] u^t(a^t_k, a^{-t})$. For each t, given α, each member i^t_j observes an estimate of $U^t_k(\alpha)$ equal to true expected payoff plus an estimation error term. Denote this estimate by $\hat{U}^t_{ik} = U^t_k(\alpha) + \varepsilon^t_{ik}$, where the dependence of $\widehat{U}^t_{ik}$ on α is understood.[7] We call $\hat{U}^t_i = (\hat{U}^t_{i1}, ..., \hat{U}^t_{iK^t})$ i's estimated expected payoffs, and $\hat{U}^t = (\hat{U}^t_1, ..., \hat{U}^t_{n^t})$ is the profile of member estimated expected payoffs in team t. The estimation errors for members of team t, $\{\varepsilon^t_{ik}\}$, are assumed to be i.i.d. draws from a commonly known probability distribution F^t, which is assumed to have a continuous density function that is strictly positive on the real line. We also assume the distribution of estimation errors are independent across teams, and allow different teams to have different distributions. Denote any such profile of estimation error distributions, $F = (F^1, ..., F^T)$, *admissible.* A *team collective choice rule*, C^t, is a correspondence that maps profiles of estimated expected payoffs in team t into a nonempty subset of elements of A^t. That is, $C^t : \mathfrak{R}^{n^t K^t} \to \mathcal{A}^t$, where $\mathcal{A}^t$ is the set of nonempty subsets of A^t. Thus, for any α and ε^t, $C^t(\hat{U}^t) \in \mathcal{A}^t$. Different teams could be using different collective choice rules, and denote $C = (C^1, ..., C^T)$ the profile of collective choice rules.

For any strategic form game $G = [\mathbf{T}, A, u]$, and for any admissible F and profile of team choice rules C, call $\Gamma = [\mathbf{T}, A, n, u, F, C]$ a *team game in strategic form.* We assume that team t chooses randomly over $C^t(\hat{U}^t)$ when it is multivalued, according to the uniform distribution.[8] That is, the probability team t chooses a^t_k at $\hat{U}^t(\alpha)$ is given by the function g^t_k defined as:

$$g^t_k(\hat{U}^t) = \frac{1}{|C^t(\hat{U}^t)|} \text{ if } a^t_k \in C^t(\hat{U}^t) \qquad (2.1)$$
$$= 0 \text{ otherwise}$$

An example of a team choice rule is the *average rule*, which mixes uniformly over the actions with the highest average estimated payoff. That is, let $\overline{\hat{U}}^t_k = \frac{1}{n^t} \sum_{i\in T} \hat{U}^t_{ik}$

[7] In some instances later we will write out the dependence on α explicitly to avoid ambiguity.

[8] The assumption that ties are broken fairly is made for convenience to reduce notation and to avoid artificially creating a source of bias into all collective choice rules. Ties could be broken by other means, for example by choosing the lowest index element of $C^t(\hat{U}^t)$. It is not essential to the results in the paper, except for Section 5.2 and Theorem 8, where neutrality of C^t is assumed.

and define $C^t_{ave}(\hat{U}^t) = \{a^t_k \in A^t | \overline{\hat{U}}^t_k \geq \overline{\hat{U}}^t_l \text{ for all } l \neq k\}$ to be the set of actions that maximize $\overline{\hat{U}}^t$ at given values of α^{-t} and ε^t. This implies a team mixed strategy, g^{ave}, defined by:

$$g^{ave}_k(\hat{U}^t) = \frac{1}{|C^t_{ave}(\hat{U}^t)|} \text{ if } a^t_k \in C^t_{ave}(\hat{U}^t)$$
$$= 0 \text{ otherwise}$$

Another example is *plurality rule*, which chooses the action for which the greatest number of team members estimate to have to highest payoff. That is, define $V^t_k(\hat{U}^t) = \#\{i \in t | \hat{U}^t_{ik} \geq \hat{U}^t_{il} \text{ for all } l \neq k\}$ and then $C^t_{pl}(\hat{U}^t) = \{a^t_k \in A^t | V^t_k(\hat{U}^t) \geq V^t_{k'}(\hat{U}^t) \text{ for all } k' \neq k\}$. Then g^{pl} is defined by:

$$g^{pl}_k(\hat{U}^t) = \frac{1}{|C^t_{pl}(\hat{U}^t)|} \text{ if } a^t_k \in C^t_{pl}(\hat{U}^t)$$
$$= 0 \text{ otherwise}$$

Team response functions and team equilibrium

It is important to note that the team choices are generally stochastic (unless C is a constant function), and for any given distribution of other teams' action choices, the distribution of the mixed strategy by team t is inherited from the estimation error distribution via a team collective choice rule. It is this distribution of each team's choices under their team collective choice rule that is the object to which we ascribe equilibrium properties.

Given a team game $\Gamma = [\mathbf{T}, A, n, u, F, C]$, we can define a *team response function* for team t, $P^{C^t} : \Re^{K^t} \rightarrow \Delta A^t$, a function that maps profiles of expected utilities for team actions to a team distribution over actions by taking an expectation of g^{C^t} over all possible realizations of ϵ^t:

$$P^{C^t}_k(U^t(\alpha)) = \int_{\epsilon^t} g^{C^t}_k(\hat{U}^t(\alpha)) dF^t(\epsilon^t) \tag{2.2}$$

where $g^{C^t}_k(\hat{U}^t)$ is defined as in equation (2.1). An equilibrium of a team game is a fixed point of $P \circ U$.

Definition 1. *A* ***team equilibrium*** *of the team game* $\Gamma = [\mathbf{T}, A, n, u, F, C]$ *is a mixed strategy profile* $\alpha^* = (\alpha^{*1}, ..., \alpha^{*T})$ *such that, for every* t *and every* $k = 1, ..., K^t$, $\alpha^{*t}_k = P^{C^t}_k(U^t(\alpha^*))$.

Theorem 1. *For every* Γ *a team equilibrium exists.*

Proof. This follows in a straightforward way. With the admissibility assumptions on F, the integral on the right hand side of equation (2.2) is well-defined for all admissible F and $P^{C^t}_k$ is continuous in α. Brouwer's fixed point theorem then implies existence. □

It is one thing to define aggregation rules in the abstract, but quite another to model how such an aggregation rule might be implemented within a team. In a sense, team games nest a game within a game, but the definition above models the game within a game in reduced form, via the function g^{C^t}. Different communication mechanisms might correspond to different collective choice rules. For example, one possibility for the average rule would be a mechanism in which each player announces $\hat{U}_i^t$ to the other members of the team, and the team just takes the average, and then chooses the action that maximizes the average announced estimated expected payoff. Because it is a common value problem for the team, there is an implicit assumption of sincere reporting, and the team is choosing optimally. One might also conjecture that free form communication within a group would lead to group choice approximating the average rule, with a dynamic similar to what has been theoretically modeled as group consensus formation (McKelvey and Page (1986), and others).[9]

More directly, the group might implement a collective choice rule such as plurality, qualified majority, or Borda count, by voting. The next section illustrates team equilibrium in 2×2 games under simple majority and qualified majority rule. The examples illustrate three different features of team equilibrium: (a) limiting properties of team equilibrium as teams become large; (b) team equilibrium with a large team playing against a small team; and (c) team equilibrium when teams use a supermajority collective choice rule. Most of the focus of the examples is on the first of these features, i.e., how does team equilibrium change as the team sizes increase?

The examples illustrate how an increase in team sizes can be understood conceptually in terms of two different effects. One effect, which is especially intuitive in 2×2 games where teams use majority rule, is *the consensus effect.* In this case, suppose, for some fixed α, $U_1^t(\alpha) > U_2^t(\alpha)$. Then for any admissible F, the probability that $\widehat{U}_{i1}^t(\alpha) > \widehat{U}_{i2}^t(\alpha)$ for an individual member of team t, which we denote by $p_1^t(\alpha)$ is greater than $\frac{1}{2}$. Thus, *fixing* α^{-t}, under majority rule, $P_1^{C^t}(U^t(\alpha), n) = \sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k} (p_1^t(\alpha))^k (1 - p_1^t(\alpha))^{n-k} > \frac{1}{2}$ and increases in n monotonically, eventually converging to 1, because an increase in team size increases the likelihood of a majority consensus for a_1^t, and this consensus is guaranteed in the limit as n increases without bound.[10]

---

[9]The formal connection between the average rule and the consensus formation literature is not direct, as that approach assumes the members of the group share a common prior. Our model of estimation errors does not specify a common prior distribution of expected payoffs from actions. Rather individual beliefs are modeled simply as unbiased point estimates of an unknown true value.

[10]If $U_1^t(\alpha) = U_2^t(\alpha)$ then $p_1^t(\alpha)$ is exactly equal to $\frac{1}{2}$ and $P_1^{C^t}(U^t(\alpha), n) = \frac{1}{2}$ for all n, so there is

Thus, for any fixed strategy profile α, as n grows for a team, the estimated expected payoffs of different actions by individual members of a team converges on average to the "true" expected payoffs of the actions. Similarly, the individual members' ranking of an action's estimated expected payoff converges to it's true ranking, given α. Thus, for many rules, such as plurality rule, scoring rules, and the average rule, the collective choice of large teams will reflect a consensus about the relative expected payoffs of the various actions, as if all members shared common and correct beliefs about $U^t(\alpha)$. In the case of majority rule, this is simply a consensus about whether $U_1^t(\alpha)$ is greater than, less than, or equal to $U_2^t(\alpha)$.

Of course the equilibrium analysis is more complicated than this. One cannot take α as fixed when n changes, because of the second effect, *the equilibrium effect*. As n changes, the team equilibrium α^* changes, so we denote its dependence on n here by α_n^*. Even in the case where $|A^t| = 2$, it will typically be the case that for $n \neq n'$, $U_1^t(\alpha_n^*) - U_2^t(\alpha_n^*) \neq U_1^t(\alpha_{n'}^*) - U_2^t(\alpha_{n'}^*)$, so under majority rule $p_1^{*t}(\alpha_n^*) \neq p_1^{*t}(\alpha_{n'}^*)$, which feeds back and affects each team's mixed strategy response. Thus, while the consensus effect only looks at how changes in one team's size affect the actions of that team, fixing the actions of the other teams, the equilibrium effect takes into account that as team size changes (even for a single team) the actions frequencies of *all* teams will typically change. Because it is an indirect rather than a direct effect, the equilibrium effect can produce some unintuitive consequences for some games and some voting rules, and it's possible that the equilibrium effect can dampen or even work in the opposite direction of the consensus effect.

What ultimately happens in the limit of team equilibrium with large teams depends on the relative strength of these two effects, which will typically be game-dependent and rule-dependent. In Section 4, we show that these two effects interact in a way such that, for a broad class of collective choice rules, in all games every limit point of team equilibria as teams grow without bound is a Nash equilibrium of the game, which we call *Nash Convergence*. Nash convergence can fail to hold if the collective choice rules are non-neutral, i.e., biased in favor or against certain actions. The next section provides an example illustrating such a failure of Nash convergence with a non-neutral rule, where a sequence of team equilibria in the prisoner's dilemma game converges to a mixed strategy in the limit.[11]

no consensus effect.

[11]Nash convergence also generally fails with large teams if one of the team's collective choice rule is dictatorial. This is shown in one of the examples below. A related question is whether the limit points of team equilibria are restricted by familiar refinements of Nash equilibrium, such as

The interaction of these two effects can also lead to other surprising properties of team equilibrium. For example, because the equilibrium effect can work in the opposite direction of the consensus effect, with relatively small team sizes, the effect of increasing team size can drive the team equilibrium away from the Nash equilibrium of the underlying normal form game. This can even arise in games that are strictly dominance solvable with a unique rationalizable strategy profile. In 2×2 games where team equilibria with majority rule converge to a mixed strategy Nash equilibrium, the consensus effect essentially disappears with large teams, because the individual voting probabilities converge to $\frac{1}{2}$ as the expected payoffs of the two actions approach equality. These and other phenomena that can arise in team equilibrium are illustrated with examples in the next section.

2.3 Team Size Effects in 2×2 Games

The model is easiest to illustrate in the simple case of 2×2 games with majority rule as the collective choice rule for each team.. That is, the team choice is the action for which a majority of team members estimate to have a higher expected payoff, with ties broken randomly. Let $T = \{1, 2\}$, $A^t = \{a_1^t, a_2^t\}$, and the set of action profiles is $A = A^1 \times A^2$. Denote by α^t the probability that team t chooses action a_1^t. Then member i of team t estimates the expected payoff if team t chooses action a_k^t when team $-t$ uses a mixed strategy α^{-t} by $\widehat{U}_{ik}^t = U_k^t(\alpha) + \varepsilon_{ik}^t$ where $U_k^t(\alpha) = \alpha^{-t} u^t(a_k^t, a_1^{-t}) + (1 - \alpha^{-t}) u^t(a_k^t, a_2^{-t})$. Because there are only two actions for each team and majority rule depends only on each member's ranking of the estimated payoff of each action, the notation can be simplified by letting $\varepsilon_i^t = \varepsilon_{i1}^t - \varepsilon_{i2}^t$ denote the difference in estimation errors for individual i on team t and denote by H^t the distribution of the *difference* of these estimation errors, ε_i^t.[12] Given that each of the payoff estimation errors, ε_{i1}^t and ε_{i2}^t, are distributed according to an admissible error distribution, H^t is also admissible and symmetric around 0. That is, for all $z \in \Re$, $H^t(z) = 1 - H^t(-z)$, implying $H(0) = \frac{1}{2}$. Thus, we can write $\Delta\widehat{U}_i^t(\alpha) \equiv \widehat{U}_{i1}^t(\alpha) - \widehat{U}_{i2}^t(\alpha) = U_1^t(\alpha) - U_2^t(\alpha) + \varepsilon_i^t$.

trembling-hand perfection and proper equilibrium. The answer is negative. In particular, there are examples of simple games with limit points of team equilibria under plurality rule that are not perfect (and hence not proper). The refinement implied by limit points of team equilibrium seems more closely related to approachability in the sense of Harsanyi (1973).

[12] Some collective choice rules depend on more than just the profile of estimated payoff differences. An example of such a collective choice rule is the maximin rule:

$$C_{\max\min}^t(\hat{U}^t) = \{a_k^t \in A^t | Min_i\{\widehat{U}_{ik}^t\} \geq Min_i\{\widehat{U}_{il}^t\}\ \forall l \neq k\}$$

i.e., select the action that has the highest minimum estimated expected payoff.

A team choice rule, C^t, maps each profile of individual estimated expected payoff error, $\Delta\widehat{U}^t(\alpha) = (\Delta\widehat{U}_1^t(\alpha), ..., \Delta\widehat{U}_{n^t}^t(\alpha))$ into the nonempty subsets of A^t, and g_C^t randomly selects one of these choices with equal probability. We next illustrate team size effects under majority rule in three different kinds of 2×2 games.

The Prisoner's Dilemma played by teams

Consider the family of PD games displayed in Table 2.1, where the two parameters $x > 0$ and $y > 0$ are, respectively the payoff gain from defecting if the other player cooperates and the payoff gain from defecting if the other player defects.

Table 2.1: Prisoner's dilemma game

	Column Team (2)	
Row Team (1)	Cooperate(C)	Defect(D)
Cooperate(C)	$5, 5$	$3 - y, 5 + x$
Defect(D)	$5 + x, 3 - y$	$3, 3$

Suppose that both teams have n (odd) members and the same distribution of estimation error differences, H, let α_n^* be a symmetric team equilibrium probability of either team choosing D in the game, and $p_n^* = \text{Prob}(U_C^t(\alpha^*) - U_D^t(\alpha^*) < \epsilon_D^t - \epsilon_C^t)$ be a symmetric team equilibrium probability that any player votes for action D. Equilibrium requires the behavior of each team to solve the following equations simultaneously:

$$\alpha_n^* = \sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k} (p_n^*)^k (1 - p_n^*)^{n-k}, \qquad\qquad p_n^* = H(x - (x - y)\alpha_n^*).$$

If $x \neq y$, then p_n^* will generally depend on n, so there is an equilibrium effect in addition to the consensus effect. The equilibrium and consensus effects can go in opposite directions. For example, if $x > y$, the effect of increasing n is buffered by the countervailing effect that increasing α_n^* leads to a decrease in p_n^*. To see this formally, notice that the equilibrium condition for p_n^* depends on α_n^* according to $p_n^* = H(x-(x-y)\alpha_n^*)$ which is strictly decreasing in α_n^* precisely when $x > y$. Since $(x-(x-y)\alpha_n) > 0$ for all x, y, it is easy to see that as n grows large, $\lim_{n\to\infty}(\alpha_n^*) = 1$ and $\lim_{n\to\infty}(p_n^*) = H(y)$

When $x = y$ the analysis is straightforward and intuitive, because the expected payoff difference between Defect and Cooperate for either team is $U_D^t(\alpha_n^*) - U_C^t(\alpha_n^*) = x > 0$, which does not depend on α_n^* and hence is independent of n. In any team equilibrium, the probability that any team member votes for D is $H(x) > \frac{1}{2}$. Thus, in this special case, changes in the team equilibrium as n increases are entirely due to the consensus effect.

Team Equilibrium in PD games with different team sizes

Consider the case where $x > y$, the size of the row team is fixed at 1, and the size of the column team, n, is variable.[13] In this case, for any n the equilibrium is of the form $(q_n^*, \alpha_n^*, p_n^*)$, where q_n^* denotes the single row team member's defect probability, which depends on the of the column team's defect probability, α_n^*, and p_n^* denotes a column team member's probability of defection, which depends on q_n^*.

Thus, the equilibrium conditions are interdependent. Formally an equilibrium solves the following three equations:

$$q_n^* = H(x+(y-x)\alpha_n^*), \qquad \alpha_n^* = \sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k} (p_n^*)^k (1-p_n^*)^{n-k}, \qquad p_n^* = H(x+(y-x)q_n^*).$$

Since $x > y$, as noted above, the row team's defection probability, q_n^*, and the column team's defection probability, α_n^*, move in opposite directions. Similarly, q_n^* and p_n^* also move in opposite directions. Because the column team members' defection probability, p_n^* is strictly bounded above 0.5, α_n^* must eventually increase to 1 as n increases without bound, so the consensus effect eventually dominates the equilibrium effect. Similarly, the row team's defection probability, q_n^*, must eventually decrease to a limiting value of $H(y)$, and the column team members' defection probability, p_n^* eventually increases to a limiting value of $H(x + (y - x)H(y))$.[14]

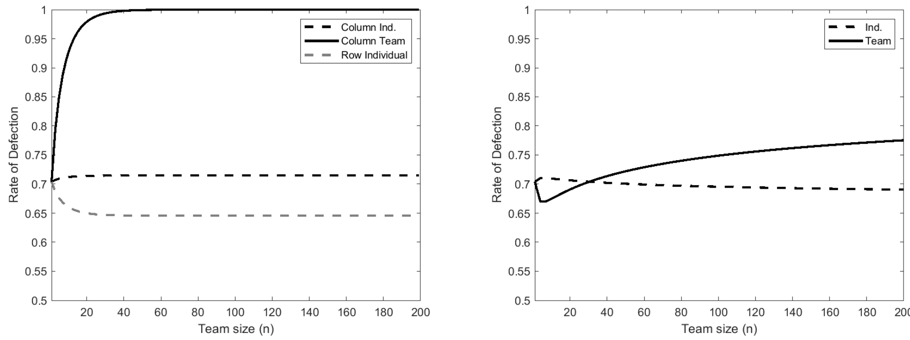
This is illustrated in the left panel of Figure 2.1, which displays the team equilibrium values, $(\alpha_n^*, q_n^*, p_n^*)$, as a function of n for $x = 5$, $y = 2$ and $H(z) = \frac{1}{1+e^{-.3z}}$, for column team sizes up to 200 (row team size fixed at 1).

Equilibrium with PD games with different voting rules

Even in games as simple as the PD game, the equilibrium effects in team games can be quite subtle. To illustrate this, we consider the same game as above, except

[13] This example also serves as an illustration where the two teams use different collective choice rules, with the row team using a dictatorial rule and the column team using simple majority rule.

[14] We conjecture that the three equilibrium values, α_n^*, q_n^*, and p_n^* each converge monotonically.



(a) Row team size = 1. Column = n. (b) 2/3 voting rule. Team sizes equal.

Figure 2.1: Team equilibrium in PD as a function of n (x=5, y=2).

where the teams are the same size and both teams use a supermajority voting rule: the team action is to defect if at least 2/3 of the team members favor defection; otherwise the team cooperates. In this case, the equilibrium conditions for a team equilibrium are:

$$\alpha_n^* = \sum_{k=\lceil \frac{2n}{3} \rceil}^{n} \binom{n}{k} (p_n^*)^k (1-p_n^*)^{n-k}, \qquad p_n^* = H(x+(y-x)\alpha_n^*),$$

where $\lceil \frac{2n}{3} \rceil$ denotes the least integer greater than or equal to $\frac{2n}{3}$. For the parameters used in the left panel of Figure 2.1 ($x = 5$, $y = 2$ and $H(z) = \frac{1}{1+e^{-.3z}}$), we get the surprising result that α_n^* does not converge to 1. The logic behind this is that if α_n^* did converge to 1 then in the limit we would have $p_n^* \to H(y) = \frac{1}{1+e^{-.3y}} \approx 0.65 < \frac{2}{3}$, which (given the $\frac{2}{3}$ rule) would imply that $\alpha_n^* \to 0$, a contradiction. A similar argument shows that α_n^* cannot converge to 0 either. For if α_n^* did converge to 0 then in the limit we would have $p_n^* \to H(x) = \frac{1}{1+e^{-.3x}} \approx 0.82 > \frac{2}{3}$, which would imply that $\alpha_n^* \to 1$, a contradiction. Hence, in the limit with large n a symmetric team equilibrium must converge to a *mixed* equilibrium! [15] The only way this can happen is if $p_n^* \to \frac{2}{3}$. Hence, the limiting equilibrium team probability of defection is calculated from the second equilibrium condition: $p_\infty^* = \frac{2}{3} = H(x-(x-y)\alpha_\infty^*)$, which gives $\alpha_\infty^* \approx 0.89$. The right panel of Figure 2.1 displays the team equilibrium (p_n^*, α_n^*) as a function of n for $x = 5$, $y = 2$ and $H(z) = \frac{1}{1+e^{-.3z}}$, for column team sizes up to 200. Convergence is much slower because the limiting team defect strategy is mixed.

[15] There are also asymmetric team equilibria that converge to pure strategies in the limit, where one team defects and the other team cooperates.

The Weak Prisoner's Dilemma

We next examine a class of team games where two teams, each of size n (odd), play a variation on the prisoner dilemma game displayed in Table 2.1, which we call the *Weak Prisoner's Dilemma*. While (D,D) is not a dominant strategy equilibrium, D is strictly dominant for the row player and is the unique solution to the game in two stages of iterated strict dominance. The column player is best off matching the row player's action choice. We characterize the team equilibrium for this class of games under majority rule. We have three free parameters to the game, which is displayed in Table 2.2. The first two, x and y, are the same as above, and for this example, we assume $x = y$. The third parameter, z, is the payoff gain the column player gets from cooperating if the row player cooperates.

Table 2.2: Weak Prisoner's Dilemma (WPD) game

	Column Team (2)	
Row Team (1)	Cooperate(C)	Defect(D)
Cooperate(C)	$5, 5$	$3 - x, 5 - z$
Defect(D)	$5 + x, 3 - x$	$3, 3$

Let $\alpha_n^* = (\alpha_n^{*1}, \alpha_n^{*2})$ be the team equilibrium probabilities of choosing D in the weak prisoner's dilemma game for row (team 1) and column (team 2). Following similar steps as in the PD example, we have for the row team:

$$\alpha_n^{*1} = \sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k} (p_n^{*1})^k (1 - p_n^{*1})^{n-k}, \qquad p_n^{*1} = H(x).$$

For a player on the column team, the probability of voting for D, p_n^{*2}, varies with n, as it depends directly on α_n^*. So the two equations for the column team are:

$$\alpha_n^{*2} = \sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k} (p_n^{*2})^k (1 - p_n^{*2})^{n-k}, \qquad p_n^{*2} = H(\alpha_n^{*1}(x + z) - z).$$

For any values of n, x, z, and H, a team equilibrium is given by any solution of this system of four equations.

Figure 2.2 illustrates how the equilibrium voting and team choice probabilities $(p_n^{*1}, \alpha_n^{*1}, p_n^{*2}, \alpha_n^{*2})$ vary with n, for the parameters $x = 1, z = 8$ and $H(z) = \frac{1}{1+e^{-.3z}}$ for team sizes up to 200. Notice that for relatively small teams sizes (< 20) the column team becomes *more* cooperative as it grows. This results from a combination

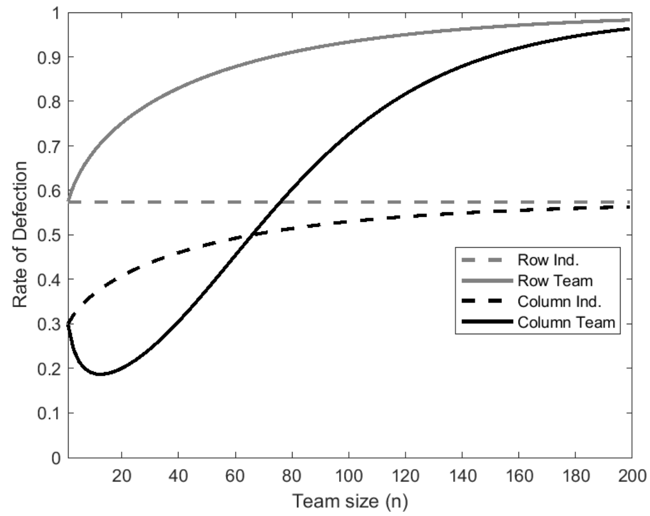


Figure 2.2: Team equilibrium in Weak Prisoner's Dilemma.

of the consensus effect (since $p_n^{*2} < .5$) and the equilibrium effect (since α_n^{*1} is increasing).

Asymmetric Matching Pennies Games

We examine a class of simple team games where two teams, each of size n (odd), play a 2×2 game with a unique (mixed-strategy) Nash equilibrium, which we refer to as asymmetric matching pennies (AMP) games. The payoffs are displayed in table 2.3, where $a > c, b < d, w < x, y > z$:

Table 2.3: Asymmetric matching pennies game

	Column L	Column R
Row U	a, w	b, x
Row D	c, y	d, z

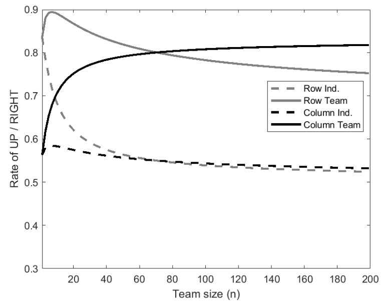
In a team equilibrium α_n^*, for any distribution of estimation errors, F, the probability that an individual player on the row team estimates that U is better than D is equal to:

$$p_n^{*1} = H(a(1-\alpha_n^{*2}) + b\alpha_n^{*2} - c(1-\alpha_n^{*2}) - d\alpha_n^{*2}) \tag{2.3}$$

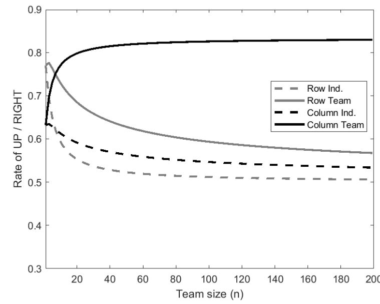
The probability that a column team member estimates that R is better than L is equal to:

$$p_n^{*2} = H(z(1-\alpha_n^{*1}) + x\alpha_n^{*1} - y(1-\alpha_n^{*1}) - w\alpha_n^{*1}) \tag{2.4}$$

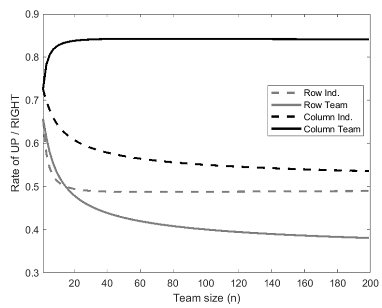
These are the voting probabilities. As in the previous examples, given p_n^{*1} and p_n^{*2} we can compute α_n^{*1} and α_n^{*2}, as the probability that at least $\frac{n+1}{2}$ members of the



(a) x=0.5



(b) x=1.0



(c) x=2.0

Figure 2.3: Team Equilibrium in Asymmetric Matching Pennies.

respective team estimate that U (R) yields a higher expected payoff than D (L). Hence:

$$\alpha_n^{*1} = \sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k} (p_n^{*1})^k (1 - p_n^{*1})^{n-k} \tag{2.5}$$

$$\alpha_n^{*2} = \sum_{k=\frac{n+1}{2}}^{n} \binom{n}{k} (p_n^{*2})^k (1 - p_n^{*2})^{n-k} \tag{2.6}$$

The team equilibrium is obtained by solving equations (2.3), (2.4), (2.5), and (2.6) simultaneously for p_n^{*1}, p_n^{*2}, α_n^{*1}, and α_n^{*2}. The team equilibrium and the equilibrium individual voting probabilities are displayed in Figure 2.3 for the parameters $w = b = c = z = 0$, $d = y = 1$, $a = 5$, $x = 0.5, 1.0, 2.0$, and $H(z) = \frac{1}{1+e^{-\cdot z}}$ for team sizes up to size 200.

From these examples, we see that $\lim_{n\to\infty} \alpha_n^* = (\frac{y-z}{y-z+x-w}, \frac{a-c}{a-c+d-b})$, the unique Nash equilibrium, in all three cases. As will be proved in Section 4, when teams use majority rule, convergent sequences of team equilibria must converge to a Nash equilibria, so this is a general feature of team equilibria of any 2×2 game with a unique mixed strategy equilibrium.

An interesting implication is that in the limit of team equilibria, *individual team members are voting randomly*. That is, $\lim_{n\to\infty} p_n^{*1} = \lim_{n\to\infty} p_n^{*2} = \frac{1}{2}$, which is necessarily the case because the expected payoffs of the two strategies are equal in the Nash equilibrium limit. The argument is similar to the earlier illustration of a mixed equilibrium in the PD game with a two-thirds voting rule. Suppose to the contrary that $\lim_{n\to\infty} p_n^{*1} > \frac{1}{2}$. Then $\lim_{n\to\infty} \alpha_n^{*1} = 1$, implying that the right hand side of equation 2.4 converges to $H(x-w) > \frac{1}{2}$ because $x - w > 0$, implying $\lim_{n\to\infty} p_n^{*2} > \frac{1}{2}$, so $\lim_{n\to\infty} \alpha_n^{*2} = 1$. This in turn implies that the right hand side of equation 2.3 converges to $H(b-d) < \frac{1}{2}$ because $b - d < 0$, and hence $\lim_{n\to\infty} p_n^{*1} < \frac{1}{2}$, a contradiction. A similar contradiction arises if one were to suppose instead that $\lim_{n\to\infty} p_n^{*1} < \frac{1}{2}$.

In this class of games, the equilibrium effect works in the opposite direction and thereby dampens the consensus effect, even in the limit. Without the equilibrium effect, which pushes p_n^* to exactly $\frac{1}{2}$ for both teams, the reinforcement effect by itself would lead to a pure strategy team actions in the limit. This is illustrated starkly in the examples shown in Figure 2.3, where *for all* n, $p_n^{*t} > \frac{1}{2}$ for both teams, yet $\lim_{n\to\infty} \alpha_n^* \neq (1,1)$.

2.4 Nash convergence

A collective choice rule satisfies the Nash convergence property if, as the size of all teams increases without bound, every convergent sequence of team equilibria converges to a Nash equilibrium. In the examples of the previous section, it was the case that the Nash convergence held if both teams used majority rule and the sizes of both teams increased without bound. This observation raises the more general question whether majority rule or 2×2 games are somehow unique in this regard, or if it is a property shared by a broader class of collective choice rules and a broader class of games. It is clearly not unique, as it is not difficult to show that the average rule also has this property. Thus the challenge is to characterize, at least partially, the class of collective choice rules and game environments with this *Nash convergence property.*

We show that all anonymous scoring rules satisfy the Nash convergence property.

Definition 2. *An* ***Individual Scoring Function*** $S_i^t : A^t \times \Re^{K^t} \to \Re^{K^t}$*, is a function defined such that* $S_{ik}^t(\hat{U}_i^t) = s_{im}$ *whenever* $|\{a_l^t \in A^t : \hat{U}_{il}^t > \hat{U}_{ik}^t\}| = m - 1$*, for some given set of* K^t *scores* $s_{i1}^t \geq s_{i2}^t \geq \ldots \geq s_{iK^t}^t \geq 0$ *such that* $s_{i1}^t > s_{iK^t}^t$*.*

Definition 3. *A team collective choice rule* C^t *is an* ***Anonymous Scoring Rule*** *if there exists a profile of individual scoring functions,* $(S_1^t, \ldots S_{n^t}^t)$ *with* $S_i^t = S_j^t = \sigma^t$

for all $i, j \in t$*, such that alternative* a_k^t *is chosen at* $\hat{U}^t$ *if and only if* $\sum_{i=1}^{n^t} \sigma_k^t(\hat{U}_i^t) \geq \sum_{i=1}^{n^t} \sigma_l^t(\hat{U}_i^t)$ *for all* $l \neq k$*.*

The individual scoring functions depend only on a team member's *ordinal* estimated expected utilities of the alternatives, and each alternative is awarded a score that is weakly increasing in its estimated expected utility rank by that team member. The individual scores for each team member are then summed to arrive at a total score for the team, and the alternatives with the highest total score are chosen. In an anonymous scoring rule, all members of the team have the same individual scoring function. Examples of common anonymous scoring rules include: plurality rule, where $s_{i1}^t = 1$ and $s_{im}^t = 0$ for all $i \in t$ and for all $m > 1$, and Borda count, where $s_{im} = K^t - m$ for all $i \in t$ and for all m. We note that in our framework, every member of every team almost always has a strict order over the K^t actions (i.e., $\hat{U}_{il}^t \neq \hat{U}_{ik}^t$ for all l, k, t, i with probability one), so ties in an individual member's ordinal rankings are irrelevant.

Theorem 2. *Consider an infinite sequence of team games,* $\{\Gamma_m\}_{m=1}^{\infty}$ *such that (1)* $A_m^t = A_{m'}^t = A^t$ *for all* t, m, m'*; (2)* $u_m^t = u_{m'}^t = u^t\ \forall t, m, m'$*; (3)* $n_{m+1}^t > n_m^t$ *for all* m, t*; and (4)* $C_m^t = \sigma$*, an anonymous scoring rule, for all* m, t*. Let* $\{\alpha_m\}_{m=1}^{\infty}$ *be a convergent sequence of team equilibria where* $\lim_{m\to\infty} \alpha_m = \alpha^*$*. Then* α^* *is a Nash equilibrium of the strategic form game* $[\mathbf{T}, A, u]$*.*

Proof. Suppose α^* is not a Nash equilibrium. Then there is some team t and some pair of actions, a_k^t, a_l^t such that $U_k^t(\alpha^*) > U_l^t(\alpha^*)$, but $\alpha_l^{t*} = \xi > 0$. Since C^t is an anonymous scoring rule, σ, we know that for all t, for all n^t, for all α,and for all $a_k^t \in A^t$, $a_k^t \in C^t(\hat{U}^t(\alpha))$ if and only if

$$\begin{aligned} \sum_{i=1}^{n^t} \sigma_k(\hat{U}_i^t(\alpha)) &\geq \sum_{i=1}^{n^t} \sigma_l(\hat{U}_i^t(\alpha)) \text{ for all } l \neq k \\ &\Leftrightarrow \\ \overline{\sigma}_{kn^t}^t(\hat{U}^t(\alpha)) &\geq \overline{\sigma}_{ln^t}^t(\hat{U}^t(\alpha)) \text{ for all } l \neq k. \end{aligned}$$

where $\overline{\sigma}_{kn^t}^t(\hat{U}^t(\alpha))$ denotes the average score of a_k^t among the n^t members of t at $\hat{U}^t(\alpha)$. That is, the score of a_k^t is maximal if and only if the average individual score of a_k^t is maximal. If $U_k^t(\alpha) > U_l^t(\alpha)$ then σ_k stochastically dominates σ_l, and hence their respective expected scores are strictly ordered. That is, $E\{\sigma_k(\hat{U}_i^t(\alpha))\} > E\{\sigma_l(\hat{U}_i^t(\alpha))\}$, where

$$E\{\sigma_k(\hat{U}_i^t(\alpha))\} = \int \sigma_k(U^t(\alpha) + \epsilon_i^t) dF^t(\epsilon_i^t)$$

Hence, at the limiting strategy profile, α^*, we have $E\{\sigma_k(\hat{U}_i^t(\alpha^*))\} > E\{\sigma_l(\hat{U}_i^t(\alpha^*))\}$. Since $U^t(\alpha_m) \to U^t(\alpha^*)$ and $E\{\sigma\}$ is continuous in α, $\overline{\sigma}^t_{kn^t_m}(\hat{U}^t(\alpha_m)) \to E\{\sigma^t_k(\hat{U}^t(\alpha^*))\}$ and $\overline{\sigma}^t_{ln^t_m}(\hat{U}^t(\alpha_m)) \to E\{\sigma^t_l(\hat{U}^t(\alpha^*))\}$ in probability as $m \to \infty$. Therefore, $E\{\sigma_k(\hat{U}_i^t(\alpha^*))\} > E\{\sigma_l(\hat{U}_i^t(\alpha^*))\}$ implies $\exists\ \overline{m}$ such that $\Pr\{\overline{\sigma}^t_{kn^t_m}(\hat{U}^t(\alpha_m)) \leq \overline{\sigma}^t_{ln^t_m}(\hat{U}^t(\alpha_m))\} < \frac{\xi}{2}\ \forall m > \overline{m}$. This leads to a contradiction to the initial hypothesis that $\alpha_l^{t*} = \xi > 0$ since $\Pr\{\overline{\sigma}^t_{kn^t_m}(\hat{U}^t(\alpha_m)) \leq \overline{\sigma}^t_{ln^t_m}(\hat{U}^t(\alpha_m))\} < \frac{\xi}{2}\ \forall m > \overline{m}$, implies that $\alpha^t_{ln^t_m} < \frac{\xi}{2}\ \forall m > \overline{m}$ and hence $\alpha_l^{t*} < \xi$. □

It is useful to clarify the implications and generality of the result with a few comments. First, the theorem does not imply that team equilibrium with larger teams is necessarily closer to Nash equilibrium than equilibrium with smaller teams. It is an asymptotic result for large teams. Examples in the last section show that the convergence can be non-monotonic even in very simple games.

Second, Nash convergence is an upper hemicontinuity property of the team equilibrium correspondence, but that correspondence is not generally lower hemicontinuous. Some Nash equilibria are not approachable, just as some weak Nash equilibria fail to be limit points of payoff disturbed games (Harsanyi (1973)) or limit points of quantal response equilibria as the error terms vanish (McKelvey and Palfrey (1995)). In the following game, (D,R) is a Nash equilibrium that cannot be approached by a sequence of large team equilibria. It is easy to see that in any team equilibrium,

Table 2.4: Example of a Nash equilibrium that is not a limit of team equilibria with majority rule.

	Column Team (2)	
Row Team (1)	Left (L)	Right (R)
Up (U)	1, 1	0, 0
Down (D)	0, 0	0, 0

for any n, the probability the row team plays Up and the probability the column team plays Left is always greater than 0.5. Hence (D, R) cannot be a limit of team equilibria.

Third, the scoring rule does not have to be the same for all teams in order to obtain Nash convergence; different teams can use different anonymous scoring rules and the result still holds. While not formally stated in the theorem, it is an obvious generalization. Fourth, the rate of convergence to large teams can differ across teams. Fifth, the result only characterizes conditions that are sufficient for the Nash convergence property; the condition is clearly not necessary as the average rule is not a scoring rule but satisfies the Nash convergence property. Finally, we conjecture that the class is much broader than scoring rules, including many other anonymous and neutral collective choice rules. Scoring rules operate only on the individual *ordinal* rankings of estimated expected payoffs. One imagines that there are many collective choice rules that operate on the cardinal values of the estimates and also have the Nash convergence property, such as weighted average rules.

2.5 Stochastic Rationality and Team Response Functions

In our framework, individual team member estimated expected utilities obey two intuitive, normatively appealing properties of stochastic rationality. First, the probability a member of team t ranks action a_k^t as having the highest estimated expected payoff (and hence would be chosen if it were an individual choice problem) is increasing in $U_k^t(\alpha)$, the 'true' equilibrium expected payoff of action k, ceteris paribus, a condition we call *payoff monotonicity*. Second, the probability a member of team t ranks action a_k^t's estimated expected payoff as highest is greater than the probability the member ranks action a_l^t's highest if and only if $U_k^t(\alpha) > U_l^t(\alpha)$, a condition we call *rank dependence*.[16] In the context of team decision making, it is the collective choice rule in combination with the error structure that determines team choice probabilities. It is easy to see that team decision making will not generally inherit these two properties for all collective choice rules. Given the normative appeal of these two properties, this naturally leads to the following question. Under what conditions on the collective choice rule will team response functions satisfy them? In addition to the normative appeal of payoff monotonicity and rank dependence, violations of these two properties might affect the incentives faced by team members during the team decision making process. While our framework does not explicitly model individual choice behavior in the team decision making process, it is plausible that stochastically irrational collective choice rules that violate payoff monotonicity or rank dependence, could hinder the ability for teams to effectively aggregate members' diverse beliefs. For example, if payoff monotonicity fails for the prescribed collective choice rule, then individual team members might profit by behaving as if their expected utility estimates for some of the actions are lower or higher than they are in truth.

This section identifies restrictions on the collective choice rule that guarantee team

[16] This follows from the i.i.d. assumption on estimation errors. Furthermore, the estimated expected payoffs are continuous in U^t and have full support on the real line. Thus, they have properties similar to the individual choice probabilities in a quantal response equilibrium.

response functions to satisfy these two properties. We first show that payoff monotonicity of team response functions requires only two weak assumptions on the collective choice rule, unanimity and positive responsiveness. On the other hand, rank dependence holds only for a more restricted class of neutral collective choice rules. Many non-neutral collective choice rules, such as those that give a status quo advantage to an action, will fail to satisfy rank dependence, as the last example in section 3.1 demonstrates. Second, we show that rank dependence is satisfied for $K^t = 2$ with any collective choice rule satisfying unanimity, positive responsiveness, and neutrality, and for $K^t > 2$ with plurality rule or weighted average rules.

Payoff Monotonicity

For any team game, Γ, P^{C^t} depends on the strategy profile α, the distribution of member's estimation errors, F^t, and the team collective choice rule, C^t. In this section we identify conditions on C^t that are sufficient for P^{C^t} to be payoff monotone for all admissible F^t. The formal definition of payoff monotonicity is given below.

Definition 4. *A team collective choice rule C^t satisfies **Payoff Monotonicity** if, for all a^t_k, α, α':*

$$U^t_k(\alpha) > U^t_k(\alpha') \text{ and } U^t_l(\alpha) = U^t_l(\alpha') \; \forall l \neq k \Rightarrow$$
$$P^{C^t}_k(U^t(\alpha)) > P^{C^t}_k(U^t(\alpha')).$$

Specifically, we require team collective choice rules to satisfy two axioms: unanimity and positive responsiveness. The first condition, unanimity, simply states that if all members of the team estimate that a^t_k has the highest expected utility, then it is uniquely chosen by C^t.[17]

Definition 5. *A team collective choice rule C^t satisfies **Unanimity** if:*

$$\widehat{U}^t_{ik} > \widehat{U}^t_{il} \text{ for all } i \in t \text{ and for all } l \neq k \Rightarrow C^t(\widehat{U}^t) = \{a^t_k\}.$$

In addition to using this axiom to prove payoff monotonicity, it also guarantees that team response functions are interior, in the sense that every action is chosen with positive probability. The second axiom, positive responsiveness, requires that the team choice responds positively to all members of a team increasing their estimated expected payoff of an action, keeping all other estimated expected payoffs the same. The following definition is used in the statement of the axiom.

Definition 6. *A profile $\tilde{U}^t$ of member estimated expected utilities is **a monotonic transformation of $\hat{U}^t$ with respect to action** a^t_k if, for all members $i \in t$, we have $\tilde{U}^t_{ik} \geq \hat{U}^t_{ik}$ and $\tilde{U}^t_{il} = \hat{U}^t_{il}$ for all $l \neq k$.*

[17] For the "standard" case of games played by one-person teams, unanimity implies that if $n = 1$ then every team equilibrium is equivalent to a quantal response equilibrium of the strategic form game, $[T, A, u]$.

Definition 7. *A team collective choice rule C^t satisfies* ***Positive Responsiveness*** *if $a_k^t \in C^t(\hat{U}^t) \Rightarrow a_k^t \in C^t(\tilde{U}^t) \subseteq C^t(\hat{U}^t)$, for all a_k^t, $\hat{U}^t$ and all monotonic transformations $\tilde{U}^t$ of $\hat{U}^t$ with respect to action a_k^t.*

This definition of Positive Responsiveness is essentially a cardinal version of the usual definition of positive responsiveness from the social choice literature. It says that if an action a_k^t is chosen at some profile of estimated expected utilities, and all team members' estimates of the expected utility of that action weakly increase, ceteris paribus, then a_k^t must still be chosen, and no new actions can be added to the choice set.

Many collective choice rules satisfy positive responsiveness. For example, any weighted average rule, where the team choice corresponds to the action with the highest weighted average of individual members estimates, is positively responsive. Plurality rule also clearly satisfies this condition. In this section we consider a class of collective choice rules, called *generalized* scoring rules, and show that positive responsiveness is satisfied for any such collective choice rule. A generalized scoring rule is substantially more general than the standard definition of a scoring rule in the social choice literature, which was defined in the previous section as an *anonymous* scoring rules (i.e., all the individual scoring functions are the same). Generalized scoring rules relax the anonymity requirement that all individual scoring functions are the same. It includes a wide range of non-anonymous collective choice rules, including dictatorial rules.

Definition 8. *A team collective choice rule C^t is a* ***Generalized Scoring Rule*** *if there exists a profile of individual scoring functions, $(S_1^t, \ldots S_{n^t}^t)$, such that for all $a_k^t \in A^t$ and for all $\hat{U}^t \in \Re^{K^t n^t}$, $a_k^t \in C^t(\hat{U}^t)$ if and only if $\sum_{i=1}^{n^t} S_{ik}^t(\hat{U}_i^t) \geq \sum_{i=1}^{n^t} S_{il}^t(\hat{U}_i^t)$ for all $l \neq k$.*

Proposition 1. *All generalized scoring rules satisfy positive responsiveness.* [18]

Proof. Positive responsiveness follows from the fact that the value of the team score function evaluated at any alternative is weakly increasing in that alternative's estimated expected utility for each team member, and weakly decreasing in every other alternative's estimated expected utility. □

We can now state the main result of this subsection.

[18] If $s_{i1}^t > s_{i2}^t$ for all $i \in t$, then the scoring rule also satisfies unanimity.

Theorem 3. *If F^t is admissible and C^t satisfies unanimity and positive responsiveness then P^{C^t} satisfies payoff monotonicity.*

Proof. Let C^t satisfy positive responsiveness and unanimity and F^t admissible. Suppose that $U_k^t - U_k^{'t} = \delta > 0$, and $U_l^t = U_l^{'t}, \forall l \neq k$. Then for all realizations of the estimation errors ϵ^t, we have that $U^t + \epsilon^t$ is a monotonic transformation of $U^{'t} + \epsilon^t$ with respect to a_k^t. So by positive responsiveness of C^t we have that if $a_k^t \in C^t(U^{'t} + \epsilon^t)$ then $a_k^t \in C^t(U^t + \epsilon^t)$, and if $a_l^t \in C^t(U^t + \epsilon^t)$ then $a_l^t \in C^t(U^{'t} + \epsilon^t)$. So $g_k^{C^t}(U^t + \epsilon^t) \geq g_k^{C^t}(U^{'t} + \epsilon^t)$ for all ϵ^t, and therefore $P_k^{C^t}(U^t) \geq P_k^{C^t}(U^{'t})$. To show the strict inequality, $P_k^{C^t}(U^t) > P_k^{C^t}(U^{'t})$, we show that there exists a region $\beta \subset \Re^{K^t \times N^t}$ with positive measure such that if $\epsilon^t \in \beta$, then $g_k^{C^t}(U^t + \epsilon^t) > g_k^{C^t}(U^{'t} + \epsilon^t)$. In particular, unanimity of C^t is used as follows to construct β such that if $\epsilon^t \in \beta$, then $g_k^{C^t}(U^t + \epsilon^t) = 1 > g_k^{C^t}(U^{'t} + \epsilon^t) = 0$. That is, such that a_k^t is uniquely chosen under $U^t + \epsilon^t$, and not chosen under $U^{'t} + \epsilon^t$. Let $\tilde{U}^t$ be an estimated expected utility profile such that all team members strictly prefer some action a_l^t to action a_k^t, all members prefer a_k^t to all other actions a_m^t (i.e., all members rank a_k^t second), and for all members we have $\tilde{U}_l^t - \tilde{U}_k^t = \frac{\delta}{2}$. Define:

$$\beta = \{\tilde{U}^t - U^{'t} + \xi : \xi_k \in (0, \frac{\delta}{4}), \xi_l \in (-\frac{\delta}{4}, 0), \xi_m < 0\}$$

Then if $\epsilon^t \in \beta$, by unanimity we have $C^t(U^{'t} + \epsilon^t) = \{a_l^t\}$ and $C^t(U^t + \epsilon^t) = \{a_k^t\}$, so $g_k^{C^t}(U^t + \epsilon^t) = 1 > g_k^{C^t}(U^{'t} + \epsilon^t) = 0$. β is an open set and hence has positive measure since the distribution of ϵ^t has full support. Therefore $P_k^{C^t}(U^t) > P_k^{C^t}(U^{'t})$, as desired. □

Rank Dependence

In this section we show that P^{C^t} satisfies rank dependence for $K^t = 2$ with any collective choice rule satisfying unanimity, positive responsiveness and neutrality, and for $K^t > 2$ with plurality rule and weighted average rules. The formal definition of rank dependence is:

Definition 9. *A team collective choice rule C^t satisfies* ***Rank Dependence*** *if, for all $a_k^t, a_l^t, \alpha, U_k^t(\alpha) > U_l^t(\alpha) \Rightarrow P_k^{C^t}(U^t(\alpha)) > P_l^{C^t}(U^t(\alpha))$*

Neutrality is an essential property for proving that team response functions satisfy rank dependence. Informally a neutral team collective choice rule is one that is not biased against or in favor of any particular action. This is analogous to the neutrality axiom from the social choice literature.

Let $\psi : A^t \rightarrow A^t$ be any permutation of team actions. Denote by $U^{t,\psi} = (U^t_{\psi(1)}, ..., U^t_{\psi(K^t)})$ the permuted profile of expected utilities and by $\hat{U}^{t,\psi} \equiv (U^t_{\psi(1)} + \epsilon^t_{\psi(1)}, ..., U^t_{\psi(K^t)} + \epsilon^t_{\psi(K^t)})$ the permuted profile of estimated expected utilities. We can then define neutrality formally.

Definition 10. *A team collective choice rule C^t satisfies* ***Neutrality*** *if, for all a^t_k, $\hat{U}^t$, for all permutations ψ, $a^t_k \in C^t(\hat{U}^t) \Leftrightarrow a^t_{\psi(k)} \in C^t(\hat{U}^{t,\psi})$*

Neutrality, along with admissibility of F^t, imply that when the expected payoffs of team actions are permuted, the team choice probabilities are permuted.

Lemma 1. *If F^t is admissible and C^t satisfies neutrality, then $P^{C^t}_k(U^t) = P^{C^t}_{\psi(k)}(U^{t,\psi})$ for all expected utility profiles, U^t, actions, a^t_k, and permutations ψ.*

Proof. By neutrality of C^t, for any expected utility profile, U^t, action, a^t_k, belief error profile, ϵ^t, and permutation ψ, $g^{C^t}_k(\hat{U}^t) = g^{C^t}_{\psi(k)}(\hat{U}^{t,\psi})$, that is the probability that a^t_k is chosen at $\hat{U}^t$ is equal to the probability that $a^t_{\psi(k)}$ is chosen at $\hat{U}^{t,\psi}$. Therefore $P^{C^t}_k(U^t) = \int_{\epsilon^t} g^{C^t}_k(\hat{U}^t)dF^t(\epsilon^t) = \int_{\epsilon^t} g^{C^t}_{\psi(k)}(\hat{U}^{t,\psi})dF^t(\epsilon^t)$. Finally, since the estimation errors are i.i.d, $\int_{\epsilon^t} g^{C^t}_{\psi(k)}(\hat{U}^{t,\psi})dF^t(\epsilon^t) = \int_{\epsilon^t} g^{C^t}_{\psi(k)}(\hat{U}^{t,\psi})dF^t(\epsilon^{t,\psi}) = P^{C^t}_{\psi(k)}(U^{t,\psi})$. □

A corollary to the lemma is that when two actions have equal expected payoffs, the team must play these actions with equal probability. It is easy to see that non-neutral collective choice rules can lead to violations of rank dependence. For example, collective choice rules that favor one action (e.g. a status quo action) over another will generally lead to violations, as in the last example of Section 3.1 with a 2/3 voting rule. Consider $K^t = 2$ and a choice rule that selects action a^t_1 if and only if all team members estimate its expected utility to be greater than that of action a^t_2, and selects action a^t_2 otherwise. For any admissible F^t, if the size of the team is large enough, a team using this choice rule will select action a^t_2 more often than action a^t_1 even when $U^t_1(\alpha) > U^t_2(\alpha)$.

Next, for the case of $K^t = 2$, we prove that neutrality, together with unanimity and positive responsiveness is sufficient to guarantee that a team response function satisfies rank dependence for all admissible F^t. This is proved below.

Theorem 4. *If $K^t = 2$, F^t is admissible and C^t satisfies unanimity, positive responsiveness and neutrality, then P^t satisfies rank dependence.*

Proof. Pick any U^t such that $U_1^t > U_2^t$ and let $\delta = U_1^t - U_2^t$. Let $U'^t = (U_1^t - \delta, U_2^t)$, then by lemma 1, $P_1^{C^t}(U'^t) = P_2^{C^t}(U'^t) = \frac{1}{2}$. Since C^t satisfies positive responsiveness and unanimity, theorem 3, together with the fact that $P_1^{C^t}(U^t) + P_2^{C^t}(U^t) = 1$, implies that $P_1^{C^t}(U^t) > P_1^{C^t}(U'^t) = P_2^{C^t}(U'^t) > P_2^{C^t}(U^t)$. □

If $K^t > 2$ the next two propositions prove rank dependence with additional restrictions on the team collective choice rule.

Theorem 5. *If F^t is admissible and C^t is plurality rule, then P^{C^t} satisfies rank dependence.*

Proof. Consider any profile of expected payoffs U^t. By neutrality and admissibility we can without loss of generality label the actions such that $U_1^t \geq U_2^t \geq ... \geq U_{K^t}^t$. By lemma 1, if $U_k^t = U_l^t$, then $P_k^{C^t} = P_l^{C^t}$.

Suppose $U_k^t > U_l^t$. The probability that any team member i ranks action k highest is $p_k = Prob(U_k^t + \epsilon_{ik}^t - \max_{j \neq k}\{U_j^t + \epsilon_{ij}^t\} \geq 0)$. Let $\psi : \{1, ..., K^t\} \to \{1, ..., K^t\}$ be the pairwise permutation of k and l, that is the permutation that maps k to l and l to k and all else to itself. By exchangeability of the error terms, $((U_1^t + \epsilon_1^t, ..., U_{K^t}^t + \epsilon_{K^t}^t)$ has the same joint distribution as $(U_1^t + \epsilon_{\psi(1)}^t, ..., U_{K^t}^t + \epsilon_{\psi(K^t)}^t)$, and so $U_k^t + \epsilon_{ik}^t - \max_{j \neq k}\{U_j^t + \epsilon_{ij}^t\}$ has the same distribution as $U_k^t + \epsilon_{i\psi(k)}^t - \max_{j \neq k}\{U_j^t + \epsilon_{i\psi(j)}^t\}$. Since $U_k^t > U_l^t$, we have for all ϵ_i^t, $U_{\psi(k)}^t + \epsilon_{i\psi(k)}^t < U_k^t + \epsilon_{i\psi(k)}^t$, and $\max_{j \neq k}\{U_{\psi(j)}^t + \epsilon_{i\psi(j)}^t\} \geq \max_{j \neq k}\{U_j^t + \epsilon_{i\psi(j)}^t\}$. By the full support assumption, it follows that $Prob(U_k^t + \epsilon_{i\psi(k)}^t - \max_{j \neq k}\{U_j^t + \epsilon_{i\psi(j)}^t\} \geq 0) > Prob(U_{\psi(k)}^t + \epsilon_{i\psi(k)}^t - \max_{j \neq k}\{U_j^t + \epsilon_{i\psi(j)}^t\} \geq 0) = p_l$, so $p_k > p_l$.

Now, still supposing that $U_k^t > U_l^t$, and therefore that $p_k > p_l$, denote by $(n_1, ..., n_K)$ the tuple of number of team members that rank each action first for a given $\hat{U}^t$. For any choice set $B \subseteq A^t$, let $V^B = \{(n_1, ..., n_K) | \forall a_k \in B, \forall a_j, n_k \geq n_j \text{ and } \sum_{j=1}^K n_j = n\}$ be the set of feasible 'vote' totals that result in B being chosen. Then, since estimated expected utilities are independent across individuals conditional on U^t, we can write the probability of this subset being chosen as

$$Prob(C^t(\hat{U}^t) = B) = \sum_{V^B} \frac{n!}{n_1! n_2! ... n_K!} \Pi_{j=1}^K p_j^{n_j}$$

Let $\psi : \{1, ..., K^t\} \to \{1, ..., K^t\}$ be the pairwise permutation between k and l as defined earlier. Pick any B that contains a_k and not a_l. Then $(n_1, ..., n_K) \in V^B$

if and only if $(n_{\psi(1)},, ..., n_{\psi(K)}) \in V^{(B-\{a_k\})\cup\{a_l\}}$, the set of vote totals that results in the choice set being B, minus a_k and adding a_l. Then, since $n_k > n_l$, we have that $p_k^{n_k} p_l^{n_l} > p_k^{n_l} p_l^{n_k}$, so every term of the sum in $Prob(C^t(\hat{U}^t) = B)$ is greater than the corresponding term in $Prob(C^t(\hat{U}^t) = (B - \{a_k\}) \cup \{a_l\})$. So we have for all B containing a_k and not a_l, that $Prob(C^t(\hat{U}^t) = B) > Prob(C^t(\hat{U}^t) = (B - \{a_k\}) \cup \{a_l\})$. Finally, define B_0 to be the subsets of A^t that contain neither a_k nor a_l, B_k the subsets containing only a_k, B_l the subsets containing only a_l and not a_k, and B_{kl} the set containing both. Then

$$
\begin{aligned}
P_k^{C^t}(U^t) &= 0 \times \sum_{B \in B_0} Prob(C^t(\hat{U}^t) = B) + \sum_{B \in B_k} \frac{1}{|B|} Prob(C^t(\hat{U}^t) = B) \\
&\quad + 0 \times \sum_{B \in B_l} Prob(C^t(\hat{U}^t) = B) + \sum_{B \in B_{kl}} \frac{1}{|B|} Prob(C^t(\hat{U}^t) = B) \\
P_l^{C^t}(U^t) &= 0 \times \sum_{B \in B_0} Prob(C^t(\hat{U}^t) = B) + 0 \times \sum_{B \in B_k} Prob(C^t(\hat{U}^t) = B) \\
&\quad + \sum_{B \in B_l} \frac{1}{|B|} Prob(C^t(\hat{U}^t) = B) + \sum_{B \in B_{kl}} \frac{1}{|B|} Prob(C^t(\hat{U}^t) = B)
\end{aligned}
$$

$P_k^{C^t}$ and $P_l^{C^t}$ share all terms of the fourth sum, so

$$
P_k^{C^t}[U^t] - P_l^{C^t}[U^t] = \sum_{B \in B_k} \frac{1}{|B|} Prob[C^t(\hat{U}^t) = B] - \sum_{B \in B_l} \frac{1}{|B|} Prob[C^t(\hat{U}^t) = B]
$$

$$
P_k^{C^t}[U^t] - P_l^{C^t}[U^t] = \sum_{B \in B_k} \frac{1}{|B|} [Prob[C^t(\hat{U}^t) = B] - Prob[C^t(\hat{U}^t) = (B - \{a_k\}) \cup \{a_l\}]] > 0
$$

Therefore, whenever $U_k^t > U_l^t$, we have $p_k > p_l$, which implies $P_k^{C^t}(U^t) > P_l^{C^t}(U^t)$.

□

Define a weighted average rule as follows:

Definition 11. *A team collective choice rule C^t is a **Weighted Average Rule** if there exists a profile of non-negative individual voting weights, $(w_1^t, ..., w_{n^t}^t)$ with $\sum_{i=1}^{n^t} w_i^t = 1$ such that for all $a_k^t \in A^t$ and for all $\hat{U}^t \in \Re^{K^t n^t}$, $a_k^t \in C^t(\hat{U}^t)$ if and only if $\sum_{i=1}^{n^t} w_i^t \hat{U}_{ik}^t \geq \sum_{i=1}^{n^t} w_i^t \hat{U}_{il}^t$ for all $l \neq k$.*

Theorem 6. *If F^t is admissible and C^t is a weighted average rule, then P^{C^t} satisfies rank dependence.*

Proof. Consider any profile of expected payoffs U^t, and suppose $U_k^t > U_l^t$. We have $P_k^{C^t}(U^t) = \int \mathbb{1}\{\sum_{i=1}^{n^t} w_i^t \hat{U}_{ik}^t \geq \max\{\sum_{i=1}^{n^t} w_i \hat{U}_{ij}^t\}_{j=1}^{K^t}\} dF^t$. Note that the probability that any of these weighted averages are exactly equal is 0. Now, since

$F^t(y - U_k^t) < F^t(y - U_l^t)$ for all $y \in \Re$, we have $\hat{U}_{ik}^t >_{st} \hat{U}_{il}^t$, where $>_{st}$ denotes the *strict* first stochastic order, for all members i. This order is closed under convolutions, so $\sum_{i=1}^{n^t} w_i^t \hat{U}_{ik}^t >_{st} \sum_{i=1}^{n^t} w_i^t \hat{U}_{il}^t$. Since $\mathbb{1}\{z > 0\}$ is increasing, non-constant and bounded, we therefore have

$$\int \mathbb{1}\{\sum_{i=1}^{n} w_i \hat{U}_{ik}^t \geq \max\{\sum_{i=1}^{n} w_i \hat{U}_{ij}^t\}_{j=1}^K\}\}dF > \int \mathbb{1}\{\sum_{i=1}^{n} w_i \hat{U}_{il}^t \geq \max\{\sum_{i=1}^{n} w_i \hat{U}_{ij}^t\}_{j=1}^K\}\}dF$$

$$P_k^{C^t}(U^t) > P_l^{C^t}(U^t)$$

□

2.6 Team Equilibrium in Extensive Form Games

A finite extensive form game consists of a Player set $\mathbf{I} = \{1, ..., I\}$, an action set $\mathbf{A}$, a set of sequences contained in A called histories, Ξ, a subset of these being terminal histories, Z, initial chance moves, b^0, a player function, ι, information sets for each player, Π^i, a feasible action function, A, that specifies the set of actions available at each information set, and payoff functions, $u = (u^1, ..., u^i, ..., u^I)$, defined on Z. Thus an extensive form game, in shorthand, can be written as $G^{EF} = (\mathbf{I}, \mathbf{A}, \Xi, A, \iota, \Pi, b^0, u)$.

We extend this definition to (finite) *team extensive form games*, modifying the notation in the following ways. As in the definition of team games in strategic form, let $\mathbf{T} = \{0, 1, ..., t, ..., T\}$ be a finite collection of *teams*, $t = \{i_1^t, ..., i_j^t, ..., i_{n^t}^t\}$, where each team consists of n^t members, or individuals, and denote the *team size profile* by $n = (n^1, ..., n^T)$. Team "0" is designated *chance*. Let $\mathbf{A}$ be a finite set of *actions* and Ξ be a finite set of *histories*, that satisfy two properties: $\emptyset \in \Xi$; and $(a_1, ..., a_K) \in \Xi \Rightarrow (a_1, ..., a_L) \in \Xi$ for all $L < K$. A history $h = (a_1, ..., a_K) \in Z \subseteq \Xi$ is *terminal* if there does not exist $a \in \mathbf{A}$ such that $(h, a) \in \Xi$, and the set of terminal histories is denoted Z. The set of actions available at any non-terminal history h is determined by the function $A : \Xi \rightarrow 2^{\mathbf{A}}$, where $A(h) = \{a | (h, a) \in \Xi\}$. There is a *team function* $\iota : \Xi - Z \rightarrow \mathbf{T}$ that assigns each history to a unique team. Without loss of generality, assume that $\iota(\emptyset) = 0$ and $\iota(h) \neq 0$ for all $h \neq \emptyset$, and denote by b^0 the probability distribution of chance actions at $h = \emptyset$ and assume without loss of generality that $b^0(a) > 0$ for all $a \in A(\emptyset)$. For each $t \in \{0, 1, ..., t, ..., T\}$, there is an information partition Π^t of $\{h \in \Xi | \iota(h) = t\}$. Elements of Π^t are t's *information sets*, and are denoted H_l^t, where l indexes t's information sets. The set of available actions to t are the same in all histories that belong to the same information set. That is, if $h \in H_l^t$ and $h' \in H_l^t$ for some $H_l^t \in \Pi_t$ then $A(h) = A(h') \equiv A(H_l^t) = \{a_{l1}^t, ..., a_{lk}^t, ..., a_{lK_l^t}^t\}$ where $K_l^t = |A(H_l^t)|$.

The *payoff function* of the game for team $t \neq 0$ is given by $u^t : Z \rightarrow \Re$. Given any terminal history $z \in Z$, all members of team t receive the payoff $u^t(z)$. A behavioral strategy for team t is a function $b^t = (b_1^t, ..., b_l^t, ..., b_{L^t}^t)$, where $L^t = |\Pi^t|$ and $b_j^t : H_l^t \rightarrow \Delta A(H_l^t)$, where $\Delta A(H_l^t)$ is the set of probability distributions over $A(H_l^t)$. Denote by B the set of behavioral strategy profiles, and B^o the interior of B, i.e., the set of totally mixed behavioral strategy profiles.

Each behavioral strategy profile $b \in B^o$ determines a strictly positive *realization probability* $\rho(z|b)$ for each $z \in Z$. For any $t \neq 0$ and $b \in B^o$, define the expected payoff function for team t, $v_i : B^o \rightarrow \mathcal{R}$ by:

$$v^t(b) = \sum_{z \in Z} \rho(z|b) u_t(z).$$

Similarly, for any $t \neq 0$ and any information set $H_l^t \in \Pi^t$, and for each $a_{lk} \in A(H_l^t)$, any $b \in B^o$ determines a strictly positive *conditional realization probability* $\rho(z|H_l^t, b, a_{lk})$ for each $z \in Z$.[19] This is the probability distribution over Z, conditional on reaching H_l^t given the behavioral strategy profile b, with b_l^t replaced with the pure action a_{lk}. For each $t \neq 0$, for each $H_l^t \in \Pi^t$, and for each $a_{lk} \in A(H_l^t)$, define the *conditional payoff function* by

$$U_{lk}^t(b) = \sum_{z \in Z} \rho(z|H_l^t, b, a_{lk}) u^t(z).$$

This is the conditional payoff to t of playing the (pure) action $a_{lk} \in A(H_l^t)$ at H_l^t with probability one, and otherwise all teams (including t) playing b elsewhere.

The rest of the formal description of extensive form team games closely follows the structure of team games in strategic form. At each information set H_l^t, each member of team t gets a noisy estimate of the *true* conditional payoff of each currently available action, given a behavioral strategy profile of the other teams, b^{-t}. These estimates are aggregated into a team decision via a team collective choice rule, C_l^t.

Formally, given $b \in B^o$, for every t, and each of t's information sets, $H_l^t \in \Pi^t$, at any history in H_l^t, for each $a_{lk} \in A(H_l^t)$ member $i \in t$ observes an estimate of $U_{lk}^t(b)$ equal to the true conditional expected payoff, $U_{lk}^t(b)$, plus an estimation error term. That is, $\widehat{U}_{ilk}^t = U_{lk}^t(b) + \varepsilon_{ilk}^t$, where the dependence of $\widehat{U}_{ilk}^t$ on b is understood. We call $\widehat{U}_{il}^t = (\widehat{U}_{il1}^t, ..., \widehat{U}_{ilK_l^t}^t)$ member i's estimated conditional payoffs at H_l^t, given

[19] The restriction to B^o is without loss of generality in our framework, as we will show later.

b. Denote by $\widehat{U}_l^t = (\widehat{U}_{1l}^t, ..., \widehat{U}_{n^t l}^t)$ is the profile of member estimated conditional payoffs at $H_l^t \in \Pi^t$. The estimation errors for members of team t are i.i.d. draws from a commonly known admissible probability distribution F_l^t. We also assume the estimation errors are independent across information sets, but allow different distributions at different information sets.[20]

A *team collective choice rule* at information set H_l^t, C_l^t, is a correspondence that maps profiles of member estimated payoffs at $H_l^t \in \Pi^t$ into a subset of elements of $A(H_l^t)$. Thus, C_l^t is a social choice correspondence. That is, $C_l^t : \Re^{n^t K_l^t} \to 2^{A(H_l^t)}$, so $C_l^t(\widehat{U}^t) \subseteq A(H_l^t)$. In principle, teams could be using different collective choice rules at different information sets, and denote $C = (C^1, ..., C^T)$, where $C^t = (C_1^t, ..., C_{L^t}^t)$. We assume that team t always mixes uniformly over $C_l^t(\widehat{U}_l^t)$. That is, the probability team t chooses a_{lk}^t at $\widehat{U}_l^t(b)$ is given by the function g^t defined as:

$$g_{lk}^{C_l^t}(\widehat{U}_l^t) = \frac{1}{|C_l^t(\widehat{U}_l^t)|} \text{ if } a_{lk}^t \in C_l^t(\widehat{U}_l^t)$$
$$= 0 \text{ otherwise}$$

Given a behavioral strategy profile, b, for each realization of $(\varepsilon_{l1}^t, ..., \varepsilon_{ln^t}^t)$ at information set $H_l^t \in \Pi^t$,, team t using collective choice rule C_l^t at H_l^t is assumed to take the action $a_{lk} \in A(H_l^t)$ with probability $g_{lk}^{C_l^t}(\widehat{U}_l^t)$.

We require C_l^t to satisfy *Unanimity* for all t and l, defined analogously to Definition 3. That is, for every team $t \neq 0$ and every information set $H_l^t \in \Pi^t$, if $\widehat{U}_{ilk}^t > \widehat{U}_{ilk'}^t$ for all $a_{lk'} \in A(H_l^t) - \{a_{lk}\}$ and for all $i \in t$, then $C_l^t(\widehat{U}_l^t) = \{a_{lk}\}$. It is important to note that our assumptions about F, together with unanimity of the collective choice rules, imply that for every t, every $H_l^t \in \Pi^t$, and every $a \in A(H_l^t)$, the probability that ε_l^t is such that $\widehat{U}_{ilk}^t > \widehat{U}_{ilk'}^t$ for all $a_{lk'} \in A(H_l^t) - \{a_{lk}\}$ and for all $i \in t$ is strictly positive. Therefore the behavioral strategies implied by the team choice probabilities are always totally mixed. That is for all $b \in B^o$, for all $t \in \mathbf{T}$, for all $H_l^t \in \Pi^t$, and for all $a \in A(H_l^t)$, $P_{lk}^{C_l^t}(U_l^t(b)) > 0$, and hence every possible history in the game occurs with positive probability. Thus, there are no "off-path" histories, so $\rho(z|H_l^t, b, a_{lk})$ is always well-defined and computed according Bayes' rule.

Team response functions are defined in the following way. Any totally mixed behavioral strategy profile, b, implies a profile of true conditional expected continuation payoffs U_l^t for each action at each information set H_l^t . Given U_l^t, the team collective

[20]This specification of errors can be interpreted in terms of an agent model the extensive form game.

choice rule, C_l^t, and the distribution of individual estimation errors, F_l^t, together imply a behavioral strategy response to b for team t at information set H_l^t, which we denote a *team response function* for team t by $P^{C_l^t}$, where $P_{lk}^{C_l^t}$ specifies the probability ε_l^t is such that a_{lk}^t is the team's action choice at H_l^t in response to b. That is:

$$P_{lk}^{C_l^t}(U_l^t(b)) = \int_{\varepsilon_l^t} g_{lk}^{C_l^t}(\widehat{U}_{lk}^t(b))dF_l^t(\varepsilon_l^t) \tag{2.7}$$

where F_l^t denotes the distribution of $\varepsilon_l^t = (\varepsilon_{l1}^t, ..., \varepsilon_{ln^t}^t)$.

For any extensive form game, any admissible F, and any profile of collective choice rules C call $\Gamma^{EF} = [\mathbf{T}, \mathbf{A}, n, \Xi, A, \iota, \Pi, b^0, u, F, C]$ a *team game in extensive form*. An equilibrium of a team game in extensive form is a fixed point of P.

Definition 12. *A **team equilibrium** of the team extensive form game Γ^{EF} is a behavioral strategy profile b such that, for every $t \neq 0$, and every information set $H_l^t \in \Pi^t$ and every $a_{lk}^t \in A(H_l^t)$, $b_{lk}^t = P_{lk}^t(U^t(b))$.*

Theorem 7. *For every team game in extensive form a team equilibrium exists and is in totally mixed behavioral strategies.*

Proof. For each t, and each $H_l^t \in \Pi^t$ and each $a \in A(H_l^t)$, define $\underline{b}_{lk}^t = \inf_{b \in B^o} P_{lk}^t(U_l^t(b))$. By Unanimity of C_l^t, $P_{lk}^t(U_l^t(b)) > 0$ so $\underline{b}_{lk}^t \geq 0$. Furthermore, we have $\underline{b}_{lk}^t > 0$ since $U_l^t(b)$ is uniformly bounded in B^o for all t and l, and hence, from Unanimity, $P_{lk}^t(U_l^t(b))$ is bounded strictly away from 0 for all t, l and k and for all $b \in B^o$. Define $\overline{B}^o = \{b \in B | b_{lk}^t \geq \underline{b}_{lk}^t \text{ for all } t, l \text{ and } k\} \subset B^o$. Since $\overline{B}^o$ is compact and convex and P_{lk}^t is a continuous function for all t, l and k, by Brouwer's fixed point theorem there exists a $b \in \overline{B}^o$ such that $b_{lk}^t = P_{lk}^t(U^t(b))$ for all t, l and k. □

A few observations about team equilibria in extensive form games are worth noting. First, the results in Section 4 about conditions for team response functions to satisfy payoff monotonicity and rank dependence carry through to extensive form team games, as applied to the behavioral strategies of each team. This is formally stated as follows.

Theorem 8. *For each t, Π^t, and $H_l^t \in \Pi^t$, if F_l^t is admissible and C_l^t satisfies unanimity and positive responsiveness then $P_l^{C_l^t}$ satisfies payoff monotonicity. Furthermore, if C_l^t also satisfies neutrality and $|A(H_l^t)| = 2$, or if $|A(H_l^t)| > 2$ and C_l^t is plurality rule or a weighted average rule, then $P_l^{C_l^t}$ satisfies rank dependence.*

Proof. The proof is essentially the same as the proof of Theorems 3, 4, and 5. □

Second, a stronger version of the Nash convergence property holds for team extensive form games. It is stronger because any limit point of a sequence of team equilibria in an extensive form game when teams become large is not just a Nash equilibrium, but must also be sequentially rational. This follows because equilibria in team games are always in B^o, so all information sets are on the equilibrium path and continuation payoffs are always computed simply using Bayes' rule.[21] This is formally stated as follows.

Theorem 9. *Consider an infinite sequence of team extensive form games,* $\left\{\Gamma_m^{EF}\right\}_{m=1}^{\infty}$, *where* $\Gamma_m^{EF} = [\mathbf{T}, \mathbf{A}, n_m, \Xi, A, \iota, \Pi, b^0, u, F, C]$,*where* m *indexes an increasing sequence of team sizes, with all the other characteristics of the game being the same. That is,* $n_{m+1}^t > n_m^t$ *for all* m, t. *Suppose* C_l^t *is an anonymous scoring rule for all* t, l *and let* $\{b_m^*\}_{m=1}^{\infty}$ *be a convergent sequence of team equilibria where* $\lim_{m\to\infty} b_m^* = b^*$. *Then* b^* *is a sequential equilibrium strategy of the corresponding extensive form game* $(\mathbf{I}, \mathbf{A}, \Xi, A, \iota, \Pi, b^0, u)$.

Proof. The proof is essentially the same as the proof of Theorem 2 except to additionally show that limit points are sequentially rational. First, the statement is not vacuous because for each m there exists at least one team equilibrium, b_m^*, and hence exists at least one convergent sequence of equilibria $\{b_m^*\}_{m=1}^{\infty}$ by the Bolzano-Weierstrass Theorem, with $\lim_{m\to\infty} b_m^* = b^*$. What we need to show is that there exist consistent beliefs, μ^* (i.e., assignments of a probability distribution over the histories at each information set that satisfy the Kreps-Wilson (1982) consistency criterion) such that, under those beliefs b^* specifies optimal behavior at each information set. That is, b^* is sequentially rational given μ^* and μ^* is consistent with b^*. We first show that μ^* is consistent with b^*. Because F^t has full support for all t and plurality rule satisfies unanimity, it follows that $b_m^* >> 0$ for all m. That is, for every m, t, H_l^t, and $a_{lk}^t \in A(H_l^t)$, $b_{mlk}^{t*} > 0$. Consequently every history occurs with positive probability, so, by Bayes' rule, for each information set H_l^t, and for all m, μ_{lm}^* is uniquely defined, where μ_{lm}^* denotes the equilibrium beliefs over the histories in the information set H_l^t. Since μ varies continuously with b there is a unique limit, $\mu^* = \lim_{m\to\infty} \mu_m^*$. Since $\mu_m^* >> 0$, and $\lim_{m\to\infty} b_m^* = b^*$ it follows that μ^* and consistent beliefs under b^*. What remains to be shown is that b_l^{t*} is optimal for all

[21] Not all sequential equilibria are approachable as limit points of team equilibria in extensive form. The example of non-approachability provided earlier applies here as well.

t and for all H_l^t. The proof is virtually the same as the proof of Theorem 2, so we omit it. □

Third, extensive form team games include games of incomplete information where teams have private information. For example, each team may have private information about u^t, or may have imperfect information about the path of play. Classic applications include signaling games.

Fourth, there is a rough connection between team equilibria and the extensive form version of quantal response equilibrium (Mckelvey and Palfrey 1998), with the main differences being that team size is fixed at n=1 in quantal response equilibrium and that the disturbances in team equilibrium are private value payoffs, but estimation errors, so all members of a team have common values. As in the agent model of quantal response equilibrium, in team games it is assumed that estimated expected continuation payoffs of each member of t at H_l^t are not observed until H_l^t is reached. If instead, each member of t observed all its estimates at the beginning of the game, this would lead to a different formulation of the model.

A last observation is that an alternative way to model team equilibrium in extensive form games would be to represent an extensive form game by its normal form or reduced normal form, and then apply the theory developed in Section 2 of this paper for games in strategic form. However, team equilibrium are *not* invariant to inessential transformations of the extensive form, so equilibria of the normal form or the reduced normal form will in general not be observationally equivalent to team equilibria in behavioral strategies derived from the extensive form. Such theoretical differences are suggestive of possible testable implications of the team equilibrium framework.

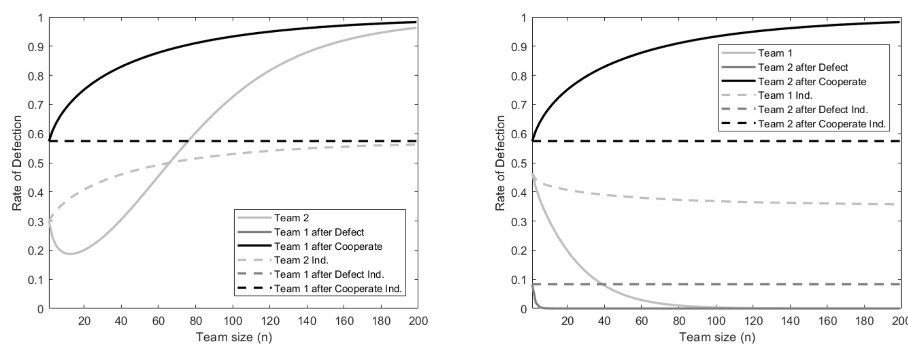
2.7 Examples of Extensive Form Games

Sequential Weak Prisoner's Dilemma Games

Next, we examine a sequential version of the Weak Prisoner's Dilemma, the simultaneous version of which we analyzed in section 3 (Table 2.2). In the sequential game, the equilibrium outcome depends on the order of moves. If column moves first, the unique subgame perfect Nash equilibrium is for both teams to defect. This is because after either cooperate or defect is chosen by column, row is better off choosing to defect. Since column optimizes by matching row's action, column should choose defect.

However, when row moves first, the unique subgame perfect equilibrium is for both teams to cooperate. This is because in equilibrium team 2 will choose whichever action team 1 chooses, defect after defect and cooperate after cooperate, and so team 1 will choose cooperate since $C > D$.

Figure 2.4 displays the team equilibrium for the two versions of this game with $x = y = 1$, and $z = 8$, the same parameter values used for the analysis of the simultaneous version in Section 3. The left panel shows the team equilibria if team 2 (column) moves first, and the right panel for the case where team 1 (row) moves first, for (odd) n ranging from 1 to 199, and $H(x) = \frac{1}{1+e^{-0.3x}}$. The solid light gray line is the team 2's (first mover) defect probability, and the dashed light gray line is team 2's individual member's defect probability. The solid black line the defect probability of team 1 after either cooperation or defection by team 2, and the dashed black line is team 1's individual member defect probability after either cooperation or defection by team 2.[22]



(a) Sequential WPD, team 2 moves first (b) Sequential WPD, team 1 moves first

Figure 2.4: Team Equilibrium in the Sequential WPD

When team 2 moves first, team 1's individual voting probabilities for defect are independent of n, because $x = y = 1$, so the payoff difference between defect and cooperate for team 1 is the same. In fact, they are the same as in the simultaneous version studied in Section 3, namely $H(y)$. So in this case, for any $H(\cdot)$ and for any n, the team equilibria of the simultaneous move version and team 2 first mover sequential version of the game are identical.

However, when team 1 moves first the team equilibria converge rapidly to the action profile (C, C). When team 1 chooses cooperate, the difference in expected utility to team 2 between choosing defect and cooperate is $-z < 0$, so the individual voting probabilities are $H(-z) < \frac{1}{2}$, and when team 1 chooses defect, the difference is y, so the individual voting probabilities in this case are $H(y) > \frac{1}{2}$.

[22]The black and dark gray curves coincide in the left panel (a) of Figure 2.4, so both appear as black.

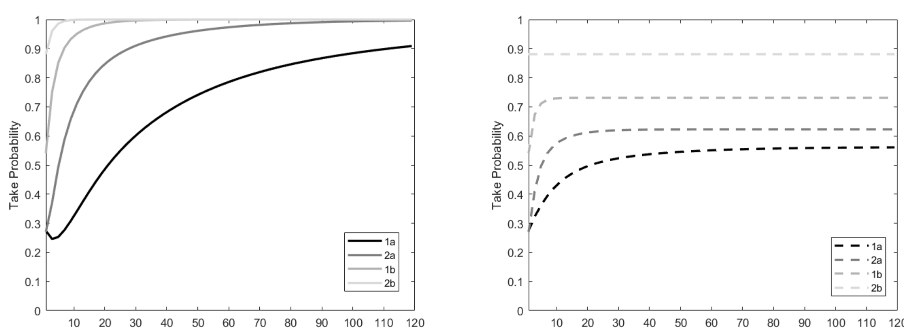
These voting probabilities are constant for all n, so team 2's individual voting probability of defection after defection converges to 1 and defection after cooperation converges to 0. Therefore, as n increases, the expected utility difference between defection and cooperation for team 1 decreases, and the team 1 individual voting probabilities and team probabilities decrease and converge to 0 probability of defection.

Centipede Games

Finally, we analyze team equilibria in a 4-move centipede game with exponentially increasing payoffs. At every outcome of this game, there is a high payoff and a low payoff, which initially equal 4 and 1, respectively. Two teams (1 and 2) take turns choosing to take or pass in sequence, starting with team 1. If team 1 chooses take the game ends, and team 1 receives the higher payoff while team 2 receives the lower payoff. If team 1 chooses pass, the two payoffs are doubled and team 2 gets to choose take or pass. This continues for up to 4 moves (fewer if one of the teams takes before the 4 node of the game), with the payoffs doubling after each pass. If pass is chosen at the last node, the game ends and team 1 receives 64 while team 2 receives 16. The two teams alternately play at most two nodes each in this game.

Since this is a finite game, the unique subgame perfect Nash equilibrium can be solved for by backward induction: choose take at every node.

In Figure 2.5 we display the team equilibrium majority-rule choice probabilities and individual voting probabilities for this game for (odd) n ranging from 1 to 119, and $H(x) = \frac{1}{1+e^{-x/8}}$. In the left panel are the team probabilities and in the right panel are the individual voting probabilities, with probability of taking and voting for take on the y axis and team size on the x axis.



(a) Centipede Game, Team Take Probabilities

(b) Centipede Game, Individual Voting Probabilities

Figure 2.5: Team Equilibrium in the Centipede Game

At the final node of the game, observe that the individual voting probabilities are fixed at $H(16) \approx 0.88$. As n increases, the team probability of taking at this node converges rapidly and monotonically to 1. Voting probabilities at the early nodes are influenced by the team choice probabilities at future nodes. As n increases, the probability that the opposing team will take at future nodes increases, decreasing the continuation value of passing. This causes the voting probabilities for take to increase and hence the team equilibrium probability of taking increases at every node as team size grows, which is consistent with experimental findings (Bornstein et al. 2004). Since, if the opposing team takes at the next node with probability 1, it is better to take at the current node than to pass, the individual voting probabilities converge to values strictly above $1/2$, and so majority rule ensures that all team take probabilities eventually converge to 1.

2.8 Discussion and Conclusions

This paper proposed and developed a theory of games played by teams of players. The framework combines the non-cooperative approach to model the strategic interaction between teams, with a collective choice approach to the decision making process within teams. The individual members of each team have correct beliefs on average about the expected payoffs to each available team strategy, given the distribution of strategy profiles being used by the other teams in the game. A team collective choice rule maps the profile of members' beliefs into a team strategy decision. Given an error structure and a collective choice rule, this induces a probability distribution over strategy choices for each team. A team equilibrium is a profile of mixed strategies, one for each team, with the property that the collective choice rule of each team will generate its equilibrium mixed strategy, given the distribution of beliefs of the individual members of the team.

The approach is initially formulated for finite games in strategic form. Four main results are proved for strategic form team games. First, team equilibria generally exist. Second, we show that all anonymous scoring rules satisfy the Nash convergence property: as team sizes become large, all limit points of team equilibria are Nash equilibria. Counterexamples are constructed to illustrate non-Nash limit points if teams do not use anonymous and/or neutral scoring rules. Third, we identify two weak conditions, unanimity and positive responsiveness, that are sufficient for team response functions to satisfy payoff monotonicity, in the sense that the probability a team chooses a particular action is increasing in the true expected payoff of that action. Fourth, we identify stronger conditions on the collective choice rule that guarantee rank dependence, i.e., the property that team choice probabilities are ordered by the actions' true expected payoffs.

Team equilibria for games in strategic form are illustrated for several 2×2 games, where the collective choice rule is majority rule. These examples illustrate two distinct effects of changing team size on outcomes. The first is the consensus effect. If the probability any individual on a team chooses one of the strategies is $p > 1/2$, then the probability a majority of the individuals on the team choose that strategy is greater than p, and is strictly increasing in the size of the team. The second effect is the equilibrium effect, which arises because in equilibrium, p will generally vary with n, and this equilibrium effect can go in the opposite direction from the consensus effect. Some of the examples suggest possible team games that might be interesting to study in the laboratory, where increasing team size can push equilibrium outcomes further away from Nash equilibrium.

The second half of the paper extends the framework to finite games in extensive form. Individuals are assumed to have correct beliefs on average at every information set about the expected continuation value of each available action at that information set. The results for strategic form team games about payoff monotonicity and rank dependence of team response functions and Nash convergence also apply to extensive form team games, with the latter result strengthened to show that limit points of team equilibrium in extensive form games are sequential equilibria.

We are hopeful that this framework is a useful starting point for the further exploration and understanding of how teams of players play games. There are many open questions that deserve further study and we mention a few. One is to generalize the class of collective choice rules that have the Nash convergence property. We identified one broad class of such collective choice rules (anonymous scoring rules), but we are well aware that there are many other rules with this property. One tempting conjecture is that Nash convergence obtains if all teams use a collective choice rule that satisfies unanimity, payoff monotonicity, rank dependence, neutrality, and anonymity. However it turns out that this is not sufficient. Consider the collective choice rule that uniquely selects action a_k^t if and only if a_k^t is unanimously ranked highest by all team members, and otherwise the collective choice rule selects the entire set, A^t. This satisfies all the above properties, but in the limit the team's strategy will always converge to uniform mixing over A^t, regardless of the mixed strategies of the other teams.

An interesting related question is how to extend the results about rank dependence with more than two actions to more general collective choice rules. We conjecture that rank dependence holds generally for collective choice rules that satisfy unanimity, positive responsiveness, and neutrality.

The framework can be expanded in several interesting directions. This paper assumed the collective choice rule of each team was exogenous, but as several of the examples suggest, teams may have preferences over collective choice rules in particular games. That is, one could pose the question, given the collective choice rule of the other teams, what would be an optimal collective choice rule for my team in the sense that the resulting team equilibrium with this profile of collective choice rules gives members the highest expected utility? Would optimal collective choice rules satisfy payoff monotonicity and rank dependence? Taking this a step further, one could define an equilibrium in collective choice rules and study its properties in different games.

Another direction to extend the framework would be to allow for a broader class of games than the finite games studied here. Many games of significant interest have infinite strategy spaces, including oligopoly, auction, and bargaining models in economics and spatial competition models in political science. In principle, the framework might be able to accommodate such an extension, for example by using finite approximations to the strategy space, but specific applications might face computational challenges. Similarly, some Bayesian games of interest, such as auctions, have a continuum of types; allowing for a continuum of types would seem to be a feasible extension if the action spaces are finite. The incorporation of behavioral biases and preferences (loss aversion, judgement biases, social preferences, etc.) would be straightforward, provided the effects are homogeneous across members of the group.

There are alternative approaches to modeling team games and extensions of the present approach that are beyond the scope of the framework presented here. For example, one might try to formalize the notion of “truth wins" - i.e., the idea that the team will adopt the choice favored by the most rational member of the group. This would require some formal notion of how to rank the rationality of the group members (such as level-k), coupled with a theory of persuasion, whereby the more rational members are able to change the beliefs of less rational members. Another alternative approach, which is more in the spirit of implementation theory and mechanism design, is to model the internal team decision making process as non-cooperative game rather than an abstract collective choice rule. This would create a nested game-within-a-game structure, which would add another layer of complexity.

Chapter 3

BIAS AND BELIEFS IN DETERRENCE AND DETECTION

3.1 Introduction

The economic analysis of crime typically focuses on the role of deterrence, through the enforcement of laws, in harm reduction and the protection of property rights. By outlawing harmful behaviors and imposing sanctions upon individuals or organization that might transgress the law, governments seek to reduce harm via deterrence. The possibility of punishment, however, places a strong incentive on offenders to actively hide their behavior. The quantity of crime is thus inherently difficult to measure. Only the portion of crime that enforcement agents successfully detect is observable to themselves, to other government officials and representatives, and ultimately to the public.

In light of this measurement problem, how do law enforcement agencies make decisions about investments into the detection of crime? In this paper, I present a law enforcement model, and an equilibrium notion, in order to address these issues. In the model, an enforcement agency makes a trade-off between a costly investment into detection that contributes to deterrence, and the harm caused by crime. The optimal investment depends on the magnitude of its deterrent effect, which is unknown to the agency.

By expending more resources on the detection of crime, enforcement agencies can increase the probability that any individual crime is detected, thereby imposing a greater expected punishment for committing the crime on potential offenders and reducing the quantity of offenses. This deterrent effect tends to decrease the quantity of crime the agency observes. By increasing its investment into detection, the agency also increases the proportion of crimes committed that it detects, which tends to increase the quantity of crime observed.

The only feedback available to the agency is the quantity of crime that it detects, and only from this data, and knowledge of its own behavior, can it learn about the quantity of crime and the response of crime to enforcement efforts. An equilibrium is an investment into detection, along with beliefs pertaining to the supply of crime, such that the investment is optimal given beliefs, and the data observed by the agency is consistent with beliefs. If the agency has an incorrectly specified model of the processes that generate the data it observes, it can hold incorrect beliefs and take sub-optimal actions in equilibrium even as it correctly predict the data.

One key insight captured by this model is the possibility of an inefficiently high or low detection investment arising from “feedback loops" between crime data and policing decisions. Many scholars in the law, criminology and machine learning communities, as well as numerous political activists, have pointed to the possibility for these feedback loops in the area of local policing. If a police department dispatches more police to patrol one neighborhood than it does to patrol another, it may detect more crime in the first neighborhood than in the second simply by virtue of having sent more police there. On the basis of this feedback it may incorrectly learn that the first neighborhood has more crime than it actually does, reinforcing its decision to assign more police to it. An equilibrium of the model in this paper captures exactly this intuition, and allows for a more precise analysis of the conditions under which a population is likely to be over-policed or under-policed.

The paper proceeds as follows. In section 2 I cover the related literature in economics. There I highlight especially connections to the theoretical economic literature on crime, the behavioral literature on overconfidence, and the empirical literature on police deterrence. In section 3, I present the model and equilibrium concept, and provide an illustrative example. Section 4 contains some general results on existence of equilibria. In section 5 I analyze a version of the model in which only equilibria with over-policing, or only equilibria with under-policing, exist under some conditions. Finally, in section 6 I conclude and discuss avenues for future study.

3.2 Related Literature

Becker (1968) is generally recognized as having inaugurated the modern theoretical economic study of crime. In this literature, the act of committing a crime is modeled as a risky gamble that imposes harm on third parties, and potential criminals as selfish expected utility maximizers. The objective of law enforcement is generally taken to be the maximization of social welfare, and the primary policy instruments are the magnitudes and types of punishments levied against offenders, and the quantity of costly effort expended toward detection and apprehension. Becker (1968) argues that the socially optimal policy is to maximize punishment magnitude and reduce detection effort, since increasing punishment magnitude is costless while detection effort is costly, and these variables are substitute determinants of the expected punishment for offending.

Subsequent studies identify conditions under which maximal punishments are socially undesirable. Since imprisonment is socially costly for many reasons, the maximal prison term length is generally not optimal, and less costly punishments should be used when possible (Becker (1968), Posner (1980)). When agents are risk averse in wealth, reducing fines may improve social welfare by reducing risk born by agents (Polinsky and Shavell (1979), Kaplow (1992)). If agents are uninformed of the legality of an illegal act (Kaplow (1990)), or if they are unsure of the probability of apprehension (Bebchuk and Kaplow (1992)), optimal fines are generally not maximal. Many other theoretical justifications of non-maximal punishments have been offered, Polinsky and Shavell (2000) and Garoupa (1997) give useful reviews of these and other related extensions of the model. In general, when there are some costs associated with increasing punishment severity, non-maximal punishments are optimal and some amount of costly investment into detection and apprehension is socially preferable, and it is the use of this policy instrument that is the focus of this study.

This previous literature on optimal law enforcement generally takes as known the distribution of benefits from crime in the population and the detection capabilities of enforcement agents. To extend the analysis to the case in which these primitives are, more realistically, unknown, I adopt a modeling approach related to several strands of recent theoretical study.

First, the equilibrium notion I use roughly corresponds to a notion of self-confirming equilibrium, introduced by Fudenberg and Levine (1993). In a self-confirming equilibrium, each agent best responds given his beliefs, and beliefs are correct along the equilibrium path but may be incorrect off-path. Similarly, I will say the law enforcement agency is in equilibrium when its beliefs about the crime data it observes are correct given its actual investment decision.

Secondly, there is a multitude of studies on the behavior of agents with incorrect beliefs about, or misspecified models of, their environment. For example, Kagel and Levin (1986) among others document a winners curse in common value auctions in the lab, especially among inexperienced bidders. Eyster and Rabin (2005) introduce the cursed equilibrium solution concept inspired by this phenomenon, in which agents do not understand how the actions of other agents depend on their private information. In another example, Rabin (2002) studies a model in which agents suffer from a belief in the “law of small numbers", the tendency to over-infer from small samples.

Most relevant for this study is the behavioral literature on overconfidence, an unrealistic belief in personal ability. Heidhues et al. (2018) study an overconfident agent that chooses an action in some production process, receives payoff feedback, and learns from this about a primitive of the production process. In my model, the mapping from detection investment to the probability of detecting crime can be interpreted as the enforcement agency's ability, and the incorrect beliefs about this mapping held by the agency can be interpreted as an institutional over- or underconfidence. Thus I contribute to the overconfidence literature by showing how an institution that is over- or underconfident at the organizational decision making level can be confirmed in its incorrect beliefs while making suboptimal decisions.

Alongside the large body of theoretical work on crime, is the large empirical literature in economics that seeks to estimate the deterrent effect of law enforcement. Chalfin and McCrary (2017) provides a helpful, recent review of methods and evidence. At least two key impediments to the estimation of the causal effect of policing on crime are generally acknowledged.

First, variance in enforcement efforts is likely not exogenous to other determinants of the quantity of crime. For example governments may increase police funding and manpower in response to increases in crime. Many early panel data studies using data on police manpower and crime found very small or incorrectly signed relationships between policing and crime. More recent instrumental variable and natural experiment studies, including Levitt (2002), Evans and Owens (2007) and Lin (2009) find large negative effects of police on crime. I am aware of little econometric work on the other direction of causation, the response of police behavior to crime.

Second, some authors have recognized that errors in the measurement of police activity or quantity of crime may influence empirical results. Levitt (1998) discusses random errors in the measurement of the quantity of crime, and dismisses the possibility that it has a major influence on empirical estimates of the deterrence effect. Chalfin and McCrary (2013, 2018) argue that mis-measurement of the quantity of police is a better explanation than simultaneity bias for the small estimates of the deterrent effect in cross-sectional and panel data studies. After correcting for measurement error they find a strongly negative relationship between policing and the quantity of violent crime.

However, to the best of my knowledge, measurement error in the quantity of crime resulting from variation in policing intensity has not been addressed in the empirical literature. Crime data collected by police agencies are likely to suffer from a missing data problem. Only those crimes detected by police show up in the data, and the proportion of crime detected depends on the allocation of police resources. It is possible that, holding the actual amount of crime constant, more heavily policed areas show up as having more crime in the data. Without a more thorough understanding of the relationship between crime data and police behavior, it is unclear what its implications are for empirical estimates of the deterrent effect.

3.3 Model

In this section I introduce a model of public enforcement of the law, extend the model to the case in which certain primitives are unknown, and finally provide conditions for existence of an equilibrium.

Public Law Enforcement Model

A population of agents with unit mass must each decide either to commit a potentially harmful act, yielding private benefit b_h, or engage in some alternative, harmless activity yielding private benefit b_l. The agents have heterogeneous net private benefits $b = b_h - b_l$, distributed according to cdf $G(\cdot)$, with pdf or pmf $g(\cdot)$. The act imposes a harm, or cost $h > 0$ on third parties, while the harmless act imposes no costs on third parties. A government authority outlaws the act, making it a crime, establishes a fine, and tasks an enforcement agency with detection, investigation and apprehension of agents that commit the act.

The enforcement agency chooses an investment into detection $x \geq 0$, which has a constant marginal cost of 1. An investment in detection may for example be an investment of resources into police patrol, or auditing. Let $p : \mathbb{R}_+ \to [0, 1]$, the “detection function", give the probability of detecting a crime given the level of investment. Assume p is twice continuously differentiable, and that $p' > 0$ and $p'' \leq 0$. Let f be the fixed expected fine conditional on being detected. Assume the potential criminals are risk neutral, then potential criminals will commit the crime when $b > p(x)f$.

The objective of the law enforcement agency is to maximize social welfare, defined here as the sum of all agents' payoffs. If the fine is treated as a cost-less transfer, the social welfare can be written as follows.

$$
\begin{aligned}
W(x) &= \int_{p(x)f}^{\infty} (b-h)g(b)db - x \\
&= \int_{p(x)f}^{\infty} bg(b)db - [1 - G(p(x)f)]h - x
\end{aligned}
$$

A typical analysis of the model would proceed with the assumption that the law enforcement agency knows the distribution $G(\cdot)$, detection function $p(\cdot)$ the harm h and the constant marginal cost 1, and selects the investment into detection that maximizes $W(x)$.

For a continuous distribution over benefits, the first order condition for a non-zero investment is

$$
g(p(x)f)p'(x)f[h - p(x)f] = 1
$$

The marginal benefit of increasing x is the product of the marginal decrease in the mass of crimes committed $g(p(x)f)p'(x)f$, with the social benefit of deterring the marginal crime $[h - p(x)f]$, and this must equal the marginal cost of 1. A sufficient second order condition for optimality is that the density $g(p(x)f)$ is weakly decreasing for all $x \geq 0$, which guarantees that the objective is strictly concave over all x such that $p(x)f \in [p(0)f, f]$, the set of feasible expected fines imposed on potential criminals.

Not the important observation first made by Polinsky and Shavell (1984) that at the optimal investment, all efficient crimes, those with private benefit greater than harm h, are committed, and some inefficient crimes, those with private benefit less than h, also go undeterred since $[h - p(x)f]$ must be positive.

However, it remains to be explained how the law enforcement agency knows the primitives that appear in their social welfare objective. In particular the distribution $G(\cdot)$ and the detection function $p(\cdot)$ that determines the supply of crime.

Beliefs and Equilibrium

In this section, I extend the law enforcement model presented in the previous section to the case of unknown $G(\cdot)$ and $p(\cdot)$. The general modeling approach taken follows in the spirit of the self-confirming equilibrium literature. The enforcement agency is endowed with some beliefs about $G(\cdot)$ and $p(\cdot)$ and maximizes social welfare given these beliefs. Crime data gathered by the agency provides it with observable feedback about the supply of crime. Only when this data is consistent with the agency's beliefs does the agency have no reason to update its beliefs. Only when the agency believes it is maximizing social welfare does it have no reason to change its investment choice. These two conditions motivate the solution concept presented in this section.

Definition 1. *An **Enforcement Environment** is a tuple* $(p, \hat{p}, \Theta, \theta^*, (G_\theta)_{\theta\in\Theta}, h, f)$.

The family of distributions $(G_\theta)_{\theta\in\Theta}$ parametrized by Θ represents the set of distributions the agency believes to be possible a priori. A (point) belief of the agency is a parameter $\theta \in \Theta$. Let $\theta^* \in \Theta$ be the parameter corresponding to the true distribution G_{θ^*}. I assume the agency is endowed with a fixed belief about the detection function, $\hat{p}(.)$, with $\hat{p}' > 0$ and $\hat{p}'' \leq 0$. The true detection function is labeled $p(\cdot)$, also with $p' > 0$ and $p'' \leq 0$. The parameters h and f are the fixed harm and fine respectively.

The assumption of a fixed, incorrect belief about some primitive of the model is similar to the assumptions made in the over-confidence literature, for example in Heidhues et al. (2018), in which the agent is assumed to have a fixed belief about his own ability, and learns about some other primitive of a production process by observing payoff feedback. The detection function in this model may in some ways be interpreted as an analog of personal ability in over-confidence models. It represents the agency's ability to use resources to fulfill its role in detecting crime. When $\hat{p}(x) > p(x)$ for all x, the agency may be said to display "over-confidence", and when $\hat{p}(x) < p(x)$ for all x we might label it "under-confident".

Let $x^*(\theta)$ and $\hat{x}(\theta)$ denote the sets of maximizers of social welfare at θ under the true detection function $p(\cdot)$, and under the belief $\hat{p}(.)$, respectively.

$$x^*(\theta) = \text{argmax}_{x\geq 0} \int_{p(x)f}^{\infty} bg_\theta(b)db - [1 - G_\theta(p(x)f)]h - x$$
$$\hat{x}(\theta) = \text{argmax}_{x\geq 0} \int_{\hat{p}(x)f}^{\infty} bg_\theta(b)db - [1 - G_\theta(\hat{p}(x)f)]h - x$$

The only observable data available to the agency is the quantity of crime that it detects, which is determined jointly by the probability the agency detects a crime, and the quantity of crime committed. The quantity of crime committed is in turn determined by the incentives faced by potential criminals: the private benefit, the fine, and the probability of being detected.

Given a distribution of benefits G_θ and fixed fine f, the supply of crime as a function of the probability of detection p is given by $S_\theta(p) = [1 - G_\theta(pf)]$. When the agency invests x into detection, its observable data is $p(x)S_{\theta^*}(p(x))$. When the agency has belief θ, it expects to observe $\hat{p}(x)S_\theta(\hat{p}(x))$.

Adopting the notation of Heidhues et al. (2018), define the surprise function $\Gamma(x, \theta; \theta^*)$ as follows:

$$\Gamma(x, \theta; \theta^*) = p(x)S_{\theta^*}(p(x)) - \hat{p}S_\theta(\hat{p}(x))$$

$\Gamma(x, \theta; \theta^*)$ gives the difference between the quantity of detected crime and the quantity the agency expects to detect when the agency holds belief θ and invests x. Equilibrium attains when the agency is not surprised by the feedback generated by its investment, and it chooses an investment that maximizes its objective given its beliefs.

Definition 2. *An* ***Equilibrium*** *of an enforcement environment* $(p, \hat{p}, \Theta, \theta^*, (G_\theta)_{\theta\in\Theta}, h, f)$, *is an ordered pair* $(\hat{x}, \hat{\theta}) \in \mathbb{R}_+ \times \Theta$ *such that*

1. $\hat{x} \in \hat{x}(\theta) = argmax_{x\geq 0} \quad \int_{\hat{p}(x)f}^{\infty} bg_{\hat{\theta}}(b)db - [1 - G_{\hat{\theta}}(p(x)f)]h - x$

2. $\Gamma(\hat{x}, \hat{\theta}; \theta^*) = 0$

Discrete Distribution Example

To illustrate the equilibrium concept, I present an example with a two point discrete distribution over the private benefit of crime and linear detection functions.

Suppose all potential criminals derive benefit $b \in \{\underline{b}, \overline{b}\}$, with $0 < \underline{b} < \overline{b}$. Let the true proportion of potential criminals with benefit $\overline{b}$ be $\theta^* \in \Theta = [0, 1]$. Let the true detection function be $p(x) = x + \beta$, and suppose the police agency believes the detection function is $\hat{p}(x) = x + \hat{\beta}$. The law enforcement agency holds incorrect beliefs about the proportion of crime detected when the investment is 0. In the interest of simplicity let $h = f = 1$.

Suppose that $\beta < \hat{\beta} < \underline{b} < 1 < \overline{b}$. At zero investment, no crime is deterred, since the true and perceived expected fine at this investment is $p(0)f = \beta$ and $\hat{p}(0)f = \hat{\beta}$. Additionally, it is efficient for the high benefit crime to be committed, since the high benefit is greater than the harm of 1. Consequently full deterrence is dominated by partial deterrence, and so will never be chosen.

The law enforcement agency therefore chooses between two investment levels: $x = 0$, the no deterrence investment, and $x = \underline{b} - \hat{\beta}$, the partial deterrence investment, to deter only the low benefit potential criminals. The agency perceives that any intermediate investment wastes resources, that there exists a lower investment that deters the same amount of crime.

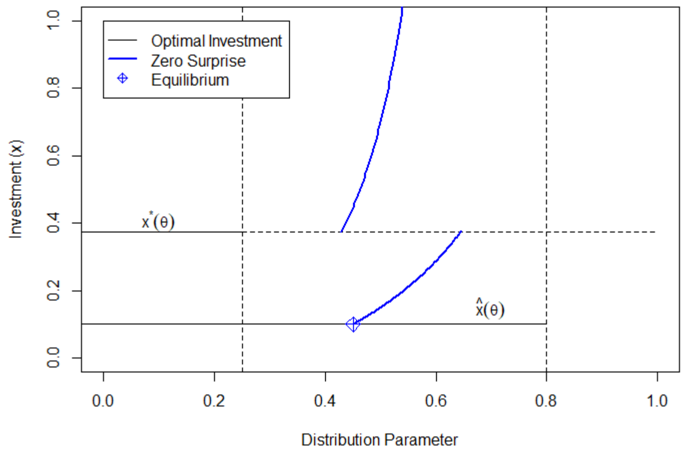
The choice between partial deterrence and no deterrence pivots on the value of θ. For low values of θ, most potential criminals have a low benefit, and so deterrence is highly effective. At higher values of θ, the high benefit group outnumbers the low benefit group, and the costs of deterrence outweighs the smaller amount of crime deterred. In particular, we have that partial deterrence is preferred when

$$\begin{aligned}
\theta(\overline{b} - 1) - (\underline{b} - \hat{\beta}) &> \theta(\overline{b} - 1) + (1 - \theta)(\underline{b} - 1) \\
\theta(\underline{b} - 1) &> (\underline{b} - 1) + (\underline{b} - \hat{\beta}) \\
\theta &< 1 - \frac{\underline{b} - \hat{\beta}}{1 - \underline{b}}
\end{aligned}$$

and otherwise no deterrence is preferred.

In figure 1 I plot $\hat{x}(\theta)$, the enforcement agency's optimal investment as a function of the distribution parameter, given their belief about the detection function $\hat{p}(x)$, and the optimal investment $x^*(\theta)$ given the true detection function, for the values

Figure 3.1: Equilibrium For Two Point Distribution Example



$$\beta = \frac{1}{8}, \quad \hat{\beta} = \frac{2}{5}$$
$$\underline{b} = \frac{1}{2}, \quad \overline{b} = 2, \quad \theta^* = 2/3$$

The actual investment into detection required for partial deterrence is $\underline{b} - \beta = \frac{3}{8}$, but the agency invests only $\frac{1}{10}$, so that at no belief θ will the agency achieve even partial deterrence.

In equilibrium, the agency must be optimizing given its beliefs, and the quantity of detected crime must be equal to the quantity of crime the agency predicts it will detect. The surprise function gives the difference between these two quantities.

$$\begin{aligned}
\Gamma(x,\theta;\theta^*) &= p(x)S_{\theta^*}(p(x)) - \hat{p}(x)S_\theta(\hat{p}(x)) \\
&= (x+\beta)\theta^* - (x+\hat{\beta})\theta && \text{if} \quad x \geq (\underline{b}-\beta) \\
&= (x+\beta) - (x+\hat{\beta})\theta && \text{if} \quad x \in [(\underline{b}-\hat{\beta}), (\underline{b}-\beta)) \\
&= (x+\beta) - (x+\hat{\beta}) < 0 && \text{if} \quad x < (\underline{b}-\hat{\beta})
\end{aligned}$$

An equilibrium is a point belief $\hat{\theta} \in \Theta$, along with investment $\hat{x}(\hat{\theta})$, such that $(\hat{x}(\hat{\theta}), \hat{\theta})$ has zero surprise, that is $\Gamma(\hat{x}(\hat{\theta}), \hat{\theta}; \theta^*) = 0$. In figure 1, this is visualized as the intersection of two sets, $\{x, \theta; x = \hat{x}(\theta)\}$, the graph of the otpimizer, and $\{x, \theta; \Gamma(x, \theta; \theta^*) = 0\}$, the region of (x, θ) space in which the surprise is zero.

We search now for such an equilibrium. First, since there does not exist a θ for which $\hat{x}(\theta) > (\underline{b} - \beta)$, this case can be discarded. Since $\hat{\beta} > \beta$, the surprise is strictly negative whenever $x < (\underline{b} - \hat{\beta})$, so this case is also eliminated.

The only remaining possibility for an equilibrium is the case in which $\hat{x}(\theta) \in [(\underline{b} - \hat{\beta}), (\underline{b} - \beta))$, in particular at the partial deterrence investment $\underline{b} - \hat{\beta}$. Computing the surprise function for this value of x:

$$\Gamma((\underline{b} - \hat{\beta}), \hat{\theta}; \theta^*) = 0$$
$$(\underline{b} - \hat{\beta} + \beta) - \underline{b}\hat{\theta} = 0$$
$$\hat{\theta} = \frac{\underline{b} - \hat{\beta} + \beta}{\underline{b}}$$

So there is an equilibrium with $\hat{\theta} = \frac{\underline{b} - \hat{\beta} + \beta}{\underline{b}}$ and $\hat{x}(\hat{\theta}) = \underline{b} - \hat{\beta}$, as long as $\hat{\theta} < 1 - \frac{\underline{b} - \hat{\beta}}{1 - \underline{b}}$.

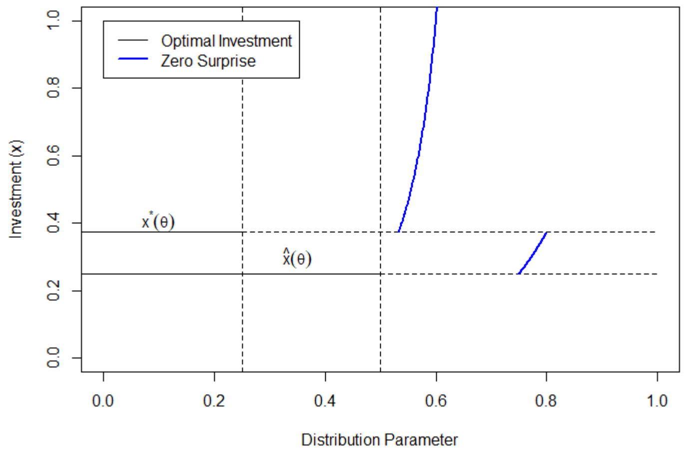
For the numerical values displayed in figure 3.1, we have that this is $\hat{\theta} = \frac{9}{20}$ and $\hat{x}(\hat{\theta}) = \frac{1}{10}$. The blue curve in figure 1 traces out set of pairs (x, θ) at which the surprise function Γ is equal to 0, for $\theta^* = 2/3$. The equilibrium is the point where the zero surprise curve intersects with $\hat{x}(\theta)$.

In this equilibrium, the enforcement agency invests in what it believes to be the partial deterrence probability of detection. In truth it does not deter any crime, so that the true supply of crime is 1. The agency believes that the high benefit type of potential criminal is undeterred, so different values of θ induce different predictions by the agency about the quantity of detected crime. Exactly one of those values of θ rationalizes the data observed by the agency, and if this belief supports partial deterrence, then it is an equilibrium belief.

3.4 Equilibrium Existence

In this section, I provide sufficient conditions on $p(\cdot)$, $\hat{p}(.)$ and $(G_\theta)_{\theta \in \Theta}$ for the existence of an equilibrium. An example of an enforcement environment in which no equilibrium exists is first given, in order to motivate the subsequent section in which the sufficient conditions are described.

Figure 3.2: Equilibrium Non-Existence with Two Point Distribution



Example of Non-Existence under Misspecification

In the two-point distribution example presented in the previous section, when $\beta < \hat{\beta} < \underline{b} < 1 < \overline{b}$, an equilibrium exists if and only if the zero surprise belief at the perceived partial deterrence investment $\underline{b} - \hat{\beta}$ is small enough so that partial deterrence is optimal. In the inequality below, the left hand side is the zero surprise belief and the right hand side is the cutoff for optimality of the partial deterrence investment.

$$\frac{(\underline{b} - \hat{\beta} + \beta)}{\underline{b}} \leq 1 - \frac{\underline{b} - \hat{\beta}}{1 - \underline{b}}$$

Fixing $\underline{b}, \overline{b}, \beta$, the parameter $\hat{\beta}$ must be large enough to guarantee existence. Non-existence is illustrated in figure 3.2. All parameters are the same as in figure 3.1 except for $\hat{\beta}$ which is now set to $\frac{1}{4}$.

When $\hat{\beta}$ is not large enough, the surprise at the optimal allocation $\Gamma(\hat{x}(\theta), \theta; \theta^*)$ is strictly positive when the partial deterrence investment is optimal, and strictly negative when 0 investment is optimal.

The key obstructions to existence in this case are the discontinuity of the surprise function and optimum investment, and the fact that at a detection investment of zero, the quantity of crime does not vary with the distribution parameter. Both of these features are consequences of the discreteness of the benefit distribution. By focusing on continuous distributions over benefits, both obstructions can be lifted with some mild assumptions.

Sufficient Conditions for Existence

In this section I focus on families of continuous distributions $(G_\theta)_{\theta\in\Theta}$ and provide sufficient conditions under which an equilibrium is guaranteed to exist.

The general strategy I follow to prove existence involves identifying conditions under which the surprise function $\Gamma(\hat{x}(\theta), \theta; \theta^*)$ is continuous in θ, positive at one extreme of Θ, and negative at the other extreme. The result then follows from the intermediate value theorem.

Assumption 1: For all $\theta \in \Theta$, $G_\theta(.)$ is continuously differentiable with derivative $g_\theta(.)$ weakly decreasing on $\mathbb{R}_+$.

First, recall from section 3.3 that if the density of benefits $g_\theta(.)$ is weakly decreasing on $\mathbb{R}_+$, then the enforcement agency's objective is strictly concave on $\mathbb{R}_+$ and the second order condition is satisfied. Strict concavity also guarantees unique maxima, so that $\hat{x}(\theta)$ and $x^*(\theta)$ are functions.

Next, note that an investment of zero dominates any investment greater than the harm h. For all θ, for all $x' > h$, we have

$$[G_\theta(\hat{p}(x')f) - G_\theta(\hat{p}(0)f)]h - x' < 0 < \int_{\hat{p}(0)f}^{\hat{p}(x')f} bg_\theta(b)db$$

$$\int_{\hat{p}(x')f}^{\infty}(b-h)g_\theta(b)db - x' < \int_{\hat{p}(0)f}^{\infty}(b-h)g_\theta(b)db$$

Intuitively, the maximum social harm net of private benefit prevented by an investment of x' is h, so an investment that costs more than h is always worse than an investment of zero. As a result we can, wlog, restrict the choice set of the agency to the compact interval $[0, h]$. The maximum theorem then implies that $\hat{x}(\theta)$ and $x^*(\theta)$ are continuous functions.

Assumption 2: $\Theta = (0, \infty)$, $G_\theta(.)$ converges point-wise to the constant function that equals 0 as $\theta \to \infty$, and point-wise to 1 as $\theta \to 0$, and $G_\theta(z)$ is continuous in θ for all z.

Assumption 2 is a technical, richness assumption on the family of distributions $(G_\theta)_{\theta\in\Theta}$. It says that a proportion ϕ of the mass of the distribution can be pushed (in a continuous way) arbitrarily far away from zero by increasing θ, and an arbitrary amount of mass can be pushed below zero by decreasing θ. In other words, the agency's model of crime does not reject, a priori, the possibility that no one would ever commit crime, or that everyone would always commit crime, so that enforcement policies have no effect on the supply of crime.

Examples of families of distributions that satisfy both assumptions 1 and 2 include the uniform on $[0, \theta]$, the exponential with parameter $1/\theta$.

Assumption 3: $p(0) = \beta > 0$ and $\hat{p}(0) = \hat{\beta} > [1 - G_{\theta^*}(\beta f)]$.

Assumption 3 says that at an investment of zero the agency actually detects, and believes it detects, a positive proportion of crime β and $\hat{\beta}$ respectively. Even when no investment is made into detecting crime, some crime may be revealed either through the self-reports of victims, or through some secondary sources of information. For example victims of theft may file a report with the police, and traffic accidents may reveal evidence of some traffic violations independently of the activity of highway patrol.

The condition that $\hat{\beta} > [1 - G_{\theta^*}(\beta f)]$ says that at an investment of zero, the proportion of crime the agency expects to detect is greater than the quantity of crime it actually detects.

Theorem 1. *If the enforcement environment $(p, \hat{p}, \Theta, \theta^*, (G_\theta)_{\theta\in\Theta}, h, f)$ satisfies assumptions 1-3, then there exists an equilibrium.*

Proof. Assumption 1 implies that $\hat{x}(\theta)$ is single valued for all $\theta \in \Theta$. Since $\hat{x}(\theta)$ is bounded above by h, the maximum theorem implies that $\hat{x}(\theta)$ is a continuous function of θ. It follows immediately that $S_{\theta^*}(p(\hat{x}(\theta))f)$ is continuous in θ.

By assumption 2, $G_\theta(.)$ is continuous in θ, so we have that $S_\theta(\hat{p}(\hat{x}(\theta))) = [1 - G_\theta(\hat{p}(\hat{x}(\theta))f)]$ is a continuous function of θ. It follows that $\Gamma(\hat{x}(\theta), \theta; \theta^*) = p(\hat{x}(\theta))S_{\theta^*}(p(\hat{x}(\theta))) - \hat{p}(\hat{x}(\theta))S_\theta(\hat{p}(\hat{x}(\theta)))$ is continuous in θ.

Assumption 2 and 3 together imply that there exist $\underline{\theta}, \overline{\theta} \in \Theta$ with $\underline{\theta} < \overline{\theta}$ such that

1. $\beta[1 - G_{\theta^*}(f)] \geq [1 - G_{\underline{\theta}}(\hat{\beta} f)]$

2. $[1 - G_{\theta^*}(\beta f)] \leq \hat{\beta}[1 - G_{\overline{\theta}}(\hat{\beta} f)]$

p increasing implies that $p(\hat{x}(\underline{\theta}))[1 - G_{\theta^*}(p(\hat{x}(\underline{\theta}))f)] \geq \beta[1 - G_{\theta^*}(f)]$ and $\hat{p}$ increasing implies that $[1 - G_{\underline{\theta}}(\hat{\beta} f)] \geq \hat{p}(\hat{x}(\underline{\theta}))[1 - G_{\theta^*}(\hat{p}(\hat{x}(\underline{\theta}))f)]$. So line 1 above guarantees that $\Gamma(\hat{x}(\underline{\theta}), \underline{\theta}; \theta^*) \geq 0$.

By the same argument, line 2 guarantees that $\Gamma(\hat{x}(\overline{\theta}), \overline{\theta}; \theta^*) \leq 0$. It follows from the intermediate value theorem that there exists a $\hat{\theta} \in [\underline{\theta}, \overline{\theta}]$ such that $\Gamma(\hat{x}(\hat{\theta}), \hat{\theta}; \theta^*) = 0$, and by definition $(\hat{x}(\hat{\theta}), \hat{\theta})$ is an equilibrium.

□

3.5 Over-Policing and Under-Policing

In this section I present an analysis of a class of enforcement environments in which the equilibrium detection investment may be either above or below the optimal investment level, which I refer to as “over-policing" and “under-policing" respectively.

Let the family of distributions over benefits from crime be the uniform distributions on $[0, \theta]$, $G_\theta(z) = \frac{z}{\theta}$, for $z \leq \theta$. Let the detection functions be linear with $p(x) = x + \beta$ and $\hat{p}(x) = \alpha x + \beta$, with $\alpha \geq 1$. The agency has the correct belief about the proportion of crime detected at an investment of 0, but incorrect beliefs about the marginal benefit of increasing the investment. In particular it is overconfident since it overestimates its probability of detecting any given crime at every x. Let $f = h = 1$. Since the fine is 1, the expected fine imposed on offenders by an investment x is $p(x)$.

As a first step in the analysis, the optimal investment under p and $\hat{p}$ for every belief θ must be derived. First note that any potential offenders with benefit $b < \beta$ are deterred from committing the crime at any investment level. When the investment is 0, the expected fine is β, so $\theta \leq \beta$ implies the entire population is deterred at an investment of zero. The optimal investment for the agency must be zero.

When $\theta > \beta$, some investment is required for full deterrence. In particular, in order for the agency to believe that it achieves full deterrence, it must be that the expected fine is greater than the maximum benefit, $\hat{p}(x) \geq \theta$ or $x \geq \frac{\theta - \beta}{\alpha}$.

When $x < \theta$, the first derivative of the social welfare objective with respect to investment x is

$$
\begin{aligned}
\frac{\partial}{\partial x} W(x) &= g_\theta(\hat{p}(x))\hat{p}'(x)[1 - \hat{p}(x)] - 1 \\
&= \frac{1}{\theta}\alpha[1 - (\alpha x - \beta)] - 1 \\
&= \frac{1}{\theta}2[1 - (2x - \frac{1}{4})] - 1
\end{aligned}
$$

This derivative is decreasing in x, so the objective is indeed strictly concave. When this derivative is positive at the full deterrence investment, it must be that full deterrence is optimal. This is true when

$$
\begin{aligned}
\frac{1}{\theta}\alpha[1 - (\alpha\frac{\theta - \beta}{\alpha} + \beta)] &\geq 1 \\
[1 - \theta] &\geq \frac{\theta}{\alpha} \\
\frac{\alpha}{\alpha + 1} &\geq \theta
\end{aligned}
$$

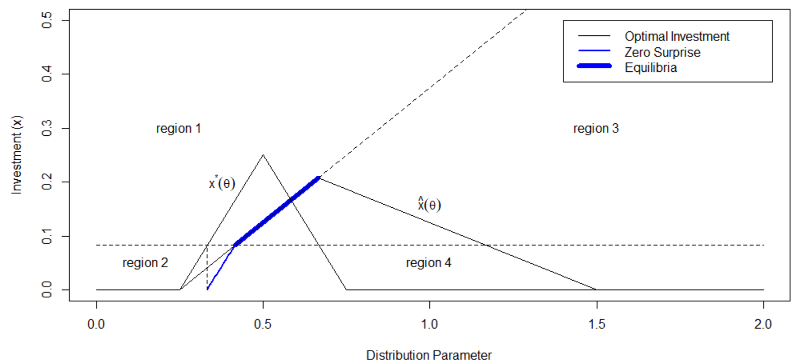
When the derivative is negative at $x = 0$, it must be negative for all x, so zero investment is optimal. That condition is written below.

$$
\begin{aligned}
\frac{1}{\theta}\alpha[1 - \beta] &\leq 1 \\
\alpha[1 - \beta] &\leq \theta
\end{aligned}
$$

For any intermediate value, the first order condition binds and the optimal investment is given by

$$
\begin{aligned}
\frac{1}{\theta}\alpha[1 - \alpha x - \beta] &= 1 \\
x &= \frac{(1 - \beta)}{\alpha} - \frac{\theta}{\alpha^2}
\end{aligned}
$$

Figure 3.3: Over-Policing with $\theta^* = \frac{1}{3}$



We therefore have that the optimal investment when the detection function is $\alpha x + \beta$, and the distribution of benefits from crime is $U[0, \theta]$, is given by the continuous, piecewise linear function:

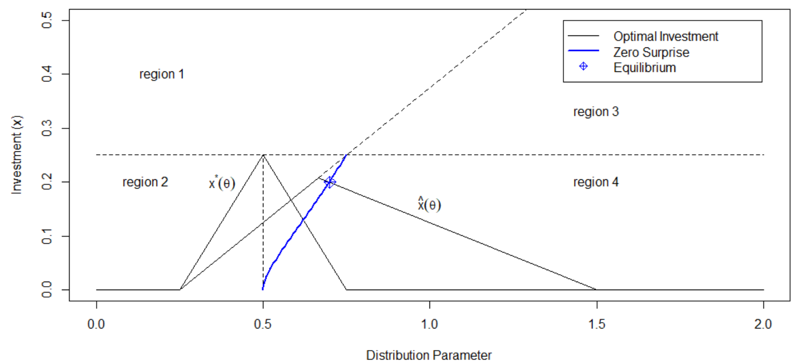
$$\begin{aligned}
\hat{x}(\theta) &= 0 && \text{if} \quad \theta < \beta \\
&= \frac{\theta - \beta}{\alpha} && \text{if} \quad \theta \in [\beta, \frac{\alpha}{\alpha + 1}) \\
&= \frac{(1-\beta)}{\alpha} - \frac{\theta}{\alpha^2} && \text{if} \quad \theta \in [\frac{\alpha}{\alpha+1}, \alpha(1-\beta)) \\
&= 0 && \text{if} \quad \theta \geq \alpha(1-\beta)
\end{aligned}$$

For $p(x) = x + \frac{1}{4}$ and $\hat{p}(x) = 2x + \frac{1}{4}$, the functions $x^*(\theta)$ and $\hat{x}(\theta)$ are displayed in figures 3.3 and 3.4. The dashed diagonal emanating from $\hat{x}(\theta)$ shows the full deterrence investments for every θ, under the agency's belief $\hat{p}$. On the upward sloping segment of these curves, full deterrence is optimal, on the downward sloping segments, partial deterrence is optimal, with the optimal investment decreasing in θ, and elsewhere the optimal investment is zero.

The intersection of the graph of $\hat{x}(\theta)$, with the set of pairs (x, θ) such that the surprise $\Gamma(x, \theta; \theta^*)$ is zero, forms the set of equilibria. We need next to derive the latter set.

There are four disjoint regions of (x, θ) space that require separate analyses: The region in which x provides full deterrence under both p and $\hat{p}$, in which x provides full deterrence under only p and not under $\hat{p}$, in which x provides full deterrence

Figure 3.4: Under-Policing with $\theta^* = \frac{1}{2}$



only under $\hat{p}$ and not under p, and in which x does not provide full deterrence under either.

Region 1: $x \geq (\theta^* - \beta)$ and $x \geq \frac{\theta - \beta}{\alpha}$

Since $x \geq (\theta^* - \beta)$, we have $p(x) \geq \theta^*$, so in truth the expected fine imposed on offenders is greater than the maximum benefit, so the true supply of crime is zero. Similarly by the second inequality, $\hat{p}(x) \geq \theta$, so the agency believes that it deters all crime. Therefore, the surprise is zero in this entire region, as the agency correctly predicts that it will observe no crime. In figures 3.3 and 3.4, this is the top left region.

When this region intersects with $\hat{x}(\theta)$, as in figure 3.3, there exists an interval of equilibria in which the agency believes it fully deters all crime, and does indeed achieve full deterrence. All of these equilibria have over-policing.

Region 2: $x < (\theta^* - \beta)$ and $x \geq \frac{\theta - \beta}{\alpha}$

In this region, the agency believes that it fully deters all crime, but in reality it does not. As a result, the difference between the quantity of crime it observes and the quantity it expects to observe (which is zero) is strictly positive. In figure 3.3 and figure 3.4, this is the lower left region, above the upward sloping portion of $\hat{x}(\theta)$, and below the full deterrence investment $(\theta^* - \beta)$.

Region 3: $x \geq (\theta^* - \beta)$ and $x < \frac{\theta - \beta}{\alpha}$

In this region, the agency deters all crime, but believes that it does not. The surprise is therefore strictly negative everywhere in this region, since the agency observes zero crime, but expects to observe some positive amount of crime. In the figures, this region is in the top right, below the upward sloping part of $\hat{x}(\theta)$ and above the full deterrence investment.

Region 4: $x < (\theta^* - \beta)$ and $x < \frac{\theta-\beta}{\alpha}$

In this region, there is only partial deterrence in truth and in the agency's expectations. The surprise function is given by

$$\begin{aligned}\Gamma(x,\theta;\theta^*) &= p(x)[1 - G_{\theta^*}(p(x))] - \hat{p}(x)[1 - G_\theta(\hat{p}(x))] \\ &= (x+\beta)[1 - \frac{x+\beta}{\theta^*}] - (\alpha x + \beta)[1 - \frac{\alpha x + \beta}{\theta}]\end{aligned}$$

Note that the surprise function is strictly decreasing in θ for all x and is continuous in x and θ, so setting it equal to zero and solving for θ in terms of x yields a formula for a continuous curve along which the surprise is zero.

Furthermore, consider the surprise at an investment of zero.

$$\begin{aligned}\Gamma(0,\theta;\theta^*) &= p(0)[1 - G_{\theta^*}(p(0))] - \hat{p}(0)[1 - G_\theta(\hat{p}(0))] \\ &= \beta[1 - \frac{\beta}{\theta^*}] - \beta[1 - \frac{\beta}{\theta}]\end{aligned}$$

At an investment of zero, the agency has correct beliefs about the proportion of crime detected, and therefore the expected fine imposed on offenders. The surprise is therefore equal to zero when $\theta = \theta^*$. This says that the zero surprise curve must hit the θ axis at θ^*.

Figures 3.3 and 3.4 display the zero surprise region 1, and the zero surprise curve in region 4 for two different values of θ^*, for $p(x) = x + \frac{1}{4}$ and $\hat{p}(x) = 2x + \frac{1}{4}$.

In figure 3.3, $\theta^* = \frac{1}{3}$. An interval of equilibria with over-policing exist for this θ^*, highlighted by a thick blue line. In these equilibria, the agency fully deters all crime and believes that it fully deters crime. These are the only equilibria for this value of θ^*

In figure 3.4, $\theta^* = \frac{1}{2}$. In this case, region 1 does not intersect the graph of $\hat{x}(\theta)$, so full deterrence equilibria do not exist for this value of θ^*. However, the zero surprise curve in region 4 does intersect with the downward sloping segment of $\hat{x}(\theta)$, and there is a unique, partial deterrence equilibrium at this intersection. This equilibrium is necessarily an under-policing equilibrium, because $x^*(\frac{1}{2}) > \hat{x}(\theta)$ for all θ.

3.6 Conclusion

In this paper, I propose a model and equilibrium concept for the analysis of an investment into the detection of crime by a law enforcement agency, when the preferences of the population of potential offenders is unknown, and the agency holds incorrect beliefs about the mapping from investment to detection probability. I provide sufficient conditions for the existence of an equilibrium. It is then shown that the model is able to explain how an enforcement agency may overpolice, or underpolice a population, even though its beliefs are entirely consistent with the data it observes, and it believes itself to be maximizing social welfare.

Future study of this model can hopefully clarify the conditions under which over-policing or underpolicing are likely to occur. An extension of the model to the problem of allocating a fixed budget, earmarked for detection efforts, between multiple neighborhoods or types of crime might shed further light on the possibility of self-reinforcing over-policing. This issue is critical, because intense policing incommensurate with the true needs of a community may not only be inefficient, but also unfair, since overpolicing of one sub-population necessitates, under a fixed budget, underpolicing of the rest of the population. This unfair and unequal application of the law may undermine dearly held democratic principles, and may contribute to an erosion of confidence in the law and justice system among the people.

Also note that it would be quite easy to apply this model to any enforcement situation, it need not only apply to government enforced law. For example, managers inside an organization such as a firm, political party or government agency may wish to enforce internal rules and regulations, and may exert effort to detect noncompliance. The same logic behind the model presented in this paper would equally apply in those settings.

This modeling approach might also be applied to contracting situations characterized by moral hazard with the possibility of costly monitoring. The idea in this application might be that, when one or more of the contracting agents have incorrect beliefs about their ability to detect noncompliance, or the incentives of the other parties not to comply, contracting may be inefficient, even though agents are not surprised by the outcomes of the contract they observe.

Bibliography

Eric J Allen, Patricia M Dechow, Devin G Pope, and George Wu. Reference-dependent preferences: Evidence from marathon runners. *Management Science*, 63(6):1657–1672, 2017.

Ashton Anderson and Etan A Green. Personal bests as reference points. *Proceedings of the National Academy of Sciences*, 115(8):1772–1776, 2018.

Ala Avoyan, Robizon Khubulashvili, and Giorgi Mekerishvili. Call it a day: History dependent stopping behavior. Technical report, CESifo Working Paper, 2020.

Lucian Arye Bebchuk and Louis Kaplow. Optimal sanctions when individuals are imperfectly informed about the probability of apprehension. *The Journal of Legal Studies*, 21(2):365–370, 1992.

Gary Becker. Crime and punishment: An economic approach. *The Journal of Political Economy*, 169:176–177, 1968.

Alan S Blinder and John Morgan. Are two heads better than one? monetary policy by committee. *Journal of Money, Credit and Banking*, pages 789–811, 2005.

Ivo Blohm and Jan Marco Leimeister. Gamification. *Business & information systems engineering*, 5(4):275–278, 2013.

Gary Bornstein and Ilan Yaniv. Individual and group behavior in the ultimatum game: are groups more “rational” players? *Experimental Economics*, 1:101–108, 1998.

Gary Bornstein, Tamar Kugler, and Anthony Ziegelmeyer. Individual and group decisions in the centipede game: Are groups more “rational” players? *Journal of Experimental Social Psychology*, 40(5):599–605, 2004.

Alexander L Brown, Taisuke Imai, Ferdinand Vieider, and Colin Camerer. Meta-analysis of empirical estimates of loss-aversion. *Available at SSRN 3772089*, 2021.

Timothy N Cason, Sau-Him Paul Lau, and Vai-Lam Mui. Prior interaction, identity, and cooperation in the inter-group prisoner’s dilemma. *Journal of Economic Behavior & Organization*, 166:613–629, 2019.

Aaron Chalfin and Justin McCrary. The effect of police on crime: New evidence from us cities, 1960-2010. Technical report, National Bureau of Economic Research, 2013.

Aaron Chalfin and Justin McCrary. Criminal deterrence: A review of the literature. *Journal of Economic Literature*, 55(1):5–48, 2017.

Gary Charness and Matthew O Jackson. Group play in games and the role of consent in network formation. *Journal of Economic Theory*, 136(1):417–445, 2007.

Gary Charness and Matthias Sutter. Groups make better self-interested decisions. *Journal of Economic Perspectives*, 26(3):157–176, 2012.

Gary Charness, Edi Karni, and Dan Levin. On the conjunction fallacy in probability judgment: New experimental evidence regarding linda. *Games and Economic Behavior*, 68(2):551–556, 2010.

David J Cooper and John H Kagel. Are two heads better than one? team versus individual play in signaling games. *American Economic Review*, 95(3):477–509, 2005.

James C Cox. Trust, reciprocity, and other-regarding preferences: Groups vs. individuals and males vs. females. *Experimental business research*, pages 331–350, 2002.

James C Cox and Stephen C Hayne. Barking up the right tree: Are small groups rational agents? *Experimental Economics*, 9:209–222, 2006.

Yannick Ferreira De Sousa and Alistair Munro. Truck, barter and exchange versus the endowment effect: Virtual field experiments in an online game environment. *Journal of Economic Psychology*, 33(3):482–493, 2012.

Stefano DellaVigna. Psychology and economics: Evidence from the field. *Journal of Economic literature*, 47(2):315–72, 2009.

John Duggan. Non-cooperative games among groups. 2001.

Alexander Elbittar, Andrei Gomberg, and Laura Sour. Group decision-making and voting in ultimatum bargaining: An experimental study. *The BE Journal of Economic Analysis & Policy*, 11(1), 2011.

Arpad E Elo. *The rating of chessplayers, past and present*. Arco Pub., 1978.

William N Evans and Emily G Owens. Cops and crime. *Journal of Public Economics*, 91(1-2):181–201, 2007.

Erik Eyster and Matthew Rabin. Cursed equilibrium. *Econometrica*, 73(5):1623–1672, 2005.

Drew Fudenberg and David K Levine. Self-confirming equilibrium. *Econometrica: Journal of the Econometric Society*, pages 523–545, 1993.

Nuno Garoupa. The theory of optimal law enforcement. *Journal of economic surveys*, 11(3):267–295, 1997.

Mark E Glickman. The glicko system. *Boston University*, 16:16–17, 1995.

Mark E Glickman. Example of the glicko-2 system. *Boston University*, pages 1–6, 2012.

John C Harsanyi. Games with randomly disturbed payoffs: A new rationale for mixed-strategy equilibrium points. *International journal of game theory*, 2(1): 1–23, 1973.

Paul Heidhues, Botond Kőszegi, and Philipp Strack. Unrealistic expectations and misguided learning. *Econometrica*, 86(4):1159–1214, 2018.

Greg Howard. A check for rational inattention. 2021.

Chester A Insko, Rick H Hoyle, Robin L Pinkley, Gui-Young Hong, Randa M Slim, Bret Dalton, Yuan-Huei W Lin, Paulette P Ruffin, Gregory J Dardis, Paul R Bernthal, et al. Individual-group discontinuity: The role of a consensus rule. *Journal of Experimental Social Psychology*, 24(6):505–519, 1988.

John H Kagel and Dan Levin. The winner's curse and public information in common value auctions. *The American economic review*, pages 894–920, 1986.

Daniel Kahneman. Prospect theory: An analysis of decisions under risk. *Econometrica*, 47:278, 1979.

Daniel Kahneman. Reference points, anchors, norms, and mixed feelings. *Organizational Behavior and Human Decision Processes*, 51(2):296–312, March 1992. URL `https://ideas.repec.org/a/eee/jobhdp/v51y1992i2p296-312.html`.

Louis Kaplow. Optimal deterrence, uninformed individuals, and acquiring information about whether acts are subject to sanctions. *JL Econ & Org.*, 6:93, 1990.

Louis Kaplow. Rules versus standards: An economic analysis. *Duke Lj*, 42:557, 1992.

Botond Kőszegi and Matthew Rabin. A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165, 2006.

Tamar Kugler, Gary Bornstein, Martin G Kocher, and Matthias Sutter. Trust between individuals and groups: Groups are less trusting than individuals but just as trustworthy. *Journal of Economic psychology*, 28(6):646–657, 2007.

Tamar Kugler, Edgar E Kausel, and Martin G Kocher. Are groups more rational than individuals? a review of interactive decision making in groups. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(4):471–482, 2012.

Steven D Levitt. Why do increased arrest rates appear to reduce crime: deterrence, incapacitation, or measurement error? *Economic inquiry*, 36(3):353–372, 1998.

Steven D Levitt. Using electoral cycles in police hiring to estimate the effects of police on crime: Reply. *American Economic Review*, 92(4):1244–1250, 2002.

Ming-Jen Lin. More police, less crime: Evidence from us state data. *International Review of Law and Economics*, 29(2):73–80, 2009.

John A List. Does market experience eliminate market anomalies? *The Quarterly Journal of Economics*, 118(1):41–71, 2003.

Alex Markle, George Wu, Rebecca White, and Aaron Sackett. Goals as reference points in marathon running: A novel test of reference dependence. *Journal of Risk and Uncertainty*, 56(1):19–50, 2018.

Jacob Marschak and Roy Radner. *Economic Theory of Teams.* 1972.

Brian W Mayhew and Adam Vitalis. Myopic loss aversion and market experience. *Journal of Economic Behavior & Organization*, 97:113–125, 2014.

Richard D McKelvey and Talbot Page. Common knowledge, consensus, and aggregate information. *Econometrica: Journal of the Econometric Society*, pages 109–127, 1986.

Richard D McKelvey and Thomas R Palfrey. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1):6–38, 1995.

Richard D McKelvey and Thomas R Palfrey. A statistical theory of equilibrium in games. *The Japanese Economic Review*, 47(2):186–209, 1996.

Richard D McKelvey and Thomas R Palfrey. Quantal response equilibria for extensive form games. *Experimental economics*, 1:9–41, 1998.

Andrea Morone, Simone Nuzzo, and Rocco Caferra. The dollar auction game: A laboratory comparison between individuals and groups. *Group Decision and Negotiation*, 28:79–98, 2019.

Jacob A Mortenson and Andrew Whitten. Bunching to maximize tax credits: Evidence from kinks in the us tax schedule. *American economic journal: Economic policy*, 12(3):402–32, 2020.

Elena Patel, Nathan Seegert, and Matthew Smith. At a loss: The real and reporting elasticity of corporate taxable income. *Available at SSRN 2608166*, 2017.

A Mitchell Polinsky and Steven Shavell. The optimal tradeoff between the probability and magnitude of fines. *The American Economic Review*, 69(5):880–891, 1979.

A Mitchell Polinsky and Steven Shavell. The optimal use of fines and imprisonment. *Journal of Public Economics*, 24(1):89–99, 1984.

A Mitchell Polinsky and Steven Shavell. The economic theory of public enforcement of law. *Journal of economic literature*, 38(1):45–76, 2000.

Richard A Posner. Optimal sentences for white-collar criminals. *American Criminal Law Review*, 17:409, 1980.

Matthew Rabin. Inference by believers in the law of small numbers. *The Quarterly Journal of Economics*, 117(3):775–816, 2002.

Bettina Rockenbach, Abdolkarim Sadrieh, and Barbara Mathauschek. Teams take the better risks. *Journal of Economic Behavior & Organization*, 63(3):412–422, 2007.

John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033, 1987.

John Rust. Nested fixed point algorithm documentation manual: Version 6. *Department of Economics, Yale University*, 2000.

Emmanuel Saez. Do taxpayers bunch at kink points? *American economic Journal: economic policy*, 2(3):180–212, 2010.

Yuval Salant and Jörg L Spenkuch. Complexity and choice. *Available at SSRN 3878469*, 2021.

Christin Schulze and Ben R Newell. More heads choose better than one: Group decision making can eliminate probability matching. *Psychonomic bulletin & review*, 23:907–914, 2016.

Roman M Sheremeta and Jingjing Zhang. Can groups solve the problem of overbidding in contests? *Social Choice and Welfare*, 35(2):175–197, 2010.

Janet A Sniezek and Rebecca A Henry. Accuracy and confidence in group judgment. *Organizational behavior and human decision processes*, 43(1):1–28, 1989.

Fei Song. Trust and reciprocity behavior and behavioral forecasts: Individuals versus group-representatives. *Games and economic behavior*, 62(2):675–696, 2008.

Dagmar Stahlberg, Frank Eller, Anne Maass, and Dieter Frey. We knew it all along: Hindsight bias in groups. *Organizational Behavior and Human Decision Processes*, 63(1):46–58, 1995.

Matthias Sutter. Are four heads better than two? an experimental beauty-contest game with teams of different size. *Economics letters*, 88(1):41–46, 2005.

Matthias Sutter, Martin G Kocher, and Sabine Strauss. Individuals and teams in auctions. *Oxford Economic Papers*, 61(2):380–394, 2009.

Tim Wildschut and Chester A Insko. Explanations of interindividual–intergroup discontinuity: A review of the evidence. *European review of social psychology*, 18(1):175–211, 2007.