Estimating the Preferences of Imperfectly Rational Agents: Behavioral Econometrics

by

Daniel Chi Yin Chui

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Computing Science

University of Alberta

Abstract

Modelling agent preferences has applications in a range of fields including economics and increasingly, artificial intelligence. These preferences are not always known and thus may need to be estimated from observed behavior, in which case a model is required to map agent preferences to behavior. Traditional models are based on the assumption that agents are perfectly rational: that is, they perfectly optimize and behave in accordance with their own interests. Work in the field of behavioral game theory has shown, however, that human agents often make decisions that are imperfectly rational, and the field has developed models that relax the perfect rationality assumption. In this thesis, we take a first step towards estimating agent preferences using this relaxed assumption. We apply models developed for predicting behavior towards the task of estimating preferences and show that they outperform both traditional and commonly used benchmark models on data collected from human subjects. In fact, Nash equilibrium and its relaxation, quantal response equilibrium (QRE), can induce an inaccurate estimate of agent preferences when compared against a known ground truth.

A key finding is that modelling non-strategic behavior, conventionally considered as uniform noise, is important for estimating preferences. To this end, we introduce quantal-linear4, a rich non-strategic model. We also propose an augmentation to the QRE model by incorporating a non-strategic component. We call this augmented model QRE+L0 and find an improvement in estimating values over the standard QRE. QRE+L0 allows for alternative models of

non-strategic behavior in addition to quantal-linear4.

Preface

Parts of this thesis are currently under submission to AAMAS 2023 as of November 2022. An earlier version of this same work is published on arxiv.org under the identifier 2208.06521 and is available at `http://arxiv.org/abs/2208.06521`. A non-archival version of the same paper was submitted and accepted, but not presented, at the 33rd Stony Brook International Conference on Game Theory (2022). Chapter 5 as well as sections 4.4 of chapter 4 are new work, and Chapters 1 and 2 are expanded versions of the paper under review.

The publication was authored along with my advisor Dr. James R. Wright as well as Dr. Jason Hartline from Northwestern University. My contributions included the writing of code, data collection, analysis and writing while my co-authors provided editorial and experimental oversight. These experiments were guided by my advisor James Wright as well as Jason Hartline and would not have been possible without their advice and mentorship. Michalis Mamakos of Northwestern University also contributed ideas and comments on an earlier iteration of this project, but no analysis from this earlier iteration are included in this thesis.

The experimental data analyzed in this thesis was collected with approval from the University of Alberta Research Ethics Board, Project Name ”Behavioural Econometrics” study ID Pro00096492, October 20, 2020. We thank the review board for their prompt approval of modifications to our application made during the data collection process.

To my grandmother

Who passed away during my studies.

I must study politics and war, that our sons may have liberty to study mathematics and philosophy.

– John Adams

Young man, in mathematics you don't understand things. You just get used to them.

– John von Neumann

If I could only have one more day, I could do a great [thesis].

– Hokusai

Acknowledgements

There are many people I would acknowledge. First and foremost, I extend my heartfelt gratitude towards my supervisor, James R. Wright for his mentorship, guidance, encouragement, and patience throughout my degree. I would also like to extend my heartfelt thanks to Jason Hartline, who acted as an unofficial co-advisor and introduced me to members of his lab. I would like to thank both of them for always pushing me to be a better researcher. This degree would not have been possible without their mentorship.

I would also like to acknowledge my peers in the ABGT lab both past and present. In particular, I would like to thank Ryan D'Orazio, for reading parts of this thesis and for our countless discussions about all topics including, but not limited to research. I would also like to thank Greg d'Eon, Revan Macqueen, and Niko Yasui for our interesting discussions, whether in the lab or online. In addition, I would also like to thank my peers at the University of Alberta including but not limited to: Zaheen Ahmad, Ifaz Kabir, Doug Rebstock, Chris Solinas, Maliha Sultana, and Kristen Yu for helping to make my grad school experience enjoyable. I would like to thank both the National Research Council of Canada (NSERC) and the Alberta Machine Intelligence Institute (Amii) for providing funding for this work, as well as the government of Alberta for the Alberta Graduate Excellence Scholarship.

I am also grateful to my friends Kevin Au, Adoria and Kevin Chan, Kirby Chau, Jenn Herbert, Zubair Kharadi, Charles Lanahan, and Andrea Law for their continued friendship over the years and for reminding me of life outside of grad school, especially during these tumultuous times.

Finally, I would like to thank my family, including D & K, and my entire extended family for their support throughout the years.

Contents

1 Introduction **1**

2 Background **5**

2.1 Framework 5
2.1.1 Normal Form Games 5
2.1.2 Nash Equilibrium 6
2.1.3 Quantal Response Equilibrium 6
2.2 Behavioral Game Theory 7
2.2.1 Level-k and Cognitive Hierarchy 8
2.3 Studying Initial Play and Level-0 Models 10
2.4 Inferring Preferences from Behavior 11
2.5 Online Experiments 12

3 Experimental Setup and Contributions **14**

3.1 Quantal Linear-4 14
3.2 Separating Models into Strategic and Non-strategic Components and QRE+L0 16
3.3 Experimental Setup 17
3.3.1 Allocation Games 18
3.3.2 Game Construction 18
3.3.3 Experimental Details 19
3.4 Estimation Methods 20
3.4.1 Estimation of Equilibrium Models 21
3.4.2 Estimation of Poisson Quantal Cognitive Hierarchy . . 23
3.4.3 Utilization of Panel Structure in Estimation of Values . 23

4 Results **25**

4.1 Evaluation 25
4.2 Contribution of Strategic vs. Non-strategic Components of the Model 27
4.3 Identification of Quantal-linear4 29
4.4 Alternate Treatments 31
4.4.1 Evaluation of ALL10 Dataset 32
4.4.2 Differences in Alternate Treatments 33

5 Theoretical Results **34**

5.1 Bounding the Error in our Value Estimate 34
5.2 Empirical Results in Treatments 36

6 Conclusion **38**

6.1 Conclusions 38
6.2 Future Work 38

References 40

Appendix A Appendix 44

- A.1 Additional Experimental Details . . . 44
  - A.1.1 Results for Different Treatments . . . 44
- A.2 Experimental Details . . . 47
- A.3 Choosing λ for Our Nash Approximation . . . 47
- A.4 Allocation Game Mapping Algorithm . . . 50
- A.5 Additional Figures and Tables . . . 51
  - A.5.1 Behavioral Parameter Estimates . . . 51

List of Tables

3.1 A summary of the strategic and non-strategic components included in our evaluation and the parameters θ for each component. 21

4.1 Relative error by v^*, with confidence interval in parentheses. Bold cells indicate best performing model for each v^*. Italicized cells indicate models which are not significantly different from the best performing one. QRE-QL4 indicates a model in which a fraction of agents are behaving non-strategically in a QL4 manner while the remaining agents are in QRE with themselves and non-strategic agents. None-QL4 indicates a model in which all agents are behaving non-strategically. 28

4.2 Relative error of predicted average per game welfare by v^*. In each scenario, the estimated valuation and model parameters θ from half the games are used to predict the average game welfare per subject on the other half and is compared against the empirically observed average welfare. Bold cells indicate models with the lowest MSE. 28

4.3 Percentage of the time that relative error is less than 10% across all values and scenarios in $\mathcal{V}$. Each cell corresponds to a model STRAT-NONSTRAT with the row indicating the strategic model and column indicating the non-strategic one. QL4 (rightmost column) outperforms all other non-strategic components regardless of the strategic model. Here, the confidence intervals are derived from a k bootstrapped samples of the observed data, with $k = 1000$. Cells marked "n/a" do not have a conceivable model that elicits an estimate of values. Cells containing 0 mean that none of the bootstrapped samples had a value estimate that fell within 10% of v^*. 30

5.1 Percentage of the time the absolute relative error is lower for the scenario with the highest minimum allocation difference vs. the lowest minimum allocation difference. 36

A.1 Relative error by v^* on randomized filtered treatment 45

A.2 46

A.3 Relative error by v^* on ordered nonfiltered treatment 46

A.4 Relative error by v^* on combined filtered treatment 46

A.5 Relative error by v^* on combined nonfiltered treatment 47

A.6 Summary Statistics for Experimental Treatments 47

A.7 Estimated τ when using QCH-QL4, with τ indicating the rate parameter for a Poisson distribution specifying the proportion of agents of level k. 51

A.8 Estimated β when using QRE-$QL4$, with β indicating the proportion of agents who are non-strategic 51

A.9 Raw value estimates when allocation games contain no payments for PQCH-QL4. 51
A.10 Comparison of estimated v^* and λ and β behavioral parameters when finding the parameter(s) which maximize the likelihood (QRE Likelihood) vs. computing QRE against the empirical distribution (QRE Empirical). T-distributed confidence interval in brackets 52

List of Figures

3.1 A visualization of our allocation algorithm mapping a payoff game to an allocation game 19

4.1 Summary plot showing values of v^* vs. the relative error. The two left-most points for all values of v^* are models containing both a strategic component and quantal-linear4 as the non-strategic component; the green point indicates a model that assumes all agents are non-strategic in a quantal-linear4 manner with no strategic behavior. 26
4.2 Synthetic data: estimated behavioral parameters vs. preferences v^* in PQCH. Asterixes indicate ground truth parameters, with each intersection representing the parameters of the synthetically generated dataset. We chose $\lambda = 0.05$ to approximate the precision in an empirical dataset, with the non-strategic agents using a maxmax strategy, where non-strategic agents pick the action that maximizes their best case payoff. τ indicates the estimated Poisson mean parameter for the distribution over levels, which seems to be biased upwards (likely due to it being truncated) but does not affect the estimate for v. 31
4.3 Summary plot showing values of v^* vs. the relative error for the ALL10 dataset. As the number of samples for ALL10 is significantly smaller, the confidence intervals are correspondingly wider, especially for models with UNIFORM as the non-strategic model 32

A.1 Summary plot for randomized filtered treatment 45
A.2 Summary plot, ordered filtered treatment 45
A.3 Summary plot, ordered nonfiltered treatment 45
A.4 Summary plot, combined filtered treatment, $n = 484$ 46
A.5 Summary plot, combined nonfiltered treatment, $n = 364$ . . . 46
A.6 The main experiment webpage presented to MTurk Participants. 48
A.7 The screening quiz presented to MTurk Participants to test their understanding of the task. Participants were allowed 3 attempts on the quiz before being rejected for the HIT. 48
A.8 The exit survey presented to MTurk Participants once they complete their HIT. The second prompt directs participants to fill out the reasoning behind their decisions. 48
A.9 QRE with fixed values of λ. As the error stabilizes around $\lambda = 100$, we use this as our Nash approximation 49

Chapter 1

Introduction

A fundamental assumption underlying modelling agent behavior is that agents have preferences which drive this behavior. The practice of estimating these preferences from their behavior goes by many names including but not limited to inverse game theory, intrinsic motivation, and revealed preferences, and has applications in economic analysis, the study of human behavior, and increasingly, artificial intelligence [30]. When in a situation where an agent is incentivized to be truthful and preferences can be mapped directly to behavior, backing out the preferences of an agent is straightforward. One such example would be certain forms of auctions, in which an agent's private valuation for an item drives their bid. In contrast, when preferences are not straightforwardly mapped towards an action, or it may be in the agent's best interest to not report their true preference (i.e. they are in a strategic situation), the problem becomes much harder. Popular card games such as bridge or poker would be an example of such a situation. The task of estimating agent preferences in such a situation then requires a model mapping preferences to behavior, commonly referred to as structural estimation in the economics literature.

Structural estimation has been used widely in empirical economics, or econometrics, to model relationships and inform decision making. The decision making process often involves constructing incentives (e.g. public policy, UX design, or rewards) to induce desired behavior and is known as mechanism design. A set of incentives for strategic agents is referred to as a *mechanism*. One key application of structural estimation is in counterfactual estimation of alternative mechanisms. The accuracy of counterfactual estimation, however, relies on the accuracy of the model. After inferring preferences from empirical data, counterfactual scenarios can be evaluated. Such evaluation can be used in mechanism design for optimizing over many mechanisms to find one with the best performance. Randomized controlled trials – called A/B testing by technology firms – can directly evaluate a novel mechanism, but the number of samples required for a given mechanism can be quite large. Chawla et al. [9] showed that methods from structural inference have an exponential improvement for sample complexity in mechanism design over randomized controlled trials, but this improvement comes with the aforementioned reliance on the accuracy of the model.

A common assumption in economics and game theory is that agents perfectly optimize. When all agents are behaving optimally given the behavior of other agents, they are said to be in equilibrium. A strong equilibrium assumption, that all agents are perfectly optimizing given the behavior of other agents, is commonly made when inferring preferences from empirical behavior. In instances approximating equilibrium, such as repeated interactions between firms [45] or for individual agents repeatedly playing the same game, this is a reasonable assumption. In situations in which agents are not in equilibrium, either due to the observed behaviors being early in the decision making process or from an unstable situation without an equilibrium, this can pose a bigger problem. Particularly in non-truth telling indirect mechanisms (i.e. mechanisms in which agents are behaving strategically, and in which there is no straightforward mapping of preferences to behavior), it is possible for agents with differing valuations to produce the same observed behavior, especially when the action space is small (e.g. in a binary decision process, an infinite number of agent valuations would map to one of only two possible outcomes).

The more recent field of behavioral game theory has shown that rich, parameterized models of behavior, including non-strategic behavior, provide better predictions of empirical behavior than classical equilibrium models. Particularly in initial play, i.e., for behavior of players who do not have prior experience playing a given game, classical notions of equilibrium are poor predictors of human behavior while behavioral models offer an improvement in predictive performance. The current result of these efforts are behavioral models in the quantal cognitive hierarchy (QCH) family, which, in addition to performing better than classical equilibrium models, also outperform other behavioral models [8, 35] in terms of predicting behavior. A key component of QCH is the concept of iterated reasoning, which requires some definition of non-strategic (level-0) agents, conventionally defined to be uniform randomizers [23]. Wright and Leyton-Brown [48] showed that going beyond the uniform randomization assumption improves performance when predicting behavior. The uniform randomization assumption also has another drawback: it is insensitive to preferences. As past empirical works have shown that a significant proportion of agents are level-0 [47], not modelling the preferences of these level-0 agents means that we are in essence discarding information given by data being generated by them.

This thesis contributes to the study of behavioral models for initial play and the study of structural inference of preferences from behavior in games. We develop behavioral models and conduct online experiments within the context of initial play to measure the accuracy of these models in inferring preferences. Our experimental analysis introduces a new level-0 model which is derived from adding quantal response to the linear4 level-0 model from Wright and Leyton-Brown [48]. The model that best predicts behavior and admits the most accurate inference of values is quantal cognitive hierarchy with this quantal linear4 level-0 model. We compared this model to classical equilibrium and behavioral models without rich level-0 behavior of Nash equilibrium, quantal response equilibrium [35], and quantal cognitive hierarchy (with uniform level-0 behavior). We also introduce the concept of considering models of behavior to contain both a strategic and a non-strategic component. Under this framework, existing equilibrium models such as quantal response equilibrium (QRE) can be augmented with non-strategic models, including but not limited to the aforementioned quantal-linear4 model. Models containing the new quantal-linear4 non-strategic specification outperform these classical models with Nash equilibrium being the worst at both inference and prediction.

Our experiments highlight the importance of rich level-0 models for modeling behavior in initial play. Our results on predicted behavior reinforce those of Wright and Leyton-Brown [48], showing that rich models of level-0 behavior are better predictors than uniform randomizers. Moreover, and intuitively, these models take into account payoffs and, thus, the inferred level-0 behavior aids in the inference of preferences. In fact, we find that the level-0 model drives most of the gains in both predicting behavior and inferring preferences, and the choice of strategic model is not as important.

Our experimental setup considered 3-by-3 bimatrix games with randomly generated payoffs. This family of games is commonly studied in the behavioral game theory literature [35, 38]. We assumed payoffs were derived from the classical single-dimensional linear model of auction theory where payoffs are given linearly in a value for units of a good (i.e., allocations) and a payment [34]. (Our games allow payments to be negative, i.e., some payoffs are given by some units of the good and a negative amount of money.) A key simplification of our experiment is that the players in our experiments were only aware of the payoffs in the game and not of the decomposition of those payoffs into allocation and payments. Thus, we do not see in our data behavioral artifacts related to whether or not the players can do the utility calculations from allocations and payments. Moreover, with such a design we are free in our analysis to consider counterfactual inference questions with various decompositions of payoffs into allocations and payments.

The rest of this thesis is organized as follows. Chapter 2 provides background on models from behavioral game theory and related works in the task of inferring preferences and studying initial play. In Chapter 3 we describe the novel models developed as well as our experimental setup, and report the results in Chapter 4. We then present a preliminary theoretical analysis of a surprising result of our experiments in Chapter 5. We offer a conclusion and possible directions for future work in 6.

Chapter 2

Background

In this chapter, we provide relevant background material for the thesis. First, we layout a framework of game theoretic concepts. Next, we provide background on existing behavioral game theory models as well as past works on inferring preferences from behavior and the study of initial play. Finally, we go over the existing methodology for collecting datasets online.

2.1 Framework

2.1.1 Normal Form Games

A common abstraction of strategic interactions between agents is the *Normal Form Game*.

Definition 2.1.1 (Normal Form Game)**.** A *normal form game* G consists of a tuple $G = (N, A, u)$ where

1. $N = (1, \ldots, n)$ is a set of N agents
2. $A = A_1 \times \cdots \times A_n$ is a set of *action profiles*, where A_i is the set of actions available to player i, and $a \in A = (a_1, ..., a_n)$ is a tuple consisting of an action a_i for each agent
3. $U = (u_1, \ldots, u_n)$ is a set of utility functions $u_i : A \to \mathbb{R}$, mapping each action profile to a utility for each player

A *mixed strategy* $s_i \in \Delta(A_i)$ for player i is a distribution over i's actions. The utility of a mixed strategy profile $s \in \Delta(A_1) \times \cdots \times \Delta(A_n)$ is the expected utility of an action profile sampled from the product distribution of the mixed strategies.

2.1.2 Nash Equilibrium

The classical solution concept in game theory is Nash equilibrium, first discovered by Nash in the 1950s [36] and is guaranteed to exist for every a normal form game. In a Nash equilibrium, every agent is best responding to the actions of the other agents. Formally,

Definition 2.1.2 (Nash equilibrium)**.** Let $BR_i(s_{-i}) = \{s_i \in \Delta(A_i) \mid u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i}) \forall s'_i \in \Delta(A_i)\}$ be the set of best responses to s_{-i}. Then a mixed strategy profile s is a *Nash equilibrium* if every agent i's mixed strategy s_i is a best response to the profile s_{-i} of mixed strategies of the other agents: $s_i \in BR_i(s_{-i})$.

2.1.3 Quantal Response Equilibrium

Experimental evidence shows that human subjects will play actions in normal form games that have zero probability under a Nash equilibrium [18, 47]. One popular explanation for this is that rather than responding to utility $u_i(a_j)$, agents are instead responding to some utility $u_i\hat{(a_j)} = u_i(a_j) + \epsilon_{a_j}$ where $u_i(a_j)$ is the utility modelled/observed by the analyst and ϵ_{a_j} is a random variable representing some unobserved/unmodelled utility or exogenous shock. When ϵ is distributed according to the Gumbel (extreme value) distribution, then the probability of taking a given action takes the form:

$$Pr(a_j) = \frac{\exp[\lambda \cdot u_i(a_j)]}{\sum_{a'_j \in A_j} \exp[\lambda \cdot u_i(a'_j)]}. \tag{2.1}$$

as shown in Train [44]. Here, λ denotes the sensitivity of agents to utility preferences, with $\lambda = 0$ corresponding to complete indifference between utilities. This formula is referred to in the decision theory literature as logit and is analagous to the softmax function commonly used in machine learning. We refer to it here as quantal best response.

Definition 2.1.3 (Quantal best response)**.** Let $u_i(a_i, s_{-i})$ be agent i's expected utility when playing action $a_i \in A_i$ against mixed strategy profile s_{-i} in game G. Then a *quantal best response* $QBR_i(s_{-i}; G, \lambda)$ by agent i to s_{-i} is a mixed strategy s_i such that

$$s_i(a_i) = \frac{\exp[\lambda \cdot u_i(a_i, s_{-i})]}{\sum_{a'_i \in A_i} \exp[\lambda \cdot u_i(a'_i, s_{-i})]}. \tag{2.2}$$

When every agent $i \in N$ is quantally best responding to the strategies of other agents, we then have a quantal response equilibrium:

Definition 2.1.4 (Quantal response equilibrium)**.** A strategy profile s of a game G is a *quantal response equilibrium* (QRE) with precision $\lambda > 0$ when each agent quantally best responds to the strategies of the other agents; that is, when $s_i = QBR_i(s_{-i}; G, \lambda)$ for all agents $i \in N$.

As $\lambda \to \infty$, the quantal best response for agent i coincides with best response, and so Nash equilibrium can be viewed as a special case of quantal response equilibrium.

Quantal response equilibrium is attractive as a solution concept because of its ability to fit a wide range of behavior. In fact, Haile et al. [22] show that without additional restrictions, QRE does not falsify any behavior in a single normal form game (that is, any behavior can be explained as a QRE). It is important to note that other forms of QRE can exist, but logit equilibrium (as defined previously) is by far the most commonly used form.

2.2 Behavioral Game Theory

It has long been observed that humans do not behave in a fully rational manner. The famous economist John Maynard Keynes wrote of "animal spirits" which influence human behavior in the 1930s [28]. Examples of behavior inconsistent with expected utility theory was shown by the Allais Paradox [2]. The work of Kahneman and Tversky in the 1970s, culminating in the model of prospect theory [27], marked the beginning of the field of behavioral game theory. Behavioral game theory aims to produce models which better predict empirical human behavior compared to traditional game theoretic solution concepts such as Nash equilibrium.

Within the field, the aforementioned quantal response equilibrium [35], relaxes the strict optimization assumption made by Nash equilibrium, while maintaining the assumption that agents mutually respond to each others' strategies. This relaxation has been found to produce a much better fit than Nash equilibrium and has been suggested as a replacement for Nash when fitting behavioral data [22].

Another line of work is that of iterative behavioral models such as level-k [43] and cognitive hierarchy [8] models. These models assume that agents perform a fixed number of iterations of strategic reasoning, starting from a default strategy called the level-0 strategy. Wright and Leyton-Brown [47] found that incorporating the quantal error model into a cognitive hierachy, or quantal cognitive hierarchy, performs best at predicting actual behavior in empirical data. In later work, they showed that prediction performance can be further improved by specifying parameterized level-0 models that combine simple decision rules, instead of the uniform randomization specification that is most frequently studied [48]. We now discuss each of the different behavioral models below.

2.2.1 Level-k and Cognitive Hierarchy

Another key contribution from behavioral game theory is to assume that humans are performing iterated levels of reasoning. The level-k model [13, 15] models agents as having a level $k_i \in 0, 1, 2, ...$ in which a level-k agent best responds to the strategy of level-$k-1$.

Definition 2.2.1 (Level-k thinking)**.** In a level-k model, each agent i with level k is best responding to the strategy profile level-$(k-1)$ Formally,

$$s_{i,k} = BR_i(s_{-i,k-1})$$

where the first level with no reasoning is considered level-0, and is conventionally considered uniform randomization over the support, $s_{-i,0} = \frac{1}{|A|}$.

The cognitive hierarchy model of Camerer et al. [8] also models heterogeneous levels of reasoning, but differs in that level-k agents respond to the distribution over lower level types rather than only the level immediately below k. When the agents respond in a quantal best response fashion, we end up with *Quantal Cognitive Hierarchy.*

Quantal cognitive hierarchy (QCH) is a non-equilibrium model, in which agents are heterogeneous in the number of steps of strategic reasoning they can perform. Higher-level agents choose their actions in response to the strategies of lower-level agents. The lowest level agents (level-0 agents) choose their actions non-strategically; that is, without reasoning about the actions of the other agents. Level-0 agents are commonly specified to simply play a uniform distribution over actions; we evaluate that specification, but we also evaluate QCH using a richer specification of level-0 behavior (see Section 3.1, below).

Definition 2.2.2 (Quantal cognitive hierarchy)**.** Quantal cognitive hierarchy with precision $\lambda > 0$, level distribution L, and level-0 specification f, specifies that each agent i has a level $k_i \sim L$. Let $\pi_{i,k} \in \Delta(A_i)$ be the distribution over actions predicted for an agent i with level k. Level-0 agents play actions according to $\pi_{i,0} = f(G)$, where f is some non-strategic function of the game payoffs. Agents with level $k > 0$ play according to the distribution $\pi_{i,k} = QBR_i(\pi_{-i,0:k-1}; G, \lambda)$, where

$$\pi_{i,0:k} = \frac{\sum_{\ell=0}^{k} L(\ell) s_{i,\ell}}{\sum_{\ell'=0}^{k} L(\ell)}$$

is the distribution over actions induced by conditioning on the level being at most k.

The overall distribution of actions predicted by quantal cognitive hierarchy is $\pi_i = \sum_{k=0}^{\infty} L(k) \pi_{i,k}$.

For the distribution of levels in QCH, a truncated Poisson distribution is commonly used [e.g. 8, 16].

Definition 2.2.3 (Poisson quantal cognitive hierarchy)**.** Poisson quantal cognitive hierarchy is a specification of QCH in which the level distribution L is specified by a truncated Poisson distribution with the mean parameter τ:

$$L_{\tau;0:k} = \sum_{\ell=0}^{k=M} \frac{\text{Poisson}(\ell;\tau)}{\sum_{\ell'=0}^{k} \text{Poisson}(\ell';\tau)}$$

where $L_{\tau;\ell}$ is the proportion of agents at level ℓ given mean τ and with $L_{\tau;0:k}$ sums to 1.

2.3 Studying Initial Play and Level-0 Models

Initial play in matrix games has generally been studied with the aim of modelling the behavior, and not for the task of value estimation. Initial play refers to agents playing normal form games for a short duration before any learning has been able to take effect or equilibrium has been reached. A common assumption made in these works is that initial play are the actions of a level-1 agent, with level-0 behavior consistently thought of as uniform randomization [14, 17]. Under this assumption, the first level of reasoning which is of interest is naturally level-1. With the advent of AI, another popular approach in recent years has been to model initial play using machine learning techniques [17, 24, 41]. These models often prove to be better predictors of behavior but lack a straightforward way to map behavior to preferences.

Of works that do attempt to specify a level-0 model, one approach is to apply the principle of saliency to the payoff structure[11, 15], which runs into the difficulty of defining a precise saliency rule(s), and furthermore deciding on which rule is most salient. A more systematic approach is that of Wright and Leyton-Brown [48], which constructs level-0 functions meeting a definition of non-strategic laid out in a companion paper [49]. The definition of non-strategic behavior is that which does not take into account the actions of other agents. This level-0 framework is attractive as it allows for the a definition of level-0 behavior that is systematic in its sensitivity to preferences, as well as allowing for an arbitrary definition of level-0 so long as it satisfies the definition of non-strategic behavior. This allows us to construct a model of level-0 behavior sensitive to their own preferences.

2.4 Inferring Preferences from Behavior

The task of inferring preferences from behavior has generally been studied under the game theoretic assumption that players are in equilibrium [e.g. 3, 4, 20, 39]. In cases where the equilibrium assumption has been relaxed, this has generally been under the condition of repeated play (i.e. the subjects play the same game(s) repeatedly). Crawford and Iriberri [15] and Goeree et al. [19], use non-equilibrium behavioral models (level-k thinking) to explain a widely-observed behavioral phenomenon—overbidding in private-value auctions—that is inconsistent with the bidders' being in equilibrium. However, their experimental evaluation focuses on estimating parameters of the behavioral model only, taking the values as known to the analyst. Nekipelov et al. [37] estimate private values from auction data without an equilibrium assumptions, instead relying on a weaker assumption that agents use some form of no-regret learning. Similarly, Ling et al. [31] provide a framework to learn game parameters from actions in zero-sum games, but do not validate their results on empirical data.

Perhaps the work that most closely resembles our own is that of Noti [38], which has a similar objective of inferring preferences from empirical play in normal form games where the values are known but hidden from the analyst. Our work differs in one key aspect, however; whereas Noti attempts to estimate values using player responses over repetitions in a single game, agents in our scenario only see each game once. None of these aforementioned works study value estimation in initial play; each relies upon repetition across games.

Value estimation has also been empirically studied under conditions resembling initial play within the field of school matching. Value estimation is necessary for counterfactual evaluation of mechanisms and there are several papers [e.g. 1, 7, 25, 26] which attempt to infer preferences of agents to evaluate the welfare of alternative mechanisms. The way in which preferences are modelled vary from an equilibrium model to assuming all agents use simple behavioral rules. Notably, Calsamiglia et al. [7] construct a model of strategic and non-strategic agents in which strategic agents best-respond noisily to all other agents, including non-strategic agents, similar in principle to our QRE+L0 model. Whereas non-strategic agents directly report their true preferences in the school choice setting, our framework allows us to consider scenarios in which there is an indirect mechanism mapping the preferences of non-strategic agents to actions.

2.5 Online Experiments

The usage of Amazon Mechanical Turk (MTurk) or other online crowdsourcing platforms as a source of experimental data in lieu of university laboratory experiments has become increasingly prevalent. MTurk workers (commonly referred to as Turkers) have several advantages over laboratory experiments. First, the demographic for university laboratory experiments generally consist of university students, whereas Turkers theoretically consist of anyone with access to a computer and an internet connection, although in reality the demographic of the average turker skews slightly younger and female [33]. Additionally, the subject pool for turkers remains constant throughout the year in contrast to university subject pools, which are typically largest in the fall. Second, Turkers are willing to work for lower wages than university subjects, allowing for the collection of larger amounts of data at the same price point. Finally, Turkers have a system of requirements and qualifications which allow researchers to break down their subjects by location, political affiliation, or other demographics.

Given the anonymity of the platform, as well as the comparably lower payments, questions about the validity and extensibility of results gathered from online platforms has arisen. Another concern is the competency of Turkers: they may not be qualified to perform the task assigned. Finally, concerns arise about the morality of conducting experiments online: university experiments require informed consent, which is more difficult to enforce online.

There are a number of works that address these concerns [33, 42]. On the issue of lower payments, studies have found that the quality of work produced between laboratory experiments and Turkers is comparable [29, 40], with a possible reason being that the threshold for particpating online is much lower than participating in a physical laboratory experiment. On the issue of competency, Amazon itself address this through the Masters qualification, which is awarded by Amazon through an unknown process. Masters workers are purported to be more reliable and better at their tasks and the fee Amazon charges the requester is slightly higher for master workers compared to a regular workers. Staffelbach et al. [42] found in a complex crowdsourcing engineering task that workers with a $> 95\%$ approval rating and 1000 HITs completed performed comparably to masters workers. Similarly, Loepp and Kelly [32] did a comparative study of masters workers and set the criteria of their regular workers to have a $> 90\%$ approval rating and 100 HITs completed. Overall, the consensus seems to be that as long as Turkers are paid a sufficient amount, a verification process is implemented to test their understanding of the task, and turkers are give their informed consent in accordance with university guidelines, online experiments offer a viable alternative to laboratory experiments.

Chapter 3

Experimental Setup and Contributions

In this chapter, we outline the main contributions of this thesis: the definition of quantal-linear4, the separation of models into strategic and non-strategic components and the resulting augmented behavioral model QRE+L0. We then provide details on the experimental setup we used to evaluate these models against the benchmark. Finally, we provide implementation details of our evaluated models.

3.1 Quantal Linear-4

As noted previously, it is standard in the literature to assume that non-strategic agents randomize uniformly over their actions. Recently, Wright and Leyton-Brown [48] found that using a linear combination of simple decision rules as a level-0 specification markedly improves the prediction performance of QCH. In this model, called linear4, each decision rule identifies an action from A_i that optimizes some simple criterion (e.g., maximizing the sum of all players' utilities), and predicts that player i will play that action.[1] The predictions of the simple decision rules are then linearly combined into an overall prediction, using weights that are free parameters of the model.

We evaluate a level-0 model adapted from linear4 that we refer to as quantal-linear4. The key difference between the two models is that in quantal-

[1] In the case of ties, the decision rule predicts a uniform distribution over the criterion-optimizing actions.

linear4, each decision rule computes its prediction as a quantal response to the different actions' criterion values. In contrast, in linear4, the predictions are computed using strict optimization—each decision rule assigns probability 0 to each action that does not optimize its criterion. This extension is motivated by two considerations. First, behavioral models that assume quantal response to preferences have tended to predict better than equivalent models based on strict optimization: QRE predicts better than Nash equilibrium, QCH predicts better than cognitive hierarchy, and the level-k model using quantal response predicts better than level-k using best response [47]. It is thus natural to expect that modeling non-strategic agents as responding quantally will also improve prediction performance. Second, the likelihood for linear4 is continuous in the weights of the decision rules (i.e., in its behavioral parameters), but discontinuous in the valuation parameter. This leads to poor optimization performance when attempting to learn the agent valuations. In contrast, the likelihood for quantal-linear4 is continuous and differentiable in both its behavioral parameters and the valuation.

Definition 3.1.1 (Quantal-linear4)**.** A quantal-linear4 (QL4) strategy for a player i in a game G with precision $\lambda_0 > 0$ and weights $w_{\max}, w_{\min}, w_{\text{eff}}, w_{\text{fair}}$ is a linear sum of the form

$$f_i(G) = \sum_{d \in \{\max,\min,\text{eff},\text{fair},\text{unif}\}} w_d f_i^d(G),$$

where the weights are constrained to lie between 0 and 1 and to sum to exactly 1.

Each function f is a soft maximization over a specific feature for each action. The features are: the maximum utility that i can receive by playing an action; the minimum utility that i can receive by playing an action; the smallest unfairness attainable by playing an action (defined as the difference between the smallest utility and the largest; this is always negative); and the largest sum of utilities that is possible by playing a given action.[2] Formally,

$$
\begin{aligned}
f_i^{\max}(G)(a_i) &= \frac{\exp[\lambda_0 \max_{a_{-i}\in A_{-i}} u_i(a_i, a_{-i})]}{\sum_{a_i'\in A_i} \exp[\lambda_0 \max_{a_{-i}\in A_{-i}} u_i(a_i', a_{-i})]}\\
f_i^{\min}(G)(a_i) &= \frac{\exp[\lambda_0 \min_{a_{-i}\in A_{-i}} u_i(a_i, a_{-i})]}{\sum_{a_i'\in A_i} \exp[\lambda_0 \min_{a_{-i}\in A_{-i}} u_i(a_i', a_{-i})]}\\
f_i^{\text{fair}}(G)(a_i) &= \frac{\exp[\lambda_0 \max_{a_{-i}\in A_{-i}} \min_{j,j'\in N}(u_j(a_i, a_{-i}) - u_{j'}(a_i, a_{-i}))]}{\sum_{a_i'\in A_i} \exp[\lambda_0 \max_{a_{-i}\in A_{-i}} \min_{j,j'\in N}(u_j(a_i', a_{-i}) - u_{j'}(a_i', a_{-i}))]}\\
f_i^{\text{eff}}(G)(a_i) &= \frac{\exp[\lambda_0 \max_{a_{-i}\in A_{-i}} \sum_{j\in N} u_j(a_i, a_{-i})]}{\sum_{a_i'\in A_i} \exp[\lambda_0 \max_{a_{-i}\in A_{-i}} \sum_{j\in N} u_j(a_i', a_{-i})]}\\
f_i^{\text{unif}}(G)(a_i) &= \frac{1}{|A_i|}.
\end{aligned}
$$

3.2 Separating Models into Strategic and Non-strategic Components and QRE+L0

Recalling that quantal cognitive hierarchy requires a non-strategic model in its inductive definition of behavior, it is straightforward to combine QCH and non-strategic models. Expanding upon the concept of heterogeneous reasoning types, equilibrium models such as Nash equilibrium and quantal response equilibrium can also be augmented with non-strategic behavior. To combine equilibrium strategic models with a non-strategic component, we assume that some fraction of agents behave non-strategically, and that the strategic agents respond to this probability of non-strategic behavior as well as the behavior of the remaining probability of strategic agents.

We separate each of our models into a non-strategic component and a strategic component that is responding to the non-strategic component, where each model is denoted by the naming convention "STRAT-NONSTRAT". In this way, conventional models such as PQCH can be rethought of as PQCH-uniform, and QRE can be rethought of as QRE-none (for the sake of simplicity and in keeping with convention, we do not list the non-strategic component in a model if there is none and so QRE-none remains QRE). For equilibrium

[2] These decision rules are not meant to be necessary and sufficient; rather, these were 4 most predictive features as discovered by Wright and Leyton-Brown. We extend the model of linear4 and leave the problem of identifying a necessary and sufficient model for future work.

models we augment with a non-strategic component, we assign the parameter $\beta \in (0, 1)$, to the probability of agents behaving non-strategically, with the strategic agents being assigned the remaining probability $1 - \beta$.

Definition 3.2.1 (QRE+L0)**.** A QRE+L0 with precision $\lambda > 0$ and level-0 specification f^{L0} is a strategy profile s in which each agent plays according to the level-0 specification with probability β, and with probability $1-\beta$ quantally best responds to the strategies of all other agents. Formally, for every agent i,

$$s_i = \beta f_i^{L0}(G) + (1 - \beta)QBR(s_{-i}; G, \lambda).$$

Unlike QCH, in which agents have heterogeneous and incorrect beliefs about the strategies of the other agents, the strategic agents in our equilibrium models augmented in this way are assumed to have correct beliefs.

3.3 Experimental Setup

The analyst's objective is to predict how much participants value a unit of some good, given their behavior in a set of normal form games. Each player i receives both an allocation of x_i units of the good, and a payment p_i in currency. We assume that player i's utility is linear in both payments and allocations; i.e., player i's utility is $u_i(x, p) = vx_i + p_i$, where $v \in \mathbb{R}$ is i's value (in currency units) for each unit of the good. In this work, we will assume that the valuation v is common across all players.

It is challenging to translate this setting into an experiment. The main challenge is that we need to endow our experimental participants with a specific value for the good, which is common knowledge across all participants. Presenting participants in the experiment with a valuation is not sufficient: participants may not believe the valuation presented to them (i.e. they will believe the purpose of the study is something other than what is presented to them), or behavioral issues (e.g., arithmetic errors, magnitude effects) may arise. To address these issues, we translate our setting into an experiment in a slightly less direct way. Instead of presenting the outcomes of a game as a decomposition of units of good allocated and units of payment to the participant, we instead present the induced utilities. That is, we perform the arithmetic for the participants of converting from an allocation and a payment to a utility. We then map the behavior observed in these translated games (which we refer to as *payoff games*) to a utility-equivalent decomposed game (which we refer to as *allocation games*), and perform our analysis as if the players had chosen their actions in the allocation games. Notice that, for any given payoff game, any number of utility-equivalent allocation games can be constructed. As we will see, this allows us to repeat our analysis for different games and even different valuations using the same set of observations. We provide a more detailed explanation below.

3.3.1 Allocation Games

We first define an *allocation game* in which each action profile maps to an *allocation profile* $x \in R^n$ and a *payment profile* $p \in \mathbb{R}^n$. Player i's utility is quasilinear, with $u_i(x, p) = vx_i + p_i$, where $v \in \mathbb{R}$ is i's valuation for the goods being allocated. The valuation v is a common value and is shared between agents.

An allocation game has a corresponding *payoff game*, a normal form game in which action profiles map directly to a scalar utility for each player, rather than being specified via a decomposition into allocation and payoff.

3.3.2 Game Construction

In our experiment, n participants play a set of bimatrix payoff games $\mathcal{G}$. To simulate our participants playing a set of allocation games $\mathcal{A}$, we map each payoff game the participant plays to a corresponding allocation game based on an endowed valuation v^* of our choosing.

To transform from the payoff games presented to the participants to the desired allocation games, we construct an allocation game in the following way:

We first select an endowed value v^* that is hidden from the models we evaluate. Then, for every cell in every payoff game $G \in \mathcal{G}$:

1. We sample an allocation from a uniform distribution bounded between 0 and $\max(u(G))/v^*$.

2. We add a payment $p \in \mathbb{R}$ such that the payoffs for each player and action match that in the original payoff game. Note that in our setup, payments can be negative.

This setup resolves the aforementioned issues with presenting allocation games to participants directly. Since only the utilities are presented to participants, we are able to abstract away from other effects such as arithmetic errors that might arise from having participants play the actual allocation game. This setup also allows us to specify an infinite number of games for any known v^*, which is useful because it allows us both to reuse the same dataset for multiple sets of allocation games, and also to specify a ground truth value with which to evaluate model performance. A more detailed pseudocode explanation can be found in the appendix in Section A.4.

	L	***M***	***R***
X	3, 4	1, 1	5, 0
Y	1, 3	2, 4	4, 3
Z	-1, 5	0, 6	4, 5

🍎 = 2

	L	***M***	***R***
X	(🍎, 1), (🍎 🍎)	1, 1	(🍎🍎, 1), 0
Y	(🍎, -1), (🍎, 1)	(🍎), (🍎, 2)	(🍎🍎🍎, -2), (🍎, 1)
Z	(🍎, -3), (🍎🍎, 1)	(🍎🍎, -4), (🍎, 4)	(🍎🍎), (🍎, 3)

Figure 3.1: A visualization of our allocation algorithm mapping a payoff game to an allocation game

3.3.3 Experimental Details

We tested our approach on experimental data collected from participants on Amazon Mechanical Turk. Mechanical Turk allows researchers to present participants with human intelligence tasks (HITs) to complete. We presented participants with a set of 24 3×3 symmetric normal form games (the payoff games) in which each participant played against the actions of the previous participant. Turkers first had to complete a quiz testing their understanding of the parameters of a normal form game before being allowed to continue in the HIT. All participants had at least 95% HIT approval and had completed at least 100 HITs. The first 200 datapoints, mainly in the ordered filtered treatment, consisted of turkers who also had the masters qualification, but this requirement was removed as it slowed down the rate of data collection. Previous studies have shown that the quality of work of master workers and those with a high enough HIT approval rating are comparable [32, 33, 42].

We removed all data from participants who completed the HIT in fewer than 120 seconds, or 5 seconds per game, as there was a high correlation between participants who did this and responses at the end of the survey that were either left empty or spurious.[3]

The payoff games were generated by randomly sampling payoffs from a uniform distribution on $[0, 100]$. The games were played by participants in a randomized order. We collected additional treatments, analyzed in Section A.1.1 of the appendix, in which the order and type of games were varied.

Participants were paid \$1.50 for completing the HIT, as well as a performance bonus based on their total payoffs in the games. The performance bonus was calculated by multiplying the payoffs achieved by the participant by \$0.02, with the goal being to have all participants achieve an equivalent wage of at least \$10 an hour[4] between the bonus and base payment if they had uniformly randomized and taken the maximum allowed time of 30 minutes.

3.4 Estimation Methods

The strategic model components we consider are: Nash equilibrium, quantal response equilibrium (QRE), quantal cognitive hierarchy (QCH), and no strategic behavior. Of the strategic models, Nash and QRE are equilibrium models while QCH is not. The non-strategic models we consider are: uniform randomization, quantal linear4 (QL4), and no non-strategic behavior. These non-strategic models satisfy the formal definition of non-strategic behavior given in Wright and Leyton-Brown [50]. The parameters for each model are thus $\theta = (\theta_S, \theta_{NS})$, and the parameters for each component included in our

[3]For example, two participants had the exact same input in the feedback field, seemingly referring to a task in an entirely different HIT.

[4]All dollar amounts listed here are in USD

main evaluation are summarized in Table 3.1.

Table 3.1: A summary of the strategic and non-strategic components included in our evaluation and the parameters θ for each component.

Strategic Component	$\hat{\theta}_S$	Non-strategic Component	$\hat{\theta}_{NS}$
Nash	$\emptyset$	none	$\emptyset$
QRE	β, λ	uniform randomization	$\emptyset$
PQCH	τ, λ	quantal-linear4	$w_{L0}, \lambda_0,$
None	$\emptyset$		

To obtain our parameter estimate $\hat{\theta}$ of θ and $\hat{v}$ of v, we performed log-likelihood maximization with respect to $\hat{v}$ and $\hat{\theta}$ jointly using L-BFGS-B [6, 51]. The data passed to the analyst are a set of allocation games generated from $\mathcal{A} \in \mathcal{G}|v^*$ and the empirically observed behavior of subjects from our online experiment. To evaluate the performance of our value estimation, we take the estimated value $\hat{v}$ at the maximum likelihood estimate of each model and compare it to the endowed value.

3.4.1 Estimation of Equilibrium Models

To estimate equilibrium models (QRE, QRE+L0, and Nash), we chose the $(v, \lambda, \beta, \lambda_0, w_{L0})$ that maximized the likelihood of the empirical behavior of the participants under the assumption that each strategic agent was quantally best responding to the empirically observed distribution s^g_{-i} defined by

$$s^g_{-i}(a) = \frac{\left|\{j \neq i \mid g \in G(j) \wedge a^g_j = a\}\right|}{|\{j \neq i \mid g \in G(j)\}|}.$$

For equilibrium models we maximize the following likelihood:

$$\log\mathcal{L}(\lambda, v, \beta, \lambda_0, w_{L0}) = \sum_i \sum_{g \in G(i)} \log\left[\beta f^{L0}(G_g(v); \lambda_0, w_{L0})(a^g_i) + (1-\beta) QBR_i(s^g_{-i} \mid G_g(v), \lambda)(a^g_i)\right] \tag{3.1}$$

where v is the value parameter being estimated; λ is the behavioral precision parameter; β is the proportion of non-strategic agents between 0 and

1; (λ_0, w_{L0}) are the behavioral precision and weight parameters for quantal-linear4, respectively; $G_g(v)$ is the payoff game induced from an allocation game g and valuation v; $\hat{s}^g_{-i}$ is the empirical distribution of play in allocation game g; and a^g_i is the action taken by participant i in game g.[5][6]

The econometric approach of computing QRE by assuming all agents are quantally responding against other agents in the empirically observed distribution of actions is commonly used [e.g. 5, 10, 19, 38]. What is not common is the simultaneous estimation of both the precision of agents λ as well as the value parameter v. The previously listed works all do a two-step estimation method of either estimating $\lambda \mid v$ or vice versa. This is because given an observed action $s_i(a_i)$ generated from a logit model which takes as an input observed utility u_i of the form $u_i = \lambda v$, there are infinitely many combinations of λ and value that could result in the same observed utility. This motivates the inclusion of a payment profile p in our allocation games. Including a static payoff p allows us to simultaneously estimate both v and λ by anchoring λ to a specific scale; indeed we find that when constraining $p = 0$, our estimates are off by up to an order of magnitude (refer to Table A.9 in the appendix).

Nash equilibrium does not have model parameters to estimate. When estimating values using the Nash equilibrium model, we approximate best response using quantal best response with a high value of λ,[7] and select the value that maximizes (3.1). This approach allows us to select a single value that is *most* consistent with best response, rather than a set of values that are consistent with all agents' best-responding. More critically, it also ensure that every possible action has positive probability. When assuming best response with no error model, a single action by a single agent that is not consistent with best response can lead to the entire dataset's having probability 0. Under our approach, actions inconsistent with best response will instead be assigned a very low, but positive, probability.

[5]For models with uniform randomization as the non-strategic component, we do not estimate (λ_0, w_{L0}).

[6]We fix $\beta = 0$ when estimating models without a non-strategic component

[7]We used $\lambda = 100$ in our experiments, as we found that both the predictive performance and value estimate converge at precision $\lambda \geq 100$; refer to Figure A.9 in the appendix for details.

3.4.2 Estimation of Poisson Quantal Cognitive Hierarchy

For Poisson quantal cognitive hierarchy, we estimate $(v, \lambda, \tau, \lambda_0, w_{L0})$ by maximizing the following likelihood:

$$\log\mathcal{L}(\lambda, v, \tau, \lambda_0, w_{L0}) =$$
$$\sum_i \sum_{g \in G(i)} \log \left[L_{\tau;0} f^{L0}(G_g(v); \lambda_0, w_{L0})(a_i^g) + \sum_{\ell=1}^{3} L_{\tau;\ell} QBR_i(G_g(v), \lambda \mid \ell_{i|0:\ell-1})(a_i^g) \right] \tag{3.2}$$

where τ is a mean parameter on a truncated Poisson distribution where the max possible level of an agent is 3. In contrast to the equilibrium models, the likelihood for PQCH does not treat the empirically observed distribution as the distribution of actions being responded to; we instead find the mean parameter τ that generates a distribution which maximizes the likelihood against the empirical data. Assuming that strategic QCH agents respond to the empirical distribution of lower-level agents would require us to estimate the levels (or posterior level distributions) for each agent, in order to estimate which agents' empirical behavior is being responded to; e.g., to determine what empirical distribution is being responded to by level-2 agents, we must first determine which agents are level-0 and level-1. This is a much more complex estimation problem, both statistically and computationally. For this reason, we take the simpler approach of estimating the mean parameter τ instead.[8]

3.4.3 Utilization of Panel Structure in Estimation of Values

In our experimental setup, we collected panel data where the individual actions of each player for each game are recorded, in contrast to other common data sets in which the actions of all agents are pooled together. This panel structure allows the estimation of model parameters that are heterogeneous across agents

[8] We also estimated our equilibrium models by finding the optimal parameter λ that maximizes the likelihood against our data; we did not find a significant difference in the estimated v, which provides assurances that estimating against the empirical distribution provides a reasonable approximation while being much simpler to compute. Refer to Table A.10 in the appendix.

but stable for a given agent i. The level of an individual agent in QCH-based models is an example of a parameter that could match this description.[9] [10] If a player's level is the same in every game, then using a likelihood that explicitly encodes this has the potential to provide more accurate estimates than one that assumes that each player's level is re-sampled before every action. (3.3) gives the likelihood for a model with parameters θ, a stable level ℓ_i for agent i distributed according to $\Pr(\ell_i = \ell \mid \theta)$, in which agent i takes action a in a game g with probability $\Pr(a_i^g \mid \ell_i = \ell, \theta)$.

$$\log \Pr(D \mid \theta, v) = \sum_i \sum_{g \in G(i)} \log \left[\sum_\ell \Pr(\ell_i = \ell \mid \theta) \Pr\left(a_i^{G_g(v)} \mid \ell_i = \ell, \theta\right) \right] \tag{3.3}$$

In contrast, the likelihood for an otherwise-identical model in which each agent's level can vary between games is given by (3.4).

$$\log \Pr(D \mid \theta, v) = \sum_i \log \left[\sum_\ell \Pr(\ell_i = \ell \mid \theta) \prod_{g \in G(i)} \Pr\left(a_i^{G_g(v)} \,\middle|\, \ell_i = \ell, \theta\right) \right] \tag{3.4}$$

When running our analysis on synthetic data we find that the likelihood of (3.4) is more numerically stable than that of (3.3), while returning a similar value estimate. We therefore report the parameters estimated using (3.4) in this paper. Our dataset is available for future research questions or models that require panel data.

[9] This same discussion applies to QRE+L0, if we treat strategic agents as having a level $\ell_i = 1$ and non-strategic agents as having a level $\ell_i = 0$.

[10] Our definition implies that every non-strategic agent plays a mixture over a number of level-0 decision rules. However, one could also imagine a definition in which there is a population of non-strategic agents, each using a single level-0 rule. Under this assumption, the assignment of decision rules to agents could also fit this description of a heterogeneous but stable behavioral parameter.

Chapter 4

Results

Our evaluation finds that models with a rich non-strategic component perform better in value estimation from behavior in initial play than those without a non-strategic component (e.g., Nash, QRE) or those that model non-strategic agents as uniform randomizers. Additionally, we find that models that include a strategic component perform better at value estimation than those that assume that all agents are non-strategic, but the choice of strategic model is not as important as the rich non-strategic component in a given model.

4.1 Evaluation

Our evaluation considered traditional equilibrium models with no non-strategic agents (Nash, QRE), and QCH and QRE+L0 where level-0 agents were either uniform randomizers or quantal-linear4 agents. We evaluated each model across multiple scenarios given v^* in $\mathcal{V} = [5, 10, 20, 40, 80]$. For each v^*, we generated $k = 25$ scenarios where we mapped our payoff games to a set of allocation games $\mathcal{A}_{\mathcal{G}}$ given v^*.

We measured each model's value estimation for each scenario using relative error, $\frac{|\hat{v}-v^*|}{v^*}$. We chose to normalize the error to account for the differing scale of values in $\mathcal{V}$. The value estimate for each scenario was evaluated using using the mean value estimate of 10 rounds of 10-fold cross-validation, with the test set being used to evaluate behavioral prediction The mean value estimate for each scenario are distributed according to a Student's t-distribution [e.g. 46]. We say that one model performs better in value estimation than another when

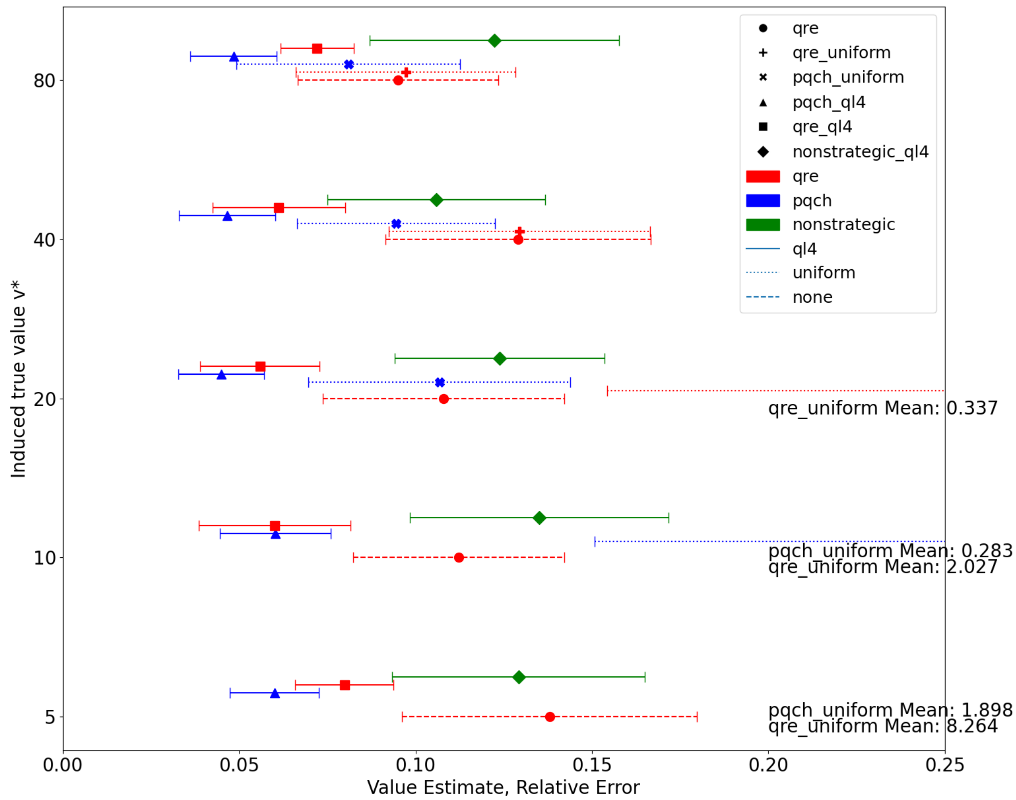

Figure 4.1: Summary plot showing values of v^* vs. the relative error. The two left-most points for all values of v^* are models containing both a strategic component and quantal-linear4 as the non-strategic component; the green point indicates a model that assumes all agents are non-strategic in a quantal-linear4 manner with no strategic behavior.

the 95% confidence intervals do not overlap.

Figure 4.1 and Table 4.1 show the performance in value estimation across models, with Figure 4.1 being a visualization of the data in Table 4.1. Behavioral models with quantal-linear4 as the non-strategic component outperform classical equilibrium models in terms of value estimation across every endowed value v^* that we evaluated. We find that using quantal-linear4 as the non-strategic component outperforms corresponding strategic models which use a uniform non-strategic component, regardless of the choice of strategic model. This leads us to conclude that modelling non-strategic behavior is more important than the choice of strategic model. Note, however, that None-QL4 does not perform as well as PQCH-QL4 or QRE-QL4, which suggests that a strategic component in the model is still necessary. Another observation is that models containing QL4 remain stable across values of v^*; the mean relative error for QL4 models varies at most by 2%, in contrast to classic equilibrium models or uniform non-strategic augmented models in which the relative errors differ by an order of magnitude from each other depending on v^*. This leads us to conclude that QL4 leads to a more *reliable* estimate of values.

We further demonstrate the importance of obtaining accurate value estimates in Table 4.2. We first obtain an estimate of θ and v on half of the games in our dataset ($m = 12$). Using the estimated value $\hat{v}$ and behavioral parameters $\hat{\theta}$, we then predict the average subject welfare on the remaining half of games that were held out. We then compute the relative error of the predicted welfare against the empirically observed average welfare of subjects. This evaluation requires a model to be accurate in both its estimation of behavioral parameters as well as that of values; a model with an accurate value estimate but a poor prediction of behavior would perform poorly, and vice versa. PQCH-QL4 and QRE-QL4 once again perform the best at this task, with Nash being noticeably poor at welfare prediction, especially at lower values of v. This pattern persists across models; welfare estimates are worse for lower values of v^* compared to higher values, albeit at a much larger scale for Nash and for models with a uniform non-strategic component. The final note here is that QRE-None outperforms QRE-uniform across the board, which shows that a level-0 model is not sufficient to improve performance: a rich level-0 model is necessary.

4.2 Contribution of Strategic vs. Non-strategic Components of the Model

We attempt to quantify the contribution of quantal-linear4 to the observed improvement in value estimation. We compare the cross-product of our strategic and non-strategic components as discussed in Chapter 3 and find that QL4 outperforms any of the other non-strategic models considered. In addition to comparing quantal-linear4 and uniformly randomization, in this section we

Table 4.1: Relative error by v^*, with confidence interval in parentheses. Bold cells indicate best performing model for each v^*. Italicized cells indicate models which are not significantly different from the best performing one. QRE-QL4 indicates a model in which a fraction of agents are behaving non-strategically in a QL4 manner while the remaining agents are in QRE with themselves and non-strategic agents. None-QL4 indicates a model in which all agents are behaving non-strategically.

Component		v^*				
Strategic	Non-strategic	5	10	20	40	80
Nash	none	10.41, (8.01, 12.8)	2.88, (2.06, 3.71)	0.64, (0.44, 0.83)	0.29, (0.18, 0.4)	0.2, (0.15, 0.25)
QRE	none	0.14, (0.1, 0.18)	0.11, (0.08, 0.14)	0.11, (0.07, 0.14)	0.13, (0.09, 0.17)	0.1, (0.07, 0.12)
QRE	uniform	8.27, (5.14, 11.4)	2.04, (1.21, 2.87)	0.37, (0.18, 0.56)	0.13, (0.09, 0.17)	0.1, (0.07, 0.13)
PQCH	uniform	1.93, (0.68, 3.19)	0.32, (0.16, 0.48)	0.12, (0.08, 0.16)	0.09, (0.07, 0.12)	*0.08, (0.05, 0.11)*
PQCH	QL4	**0.06, (0.05, 0.07)**	**0.06, (0.04, 0.08)**	**0.05, (0.03, 0.06)**	**0.05, (0.03, 0.06)**	**0.05, (0.04, 0.06)**
QRE	QL4	*0.08, (0.07, 0.09)*	*0.06, (0.04, 0.08)*	*0.06, (0.04, 0.07)*	*0.06, (0.04, 0.08)*	0.07, (0.06, 0.08)
none	QL4	0.13, (0.09, 0.17)	0.14, (0.1, 0.17)	0.12, (0.09, 0.15)	0.11, (0.08, 0.14)	0.12, (0.09, 0.16)

Table 4.2: Relative error of predicted average per game welfare by v^*. In each scenario, the estimated valuation and model parameters θ from half the games are used to predict the average game welfare per subject on the other half and is compared against the empirically observed average welfare. Bold cells indicate models with the lowest MSE.

Component		v^*				
Strategic	Non-strategic	5	10	20	40	80
Nash	none	16.81	5.42	1.75	0.41	0.30
QRE	none	0.24	0.20	0.20	0.18	0.12
QRE	uniform	12.96	4.26	0.55	0.16	0.15
PQCH-uniform	uniform	5.86	0.52	0.22	0.13	0.11
PQCH-QL4	QL4	**0.12**	0.12	**0.09**	**0.09**	**0.07**
QRE-QL4	QL4	**0.12**	**0.11**	**0.09**	0.10	0.08
none	QL4	0.56	0.27	0.39	0.18	0.18

include linear4 from [47] as well as a differentiable version of linear4 we refer to as differentiable-linear4 (DL4) where $\lambda_0 = 1$, which gives us a differentiable function with respect to v without adding an additional degree of freedom.

For each model resulting from the the cross-product of strategic and non-strategic components, we take each of our scenarios for each value v^* (n = 125) and report the percentage of the time that the relative error of $\hat{v}$ falls below a threshold α (i.e., the error falls within 10% accuracy). We sampled 1000 bootstrapped samples from our empirically observed data $\mathcal{D}$ and did this for each bootstrapped sample, reporting the median percentage each model falls within our threshold with the lower and upper bounds being the middle 95% of the bootstrapped estimates as outlined in [12]. Doing this allows us to to see how well a given non-strategic component performs at recovering v^*, regardless of the strategic component being used in the model. The results demonstrate the advantages of quantal-linear4, as it performs strictly better than uniform and linear4, and outperforms differentiable-linear4, although not significantly. The results of this test are reported in Table 4.3.

There are two reasons why linear4 performs poorly as a non-strategic model: the first is that as a non-continuous function of v, it is not differentiable with respect to v and so our optimization procedure fails to reliably find the value that maximizes likelihood; checks on synthetic data show that the likelihood returned by the estimator is often worse than the likelihood at the known ground truth value. A second possible reason is due to the lack of quantal response in non-strategic agents; if we believe that strategic agents quantally respond to their payoffs, it stands to reason that non-strategic agents do so as well. This would also be a possible explanation for why quantal-linear4 outperforms differentiable-linear4.

4.3 Identification of Quantal-linear4

We do not claim that quantal-linear4 is a complete specification of non-strategic behavior; rather, we claim only that it captures regularities of non-strategic behavior beyond that of uniform randomization, resulting in an improvement

Table 4.3: Percentage of the time that relative error is less than 10% across all values and scenarios in $\mathcal{V}$. Each cell corresponds to a model STRAT-NONSTRAT with the row indicating the strategic model and column indicating the non-strategic one. QL4 (rightmost column) outperforms all other non-strategic components regardless of the strategic model. Here, the confidence intervals are derived from a k bootstrapped samples of the observed data, with $k = 1000$. Cells marked "n/a" do not have a conceivable model that elicits an estimate of values. Cells containing 0 mean that none of the bootstrapped samples had a value estimate that fell within 10% of v^*.

Strategic Component	Non-strategic Component				
	None	Uniform	L4	DL4	QL4
Nash	0.12 (0.06 0.18)	0.10 (0.05 0.15)	0	0.05 (0.02 0.09)	0.69 (0.50 0.84)
QRE	0.50 (0.41 0.57)	0.33 (0.26 0.41)	0	0.71 (0.62 0.82)	0.70 (0.60 0.82)
PQCH	n/a	0.52 (0.42 0.64)	0	0.62 (0.50 0.74)	0.84 (0.72 0.94)
None	n/a	n/a	0	0	0.45 (0.30 0.58)

of behavioral prediction and value estimation. Constructing a level-0 specification that fully captures all the vagaries of non-strategic behavior is beyond the scope of this paper. The inclusion of a richer level-0 model leads to the concern, however, that the model may no longer be identified; if this is the case then the estimated $\hat{\theta}$ may not be unique. We argue that the tradeoff in increasing the performance in value estimation is worth the cost of introducing a possibly inconsistent model. The results in Table 4.2 indicate that when paired with a quantally responding strategic component, the behavioral estimates in quantal-linear4 provide a sufficiently close estimate of predicted behavior such that it outperforms uniform randomization in predicting welfare.

Furthermore, we find that the estimated behavioral parameters (the mean parameter τ for PQCH and proportion β of non-strategic agents for equilibrium models) on the empirical dataset remain stable; see Section A.5.1 in the appendix. Across most different values of v^*, as well as when estimating the behavioral parameters only, the confidence interval of each behavioral parameter overlap (i.e., there does not seem to be a bimodal distribution of the parameter estimate). We also find that when we estimate only the behavioral parameters using the payoff games directly, our estimates of the behavioral parameters do not differ significantly from when we jointly estimate valuation and behavioral parameters using the allocation games.

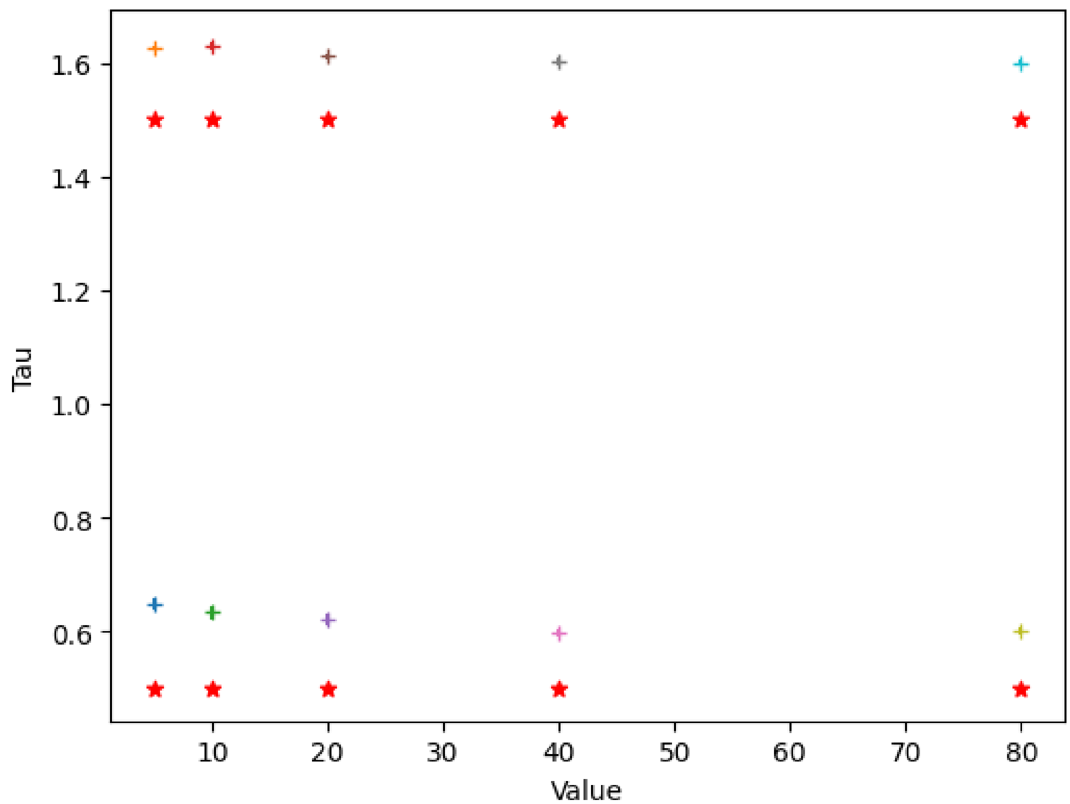

Figure 4.2: Synthetic data: estimated behavioral parameters vs. preferences v^* in PQCH. Asterixes indicate ground truth parameters, with each intersection representing the parameters of the synthetically generated dataset. We chose $\lambda = 0.05$ to approximate the precision in an empirical dataset, with the non-strategic agents using a maxmax strategy, where non-strategic agents pick the action that maximizes their best case payoff. τ indicates the estimated Poisson mean parameter for the distribution over levels, which seems to be biased upwards (likely due to it being truncated) but does not affect the estimate for v.

Finally, we generated synthetic datasets using PQCH-QL4 and find that when the model is correctly specified, our value estimates are are extremely close to the endowed value with a high degree of confidence. While the behavioral mean parameter seems to have a slight upward bias, it still falls within the neighborhood of the ground truth mean parameter. These results are summarized in Figure 4.2.

4.4 Alternate Treatments

We summarize the results of our experiments on alternate datasets below.

4.4.1 Evaluation of ALL10 Dataset

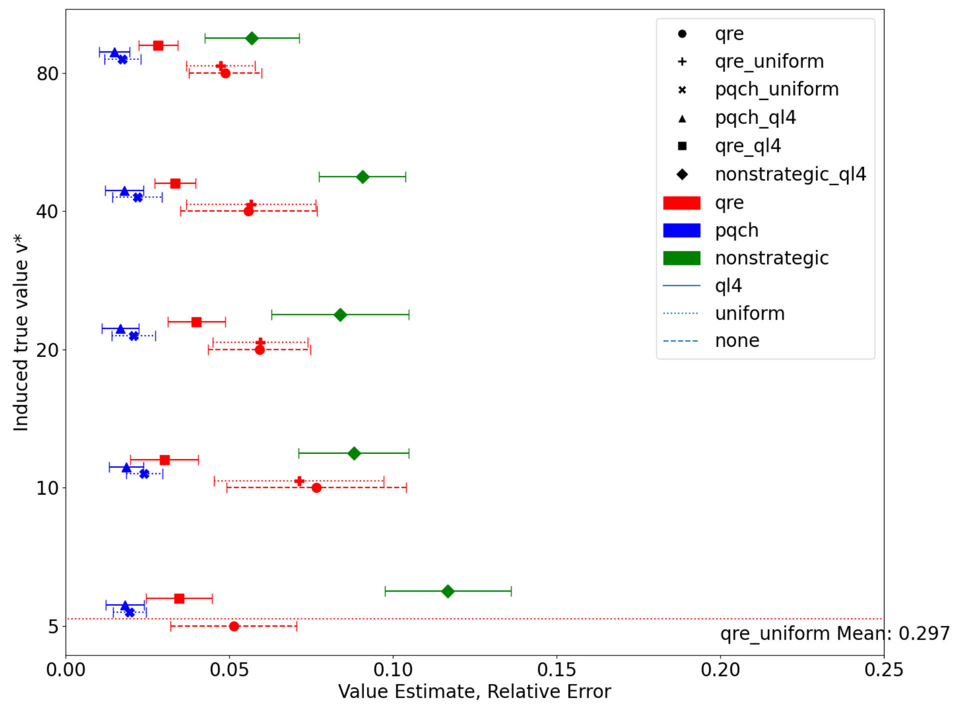

Figure 4.3: Summary plot showing values of v^* vs. the relative error for the ALL10 dataset. As the number of samples for ALL10 is significantly smaller, the confidence intervals are correspondingly wider, especially for models with UNIFORM as the non-strategic model

We evaluated our experimental setup on the set of 86 3x3 symmetric games in the ALL10 dataset from [17] and find that the relative error is similiar to that of the data collected in our experiment[1]. Notably, the ALL10 dataset contains many fewer subjects per game ($40 \leq n \leq 147$) and so the confidence interval of the value estimate ends up being much wider when there is no rich non-strategic model[2].

[1]An additional 2 treatments are available in [17], but the games are not symmetric and information about only the row player actions is included. This means our method of computing QRE cannot be done on the data and so we did not analyze them.

[2]The All10 dataset also only contains summary of player actions instead of panel data, which means that we are only able to take the pooled approach in equation 3.4

4.4.2 Differences in Alternate Treatments

As mentioned previously, we collected data with additional treatments. While there are quantitative differences in the treatments, our main finding that models containing quantal-linear4 outperform their uniform counterparts still holds. The results for the additional treatments can be found in section A.1.1 of the appendix. The most notable difference comes from the randomized filtered treatment, in which none of the models are statistically different from one another; additionally, the randomized filtered treatment also has the worst mean relative error in value estimates of all the treatments. There are a few possible reasons for this. The first is that the filtering of games captured some regularity that causes some issue in estimating values. The second is that there was a time difference between the collection of the treatments, with the collection of the data from the randomized filtered condition occurring several months later due to logistical reasons. To test this hypothesis, we also grouped the treatments together by the filtering condition, creating a combined filtered and combined non-filtered treatment respectively, and find that our key findings remain consistent across treatments, providing reassurances that our approach should work across different types of games. We examine a possible reason for the difference in relative error estimates in the following chapter.

Chapter 5

Theoretical Results

In this section, we present a preliminary theoretical analysis on the error in recovering the value from quantal response behavior. We provide an analysis on what makes an individual allocation game well or poorly suited for value estimation. This drives an empirical exploration of our collected data and provides a possible explanation for the difference in the errors of our value estimates between the filtered and non-filtered conditions.

5.1 Bounding the Error in our Value Estimate

Given our estimation approach, we want to bound the error we would get for a given number of samples. We find that having a large difference in allocations is key to bounding the error in our estimate of the value.

Our analysis is based upon a previous work. We start from the following lemma taken directly from Haghtalab et al. [21]:

Lemma 1. Let $\hat{D}$ be the empirically observed distribution of actions based on $m = \Omega(\frac{1}{\rho\epsilon^2} log(\frac{n}{\delta}))$ samples, where $\rho = poly(n)$, and n is the number of actions available within a game.
With probability $1 - \delta$, for all actions $i \in A$, $\frac{1}{1+\epsilon} \leq \hat{D}_i / D_i \leq 1 + \epsilon$.

Assuming a quantal response model, we know that each probability is generated according to:

$$D_i = \frac{exp(u_i(v))}{\sum_{i \in A} exp(u_i(v))} \tag{5.1}$$

where $u_i(v)$ is the utility of action i given valuation v. The ratio between any two actions i and j therefore follows the relation:

$$u_i(v) = ln\Big[\frac{D_i}{D_j}\Big] + u_j(v) \tag{5.2}$$

By construction, we know that the utility of an action i takes the form:

$$u_i(v) = \lambda(x_i \cdot v + p_i) \tag{5.3}$$

Here, we make a simplifying assumption and only consider the utility of an agent for a given action, holding the actions of the other agent(s) constant. That is, $u_i(v) = f(\mathcal{A}(v), s_{-i})$, but we suppress s_{-i} by converting $f(\mathcal{A}(v), s_{-i})$ to x_i and p_i, converting this to a single agent problem.

Substituting $u_i(v)$ and rearranging in terms of v, we obtain:

$$v = \frac{\frac{1}{\lambda} ln\left[\frac{D_i}{D_j}\right] + (p_j - p_i)}{x_i - x_j} \tag{5.4}$$

We note that as the ratio $\frac{D_i}{D_j}$ changes based on the empirically observed distribution of actions, the estimate of v becomes more inaccurate as $\frac{\hat{D}_i}{\hat{D}_j}$ gets farther from $\frac{D_i}{D_j}$; that is, inaccuracy in our observation of D leads to error in the value estimate. The error in our value estimate is then:

$$|\hat{v} - v^*| = \frac{\frac{1}{\lambda} ln\left[\frac{\hat{D}_i}{\hat{D}_j}\right] + (p_j - p_i)}{x_i - x_j} - \frac{\frac{1}{\lambda} ln\left[\frac{D_i}{D_j}\right] + (p_j - p_i)}{x_i - x_j} \tag{5.5}$$

$$= ln\left[\frac{D_j \hat{D}_i}{D_i \hat{D}_j}\right] \frac{1}{\lambda(x_i - x_j)} \tag{5.6}$$

$$= ln\left[\frac{D_j \hat{D}_i}{\hat{D}_j D_i}\right] \frac{1}{\lambda(x_i - x_j)} \tag{5.7}$$

from Lemma 1, we know that $\frac{\hat{D}_i}{D_i} \leq 1 + \epsilon$:

$$\leq \frac{ln(1+\epsilon)^2}{\lambda(x_i - x_j)} \leq \frac{2\epsilon}{\lambda(x_i - x_j)} \tag{5.8}$$

where the last inequality comes from the inequality $ln(1 + x) \leq x$ for all $x \in \mathbb{R}$

We now have an upper bound on the maximum error of the value estimate given the analytical estimate; the question is whether this bound is also an upper bound on the MLE estimate. We conjecture that it must be as the analytical estimate only takes into consideration two of the observed distributions while MLE takes into account all of them. Without loss of generality, let $D_i > D_{i+1}$. We run an empirical check using $\hat{D}_1$ and $\hat{D}_2$ (that is, the 2 most actions with the highest observed probability), and find that the MLE estimate is within the bound in equation 5.8 a large percentage of the time. A final note is that as the ratio between $\hat{D}$ and D is reliant on sample size from Lemma 1, a given allocation game which does poorly in estimating an agent value could still produce an error within ϵ with a larger number of samples.

Proposition 2. For a given distribution $\hat{D}$ with $\epsilon = \arg\max_{i \in A} \frac{\hat{D}_i}{D_i}$, a mapping of a payoff game $\mathcal{G}$ to an allocation game $\mathcal{A}$ will have a lower relative error when $x_i - x_j$ is larger.

5.2 Empirical Results in Treatments

Table 5.1: Percentage of the time the absolute relative error is lower for the scenario with the highest minimum allocation difference vs. the lowest minimum allocation difference.

v^*	Randomized non-filtered	Randomized filtered
5	0.83	0.83
10	0.83	0.66
20	0.88	0.75
40	0.75	0.63
80	0.625	0.75

We next attempt to empirically verify Proposition 2 on our collected data. We find that the minimum apple difference in the payoff games (and therefore, for the corresponding allocation games) affects the error in the value estimate. Table 5.1 shows that across all $k = 25$ scenarios, the scenario with highest minimum allocation difference has a lower error than the scenario with the lowest minimum allocation difference a large percentage of the time. Here, we obtain the utilities for each action by weighting the allocations and payments of each column action by the empirically observed distribution.

We conclude this analysis with the remark that this only investigates the relation in error between an observed distribution produced by quantal response and the resulting error in a linear utility function. We know the model must be inconsistent, and do not attempt to analyse the error induced between the different models we compare. We also do not attempt to quantify how jointly estimating a shared v^* across multiple games affects the error estimate. The results from Table 5.1 as well as Proposition 2 offers some insights into the problem. First, for a given observed distribution, it seems that for a given ϵ, increasing the difference in allocations between actions improves the chance of getting a better value estimate. Second, having a low difference in actions (or a high ϵ) could be overcome by having a sufficient number of samples. Finally, a decrease in λ leads to a corresponding increase in the upper bound of the error. One thing that remains unclear is the effect that payments and the endowed value v^* have on the error in the value estimate, which is worth further investigation. In regards to the randomized filtered treatment, this analysis provides a possible insight. Supposing that the filtered treatment produced allocation games which had some regularity that required a larger number of samples, whereas the ordered filtered treatment had a larger number of samples ($n = 303$) that could overcome this regularity, the randomized filtered had a number of samples closer to the other treatments. The fact that the combined filtered treatment which combined both filtered treatments offered an estimate closer to that of the ordered filtered one rather than the randomized filtered one lends further credence to this theory.

Chapter 6

Conclusion

6.1 Conclusions

This thesis examines the benefit of using behavioral models for value estimation. Behavioral models typically include parameters that must be estimated from the data. Using a novel experimental design, we demonstrate that estimating these behavioral parameters simultaneously with value parameters is feasible, and leads to more reliably accurate value estimates from initial play than models based on the standard strong equilibrium assumption.

We introduce a new specification of level-0 behavior called *quantal-linear4*, and a new behavioral model called *QRE+L0* that extends quantal response equilibrium to settings that contain non-strategic agents, who are responsive to their own preferences but do not reason about other agents. Our results show that models that include a rich level-0 specification perform better at estimating values from initial play. These results strongly argue for the importance of explicitly modeling non-strategic behavior rather than treating it as noise, especially in contexts such as initial play in which equilibrium is unlikely to have been reached.

6.2 Future Work

There are a number of directions in which the work in this thesis could be extended. We made a simplifying assumption that all agents shared a homogeneous value, but an important future direction would be to estimate individual agent values. Further to this direction, we could extend this work to estimating individual behavioral parameters of agents. This would allow us to better optimize individual welfare, in addition to modelling differences in behavior based on hetereogeneous beliefs about the values of others. As previously discussed in Section 3.4, we could also extend the approach of responding to empirically observed distributions towards models of iterated reasoning, which would allow us to move beyond needing to specify a distribution over levels.

Using our framework of separating models into a strategic and non-strategic component, expressed by QRE+L0, we could examine other models of non-strategic behavior (for example, the model proposed by Fudenberg and Liang [17] could be used as a non-strategic model) beyond those discussed in this paper. This could take the form of extending QL4 to be more predictive, or evaluating domain-specific models of nonstrategic behavior.

Finally, a surprising result was the difference in value estimates between treatments. Given these games had distinct higher level strategies, we intuitively believed that it would be easier to estimate behavior and, by extension infer preferences from these games. Further expanding upon the analysis in Chapter 5, we could investigate how jointly estimating one value across several allocation games affects the error, or how different models of non-strategic behavior affect the error bound.

References

[1] Nikhil Agarwal and Paulo Somaini. “Demand analysis using strategic reports: An application to a school choice mechanism.” In: *Econometrica* 86.2 (2018), pp. 391–444.

[2] Maurice Allais. “Le comportement de l’homme rationnel devant le risque: critique des postulats et axiomes de l’école américaine.” In: *Econometrica: Journal of the Econometric Society* (1953), pp. 503–546.

[3] Susan Athey and Philip A Haile. “Nonparametric approaches to auctions.” In: *Handbook of econometrics* 6 (2007), pp. 3847–3965.

[4] Susan Athey and Denis Nekipelov. “A structural model of sponsored search advertising auctions.” In: *Sixth ad auctions workshop*. Vol. 15. 2010.

[5] Patrick Bajari and Ali Hortacsu. “Are structural estimates of auction models reasonable? Evidence from experimental data.” In: *Journal of Political economy* 113.4 (2005), pp. 703–741.

[6] Richard H Byrd et al. “A limited memory algorithm for bound constrained optimization.” In: *SIAM Journal on scientific computing* 16.5 (1995), pp. 1190–1208.

[7] Caterina Calsamiglia, Chao Fu, and Maia Güell. “Structural estimation of a model of school choices: The boston mechanism versus its alternatives.” In: *Journal of Political Economy* 128.2 (2020), pp. 642–680.

[8] Colin F Camerer, Teck-Hua Ho, and Juin-Kuan Chong. “A cognitive hierarchy model of games.” In: *The Quarterly Journal of Economics* 119.3 (2004), pp. 861–898.

[9] Shuchi Chawla, Jason Hartline, and Denis Nekipelov. “A/B testing of auctions.” In: *Proceedings of the 2016 ACM Conference on Economics and Computation*. 2016, pp. 19–20.

[10] Yefen Chen, Xuanming Su, and Xiaobo Zhao. “Modeling bounded rationality in capacity allocation games with the quantal response equilibrium.” In: *Management Science* 58.10 (2012), pp. 1952–1962.

[11] Juin-Kuan Chong, Teck-Hua Ho, and Colin Camerer. “A generalized cognitive hierarchy model of games.” In: *Games and Economic Behavior* 99 (2016), pp. 257–274.

[12] Paul R Cohen. *Empirical methods for artificial intelligence*. Vol. 139. MIT press Cambridge, 1995.

[13] Miguel Costa-Gomes, Vincent P Crawford, and Bruno Broseta. "Cognition and behavior in normal-form games: An experimental study." In: *Econometrica* 69.5 (2001), pp. 1193–1235.

[14] Vincent P Crawford, Miguel A Costa-Gomes, and Nagore Iriberri. "Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications." In: *Journal of Economic Literature* 51.1 (2013), pp. 5–62.

[15] Vincent P Crawford and Nagore Iriberri. "Level-k auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions?" In: *Econometrica* 75.6 (2007), pp. 1721–1770.

[16] Drew Fudenberg, Wayne Gao, and Annie Liang. "How Flexible is that Functional Form? Quantifying the Restrictiveness of Theories." In: *arXiv preprint arXiv:2007.09213* (2020).

[17] Drew Fudenberg and Annie Liang. "Predicting and understanding initial play." In: *American Economic Review* 109.12 (2019), pp. 4112–41.

[18] Jacob K Goeree and Charles A Holt. "Ten little treasures of game theory and ten intuitive contradictions." In: *American Economic Review* 91.5 (2001), pp. 1402–1422.

[19] Jacob K Goeree, Charles A Holt, and Thomas R Palfrey. "Quantal response equilibrium and overbidding in private-value auctions." In: *Journal of Economic Theory* 104.1 (2002), pp. 247–272.

[20] Emmanuel Guerre, Isabelle Perrigne, and Quang Vuong. "Optimal nonparametric estimation of first-price auctions." In: *Econometrica* 68.3 (2000), pp. 525–574.

[21] Nika Haghtalab et al. "Three strategies to success: Learning adversary models in security games." In: (2016).

[22] Philip A Haile, Ali Hortaçsu, and Grigory Kosenok. "On the empirical content of quantal response equilibrium." In: *American Economic Review* 98.1 (2008), pp. 180–200.

[23] David W Harless and Colin F Camerer. "The predictive utility of generalized expected utility theories." In: *Econometrica: Journal of the Econometric Society* (1994), pp. 1251–1289.

[24] Jason S Hartford, James R Wright, and Kevin Leyton-Brown. "Deep learning for predicting human strategic behavior." In: *Advances in neural information processing systems* 29 (2016).

[25] Yinghua He. "Gaming the Boston school choice mechanism in Beijing." In: (2015).

[26] Il Myoung Hwang. “A robust redesign of high school match.” PhD thesis. The University of Chicago, 2015.

[27] Daniel Kahneman and Amos Tversky. “Prospect Theory: An Analysis of Decision under Risk.” In: *Econometrica* 47.2 (1979), pp. 263–292.

[28] Roger Koppl. “Retrospectives: animal spirits.” In: *Journal of Economic Perspectives* 5.3 (1991), pp. 203–210.

[29] Yun Shin Lee, Yong Won Seo, and Enno Siemsen. “Running behavioral operations experiments using Amazon’s Mechanical Turk.” In: *Production and Operations Management* 27.5 (2018), pp. 973–989.

[30] Joel Z Leibo et al. “Multi-agent reinforcement learning in sequential social dilemmas.” In: *arXiv preprint arXiv:1702.03037* (2017).

[31] Chun Kai Ling, Fei Fang, and J Zico Kolter. “What game are we playing? end-to-end learning in normal and extensive form games.” In: *arXiv preprint arXiv:1805.02777* (2018).

[32] Eric Loepp and Jarrod T Kelly. “Distinction without a difference? An assessment of MTurk Worker types.” In: *Research & politics* 7.1 (2020), p. 2053168019901185.

[33] Winter Mason and Siddharth Suri. “Conducting behavioral research on Amazon’s Mechanical Turk.” In: *Behavior research methods* 44.1 (2012), pp. 1–23.

[34] Daniel McFadden. “Econometric models of probabilistic choice.” In: *Structural analysis of discrete data with econometric applications* 198272 (1981).

[35] Richard D McKelvey and Thomas R Palfrey. “Quantal response equilibria for normal form games.” In: *Games and economic behavior* 10.1 (1995), pp. 6–38.

[36] John F Nash Jr. “Equilibrium points in n-person games.” In: *Proceedings of the national academy of sciences* 36.1 (1950), pp. 48–49.

[37] Denis Nekipelov, Vasilis Syrgkanis, and Eva Tardos. “Econometrics for learning agents.” In: *Proceedings of the sixteenth acm conference on economics and computation*. 2015, pp. 1–18.

[38] Gali Noti. “From Behavioral Theories to Econometrics: Inferring Preferences of Human Agents from Data on Repeated Interactions.” In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. 6. 2021, pp. 5637–5646.

[39] Harry J Paarsch, Han Hong, et al. “An introduction to the structural econometrics of auction data.” In: *MIT Press Books* 1 (2006).

[40] Daniel R Saunders, Peter J Bex, and Russell L Woods. “Crowdsourcing a normative natural language dataset: A comparison of Amazon Mechanical Turk and in-lab data collection.” In: *Journal of medical Internet research* 15.5 (2013), e2620.

[41] Daniel Sgroi and Daniel John Zizzo. "Learning to play $3\times$ 3 games: Neural networks as bounded-rational players." In: *Journal of Economic Behavior & Organization* 69.1 (2009), pp. 27–38.

[42] Matthew Staffelbach et al. "Lessons learned from an experiment in crowdsourcing complex citizen engineering tasks with Amazon Mechanical Turk." In: *arXiv preprint arXiv:1406.7588* (2014).

[43] Dale O Stahl and Paul W Wilson. "Experimental evidence on players' models of other players." In: *Journal of economic behavior & organization* 25.3 (1994), pp. 309–327.

[44] Kenneth Train. "E., 2003. Discrete choice methods with simulation." In: *Cambridge University Press* ().

[45] Hal R Varian. "Position auctions." In: *international Journal of industrial Organization* 25.6 (2007), pp. 1163–1178.

[46] Ian H Witten and Eibe Frank. "Data mining: practical machine learning tools and techniques with Java implementations." In: *Acm Sigmod Record* 31.1 (2002), pp. 76–77.

[47] James R Wright and Kevin Leyton-Brown. "Predicting human behavior in unrepeated, simultaneous-move games." In: *Games and Economic Behavior* 106 (2017), pp. 16–37.

[48] James R Wright and Kevin Leyton-Brown. "Level-0 models for predicting human behavior in games." In: *Journal of Artificial Intelligence Research* 64 (2019), pp. 357–383.

[49] James R Wright and Kevin Leyton-Brown. "A formal separation between strategic and nonstrategic behavior." In: *Proceedings of the 21st ACM Conference on Economics and Computation*. 2020, pp. 535–536.

[50] James R Wright and Kevin Leyton-Brown. "A Formal Separation Between Strategic and Nonstrategic Behavior." In: *arXiv preprint arXiv:1812.11571* (2022).

[51] Ciyou Zhu et al. "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization." In: *ACM Transactions on mathematical software (TOMS)* 23.4 (1997), pp. 550–560.

Appendix A

Appendix

A.1 Additional Experimental Details

A.1.1 Results for Different Treatments

When conducting the data collection process on Mturk, we varied 2 conditions for a total of 4 treatments.

1. We chose the payoff games according to two procedures. In the first condition, we used randomly-generated payoff games with no further filtering. In the second condition, we only used randomly generated payoff games for which no level-k strategy was similar to the level-$(k-1)$ strategy, when any of the `linear4` decision rules were used as a level-0 strategy.[1] We refer to the games from the first condition as the unfiltered games, and the games from the second condition as the filtered games.

2. In the "ordered" condition, we showed all payoff games to the participants in the same order. In the "randomized" condition, we showed the payoff games to each participant in a randomized order.

The results reported on in the main paper are that of the nonfiltered randomized treatment.

[1]Our motivation for the filtered procedure was to enable the estimation of the parameters of the cognitive hierarchy behavioral model for individual participants; however, this proved to be infeasible under realistically small values of the precision parameter λ.

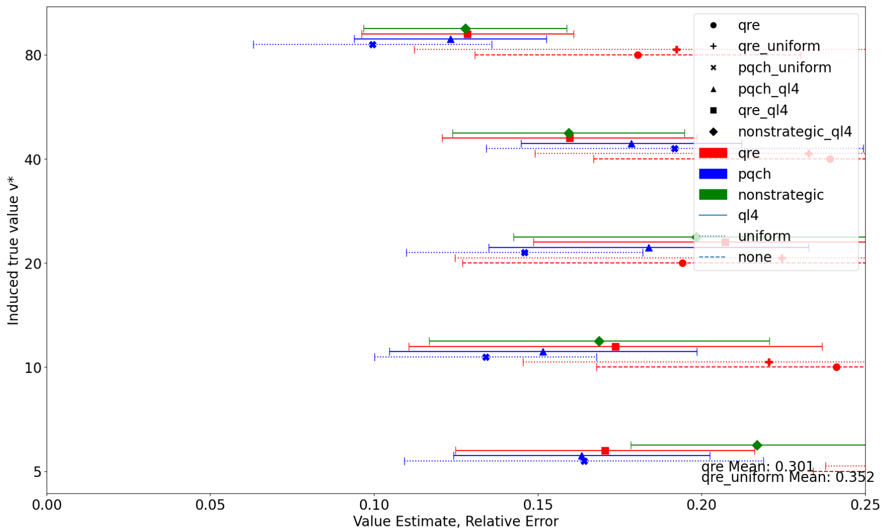

Figure A.1: Summary plot for randomized filtered treatment

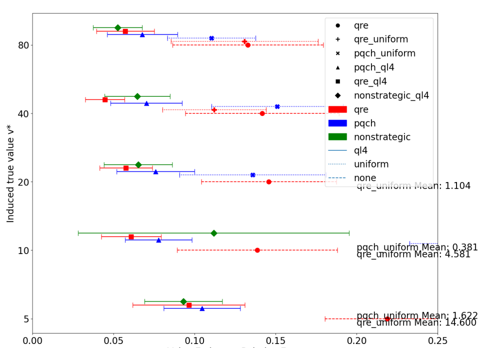

Figure A.2: Summary plot, ordered filtered treatment

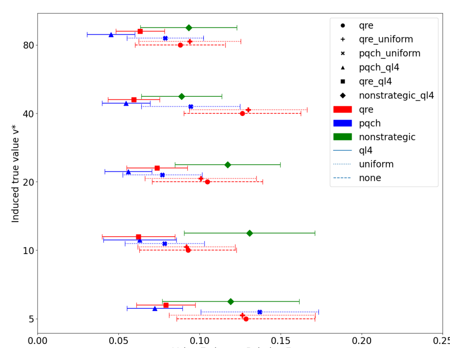

Figure A.3: Summary plot, ordered nonfiltered treatment

Table A.1: Relative error by v^* on randomized filtered treatment

Component		v^*				
Strategic	Non-strategic	5	10	20	40	80
QRE	none	0.30, (0.23, 0.37)	0.24, (0.17, 0.31)	0.19, (0.13, 0.26)	0.24, (0.17, 0.31)	0.18, (0.13, 0.23)
QRE	uniform	0.35, (0.24, 0.47)	0.22, (0.15, 0.30)	0.22, (0.12, 0.32)	0.23, (0.15, 0.32)	0.19, (0.11, 0.27)
PQCH	uniform	0.16, (0.11, 0.22)	0.13, (0.10, 0.17)	0.15, (0.11, 0.18)	0.19, (0.13, 0.25)	0.10, (0.06, 0.14)
PQCH	QL4	0.16, (0.12, 0.20)	0.15, (0.10, 0.20)	0.18, (0.14, 0.23)	0.18, (0.14, 0.21)	0.12, (0.09, 0.15)
QRE	QL4	0.17, (0.12, 0.22)	0.17, (0.11, 0.24)	0.21, (0.15, 0.27)	0.16, (0.12, 0.20)	0.13, (0.10, 0.16)
none	QL4	0.22, (0.18, 0.26)	0.17, (0.12, 0.22)	0.20, (0.14, 0.25)	0.16, (0.12, 0.19)	0.13, (0.10, 0.16)

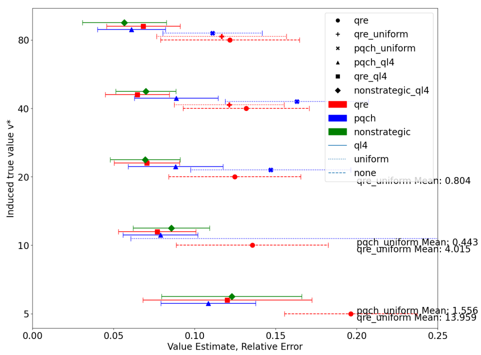

Figure A.4: Summary plot, combined filtered treatment, $n = 484$

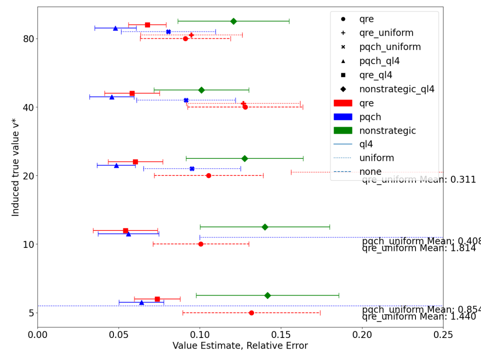

Figure A.5: Summary plot, combined nonfiltered treatment, $n = 364$

Table A.2

Component		v^*				
Strategic	Non-strategic	5	10	20	40	80
QRE	none	0.22, (0.18, 0.26)	0.14, (0.09, 0.19)	0.15, (0.10, 0.19)	0.14, (0.09, 0.19)	0.13, (0.09, 0.18)
QRE	uniform	14.60, (10.10, 19.10)	4.58, (2.91, 6.25)	1.10, (0.65, 1.56)	0.11, (0.08, 0.14)	0.13, (0.09, 0.18)
PQCH	uniform	1.62, (0.63, 2.61)	0.38, (0.23, 0.53)	0.14, (0.09, 0.18)	0.15, (0.11, 0.19)	0.11, (0.08, 0.14)
PQCH	QL4	0.10, (0.08, 0.13)	0.08, (0.06, 0.10)	0.08, (0.05, 0.10)	0.07, (0.05, 0.09)	0.07, (0.05, 0.09)
QRE	QL4	0.10, (0.06, 0.13)	0.06, (0.04, 0.08)	0.06, (0.04, 0.07)	0.04, (0.03, 0.06)	0.06, (0.04, 0.08)
none	QL4	0.09, (0.07, 0.12)	0.11, (0.03, 0.20)	0.07, (0.04, 0.09)	0.06, (0.04, 0.08)	0.05, (0.04, 0.07)

Table A.3: Relative error by v^* on ordered nonfiltered treatment

Component		v^*				
Strategic	Non-strategic	5	10	20	40	80
QRE	none	0.13, (0.09, 0.17)	0.09, (0.06, 0.12)	0.10, (0.07, 0.14)	0.13, (0.09, 0.16)	0.09, (0.06, 0.12)
QRE	uniform	0.13, (0.08, 0.17)	0.09, (0.06, 0.12)	0.10, (0.07, 0.14)	0.13, (0.09, 0.17)	0.09, (0.06, 0.13)
PQCH	uniform	0.14, (0.10, 0.17)	0.08, (0.05, 0.10)	0.08, (0.05, 0.10)	0.09, (0.06, 0.12)	0.08, (0.06, 0.10)
PQCH	QL4	0.07, (0.06, 0.09)	0.06, (0.04, 0.09)	0.06, (0.04, 0.07)	0.05, (0.04, 0.07)	0.05, (0.03, 0.06)
QRE	QL4	0.08, (0.06, 0.10)	0.06, (0.04, 0.08)	0.07, (0.06, 0.09)	0.06, (0.04, 0.08)	0.06, (0.05, 0.08)
none	QL4	0.12, (0.08, 0.16)	0.13, (0.09, 0.17)	0.12, (0.08, 0.15)	0.09, (0.06, 0.11)	0.09, (0.06, 0.12)

Table A.4: Relative error by v^* on combined filtered treatment

Component		v^*				
Strategic	Non-strategic	5	10	20	40	80
QRE	none	0.20, (0.16, 0.24)	0.14, (0.09, 0.18)	0.12, (0.08, 0.17)	0.13, (0.09, 0.17)	0.12, (0.08, 0.16)
QRE	uniform	13.96, (10.22, 17.69)	4.02, (2.51, 5.52)	0.80, (0.41, 1.20)	0.12, (0.09, 0.16)	0.12, (0.08, 0.16)
PQCH	uniform	1.56, (0.65, 2.46)	0.44, (0.06, 0.83)	0.15, (0.10, 0.20)	0.16, (0.12, 0.21)	0.11, (0.08, 0.14)
PQCH	QL4	0.11, (0.08, 0.14)	0.08, (0.06, 0.10)	0.09, (0.06, 0.12)	0.09, (0.06, 0.11)	0.06, (0.04, 0.08)
QRE	QL4	0.12, (0.07, 0.17)	0.08, (0.05, 0.10)	0.07, (0.05, 0.09)	0.06, (0.04, 0.08)	0.07, (0.05, 0.09)
none	QL4	0.12, (0.08, 0.17)	0.09, (0.06, 0.11)	0.07, (0.05, 0.09)	0.07, (0.05, 0.09)	0.06, (0.03, 0.08)

Table A.5: Relative error by v^* on combined nonfiltered treatment

Component		v^*				
Strategic	Non-strategic	5	10	20	40	80
QRE	none	0.13, (0.09, 0.17)	0.10, (0.07, 0.13)	0.11, (0.07, 0.14)	0.13, (0.09, 0.16)	0.09, (0.06, 0.12)
QRE	uniform	1.44, (0.62, 2.26)	1.81, (0.95, 2.68)	0.31, (0.16, 0.47)	0.13, (0.09, 0.16)	0.09, (0.06, 0.13)
PQCH	uniform	0.85, (-0.21, 1.92)	0.41, (0.10, 0.72)	0.10, (0.07, 0.13)	0.09, (0.06, 0.12)	0.08, (0.05, 0.11)
PQCH	QL4	0.06, (0.05, 0.08)	0.06, (0.04, 0.07)	0.05, (0.04, 0.06)	0.05, (0.03, 0.06)	0.05, (0.04, 0.06)
QRE	QL4	0.07, (0.06, 0.09)	0.05, (0.03, 0.07)	0.06, (0.04, 0.08)	0.06, (0.04, 0.08)	0.07, (0.06, 0.08)
none	QL4	0.14, (0.10, 0.19)	0.14, (0.10, 0.18)	0.13, (0.09, 0.16)	0.10, (0.07, 0.13)	0.12, (0.09, 0.16)

Table A.6: Summary Statistics for Experimental Treatments

Treatment	# Participants	Avg. Bonus	Avg. Total	Avg. Time (minutes:seconds)
Filtered ordered	303	2.50	4.00	9:30
Nonfiltered ordered	179	2.71	4.21	9:11
Filtered randomized	181	2.38	3.88	9:31
Nonfiltered randomized	185	2.58	4.08	10:18

A.2 Experimental Details

The experimental interface presented to MTurk Workers after acceptance of our HIT are shown in Figures A.6 to A.8.

A.3 Choosing λ for Our Nash Approximation

To select λ for our Nash approximation, we compared the value estimates and likelihoods of behavioral predictions for several values of λ. We choose the lowest value of λ at which both the value estimates and the likelihoods no longer change by increasing λ further. Figure A.9 shows this convergence in both the value estimate and behavioral predictions.

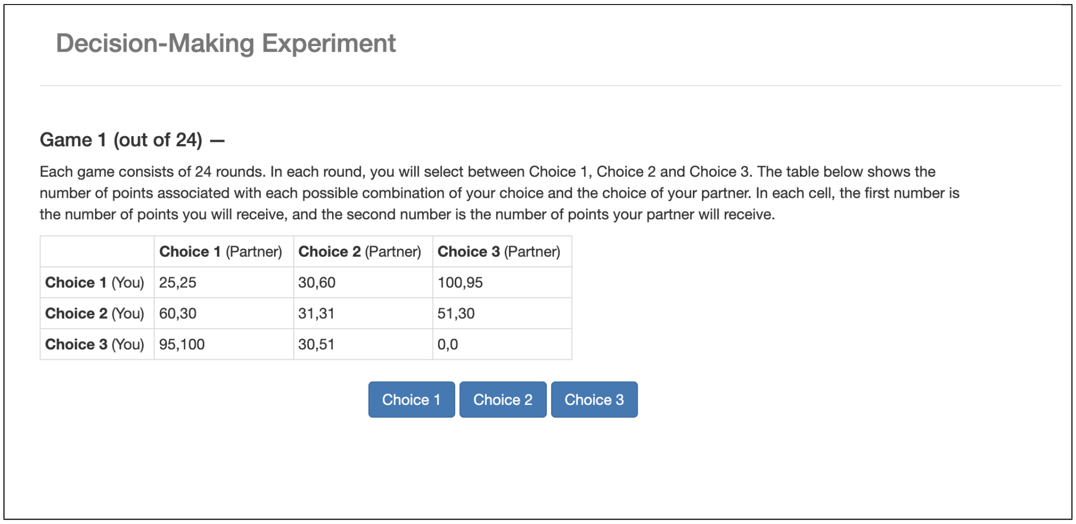

Figure A.6: The main experiment webpage presented to MTurk Participants.

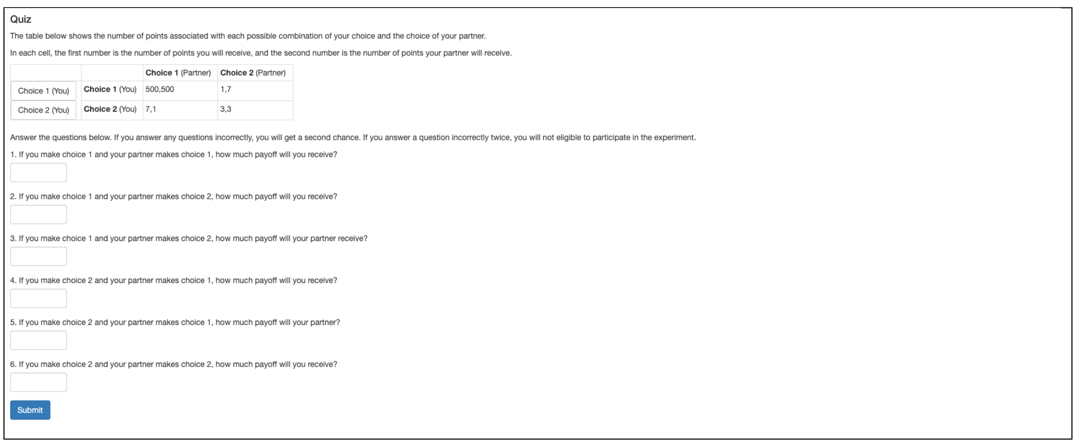

Figure A.7: The screening quiz presented to MTurk Participants to test their understanding of the task. Participants were allowed 3 attempts on the quiz before being rejected for the HIT.

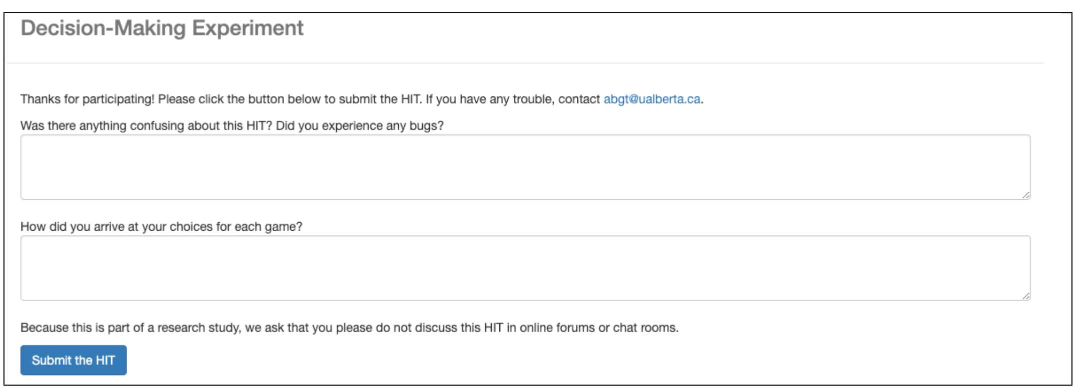

Figure A.8: The exit survey presented to MTurk Participants once they complete their HIT. The second prompt directs participants to fill out the reasoning behind their decisions.

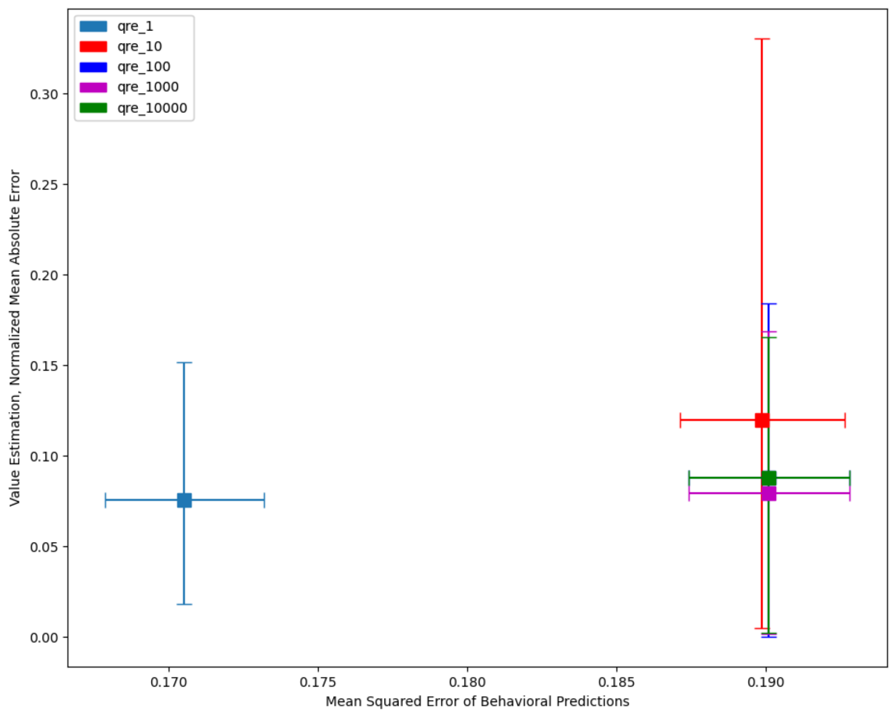

Figure A.9: QRE with fixed values of λ. As the error stabilizes around $\lambda = 100$, we use this as our Nash approximation

A.4 Allocation Game Mapping Algorithm

To convert our payoff games to arbitrary allocation games, we use algorithm 1

Algorithm 1 Random allocation game generation algorithm

```
Given set of payoff games G
Given value v*
for g ∈ G do
    for u(s_i, s_{-i}) ∈ g do
        Sample x ~ U(0, max(u(G))/v*
        Compute value p where p = u(s_i, s_{-i}) − x × v
        return (x, p)
    end for
end for
```

A.5 Additional Figures and Tables

A.5.1 Behavioral Parameter Estimates

This section gives additional information on the estimated behavioral parameters. Table A.7 gives the Poisson mean parameter τ we back out for different v^* across treatments, and A.8 gives the proportion β of non-strategic agents.

Table A.7: Estimated τ when using QCH-QL4, with τ indicating the rate parameter for a Poisson distribution specifying the proportion of agents of level k.

v^*	Combined Filtered	Combined Nonfiltered
5	0.32654 (0.22824 0.42485)	0.48453 (0.32990 0.63915)
10	0.40766 (0.27434 0.54098)	0.47658 (0.32487 0.62830)
20	0.35799 (0.27176 0.44421)	0.39598 (0.24480 0.54717)
40	0.41049 (0.35996 0.46102)	0.28461 (0.15282 0.41640)
80	0.35470 (0.17743 0.53198)	0.50106 (0.32416 0.67796)
Behavioral	0.53242 (0.41462 0.65023)	0.28660 (0.27126, 0.30193)

Table A.8: Estimated β when using QRE-$QL4$, with β indicating the proportion of agents who are non-strategic

v^*	Combined Filtered	Combined Nonfiltered
5	0.70950 (0.65445 0.76456)	0.57524 (0.45547 0.69502)
10	0.67130 (0.59895 0.74366)	0.56662 (0.46635 0.66688)
20	0.68092 (0.63536 0.72647)	0.45589 (0.40195 0.50983)
40	0.78786 (0.70963 0.86610)	0.56688 (0.52949 0.60427)
80	0.73638 (0.65591 0.81685)	0.41873 (0.39231 0.44515)
Behavioral	0.46723 (0.42931 0.50523)	0.52400 (0.49816, 0.54984)

Table A.9 shows the effect of not including payments within the allocation games. Even when using our best model in PQCH-QL4, the estimated values are incorrect, failing to scale to the correct value, especially at lower values of v^*. This issue does not seem to happen in QRE, but the relative error is worse than allocation games containing payments.

Table A.9: Raw value estimates when allocation games contain no payments for PQCH-QL4.

v^*	5	10	20	40	80
PQCH-QL4	53.75, (0.63, 106.87)	77.89, (21.59, 134.19)	44.38, (15.13, 73.64)	50.98, (29.21, 72.74)	104.07, (73.05, 135.08)
QRE	6.86, (6.79, 6.93)	11.54, (11.53, 11.55)	2.05, (1.93, 2.17)	33.99, (33.88, 34.1)	89.92, (89.83, 90)

Table A.10: Comparison of estimated v^* and λ and β behavioral parameters when finding the parameter(s) which maximize the likelihood (QRE Likelihood) vs. computing QRE against the empirical distribution (QRE Empirical). T-distributed confidence interval in brackets

	QRE Likelihood			**QRE Empirical**		
V^*	λ	β	v	λ	β	v
5	0.465 (0.127)	0.869 (0.045)	4.857 (0.327)	0.244 (0.209)	0.534 (0.116)	4.922 (0.227)
10	0.27 (0.102)	0.811 (0.081)	9.801 (0.555)	0.087 (0.103)	0.447 (0.089)	10.161 (0.474)
20	0.487 (0.148)	0.817 (0.079)	19.857 (1.12)	0.106 (0.089)	0.554 (0.1)	19.629 (0.735)
40	0.389 (0.147)	0.803 (0.075)	39.17 (1.879)	0.017 (0.006)	0.504 (0.085)	39.011 (1.454)
80	0.442 (0.131)	0.878 (0.06)	79.194 (3.053)	0.094 (0.161)	0.471 (0.078)	78.326 (2.649)