Non-Strategic Econometrics (for Initial Play)

Daniel Chui
Department of Computing Science
University of Alberta / Amii
Edmonton, Canada
dchui1@ualberta.ca

Jason Hartline
Department of Computer Science
Northwestern University
Evanston, USA
hartline@northwestern.edu

James R. Wright
Department of Computing Science
University of Alberta / Amii
Edmonton, Canada
james.wright@ualberta.ca

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, also known as structural estimation. 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. We apply models developed for predicting behavior towards 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 ground truth.

A key finding is that modelling non-strategic behavior, conventionally considered 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 popular quantal response equilibrium with 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.

KEYWORDS

Behavioral Game Theory; Bounded Rationality; Econometrics

1 INTRODUCTION

This paper[1] contributes to the study of behavioral models for initial play and of structural inference of preferences from behavior in games. The former has shown that rich, parameterized models of behavior, including non-strategic behavior, give better predictions than more classical equilibrium models. The latter uses strong equilibrium assumptions to estimate preferences from equilibrium

---

[1]The most recent version of this paper, including the appendix can be found at https://arxiv.org/abs/2208.06521

behavior. We conduct experiments that show preferences and behavioral models can be simultaneously inferred from initial play.

A main application of structural inference is in counterfactual estimation. After inferring preferences from behavioral data, counterfactual scenarios can be evaluated. Such evaluation can be used, for example, in mechanism design for optimizing over many mechanisms to find the one with the best equilibrium performance. A challenge for structural inference is that its predictions are only guaranteed to be accurate if the assumed model is correct. In contrast, randomized controlled trials – called A/B testing by technology firms – can directly evaluate a novel mechanism, but to optimize over many mechanisms the sample size that can be allocated to each mechanism is small. Chawla et al. [8], for example, showed that methods from structural inference have an exponential improvement for sample complexity in mechanism design over randomized controlled trials. However, this improvement comes with the aforementioned reliance on the accuracy of the model.

The literature in behavioral game theory considers models of behavior that relax the strong notion of equilibrium of classical game theory. Especially for initial play, i.e., for behavior of players who do not have prior experience playing a given game, classical notions of equilibrium are bad predictors of behavior while behavioral models, such as those in the quantal cognitive hierarchy (QCH) family, are good predictors [7, 21]. The QCH model combines quantal response (i.e., choosing actions with probability proportional to the exponentiated payoff of the action) with cognitive hierarchy (i.e., with levels of strategic thinking and agents at each level responding only to those at lower levels and with level-0 corresponding to non-strategic behavior). More recently, Wright and Leyton-Brown [28] showed that rich level-0 models significantly improve accuracy of predicted behavior in QCH.

This paper develops behavioral models and conducts experiments within the context of initial play to measure the accuracy of these models in predicting behavior and inferring preferences. The experiments again highlight the importance of rich level-0 models in modeling behavior in initial play. The classical model of level-0 behavior is uniform randomization. Uniform randomization ignores payoffs and is thus unhelpful for inferring preferences. Our results on predicted behavior reinforce those of Wright and Leyton-Brown [28], showing that rich models of level-0 behavior are better predictors. 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 predictive behavior and inferring preferences, and the choice of strategic model is not as important.

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 [28]. 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 with the classical equilibrium and behavioral models without rich level-0 behavior of Nash equilibrium, quantal response equilibrium [21], and quantal cognitive hierarchy (with uniform level-0 behavior). We also considered quantal response equilibrium augmented with the a non-strategic model, including the aforementioned quantal-linear4 model. Our models outperform these classical models with Nash equilibrium being the worst at both inference and prediction.

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 [e.g., 21, 23]. We assumed payoffs were derived from the classical single-dimensional linear model of auction theory where payoffs are given linearly as a value for units of a good (i.e., an allocation) and a payment [20] (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 game design 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.

2 RELATED WORK

The task of inferring preferences from observed data has generally been studied under the game theoretic assumption that players are in equilibrium [e.g. 2, 3, 16, 24]. 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 [11] and Goeree et al. [15] 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. [22] estimate private values from auction data without equilibrium assumptions, instead relying on a weaker assumption that agents use some form of no-regret learning. Similarly, Ling et al. [19] provide a framework to learn game parameters from actions in zero-sum games, but do not validate their results on empirical data.

The work that most closely resembles our own is that of Noti [23], which has a similar objective of using models from behavioral game theory to infer preferences from empirical data 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 either relies upon repetition across games, and/or does not estimate values at all.

Value estimation has also been studied empirically under conditions resembling initial play in the field of school matching. Value estimation is necessary for counterfactual evaluation of mechanisms and there are several papers [e.g. 1, 6, 17, 18] which attempt to infer preferences of agents to evaluate the welfare of alternative mechanisms. The way in which preferences are modelled vary between an equilibrium model to assuming all agents use simple behavioral rules. Notably, Calsamiglia et al. [6] 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 QRE+L0. Whereas non-strategic agents directly report their true preferences in the school choice setting, our framework allows us to consider scenarios where an indirect mechanism maps the preferences of non-strategic agents to actions.

Behavioral game theory aims to predict empirical human behavior better than traditional game theoretic concepts such as Nash equilibrium. One well known model, quantal response equilibrium [21], relaxes the strict optimization assumption made by Nash equilibrium, while maintaining the assumption that agents mutually respond to each others' strategies. In contrast, iterative behavioral models such as level-k models [25] and cognitive hierarchy [7] 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 [27] found that a combination of the two approaches, quantal cognitive hierarchy, performs best at predicting actual initial play in human subject experiments. 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 [28].

Another work that focuses on human behavior in initial play for normal form games is that of Fudenberg and Liang [14]. Starting from the premise that initial play is reasonably approximated by level-1 of iterated reasoning, they algorithmically generate games that are not captured well by level-1 reasoning and construct a decision tree based model that improves on the prediction of the modal action in normal form games over that of previous economic models. A key point to note is that their level-1 response relies on uniform randomization by non-strategic (level-0) agents. Further work by the same authors [12, 13] continues to rely on this uniform assumption. Our work focuses on inferring preferences of the agents and so exploits a richer model of level-0 behavior in which non-strategic agents are sensitive to their own preferences.

3 PRELIMINARIES AND 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) may arise.

To resolve 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 an allocation and 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 dataset of observations. We provide a more detailed explanation in Section 3.6.

Table 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	$\varnothing$	none	$\varnothing$
QRE	β, λ	uniform randomization	$\varnothing$
PQCH	τ, λ	quantal-linear4	$w_{L0}, \lambda_0,$
None	$\varnothing$		

3.1 Behavioral Models

Our behavioral models combine a component modelling strategic behavior with a component that models non-strategic behavior. At a high level, strategic behavior is that which responds to the anticipated actions of other agents while non-strategic behavior does not. The strategic models 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 [29]. With the exception of Nash equilibrium, each of the components just described have free parameters that must be learned from the data. We refer to these parameters as "behavioral parameters", to distinguish them from the valuation parameters describing agent preferences that we also estimate from data. Table 1 summarizes the behavioral parameters for each component included in our main evaluation.

3.2 Strategic Models

The strategic models of Nash equilibrium, quantal response equilibrium, and quantal cognitive hierarchy are defined formally below.

Definition 3.1 (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: $s_i \in BR_i(s_{-i})$.

The non-Nash equilibrium models that we consider are based on a relaxation of best response called quantal best response (QBR), in which agents play higher-utility strategies with higher probability (rather than strictly maximizing). QBR is parameterized by a *precision* (denoted by λ), indicating agents' sensitivity to utility differences.

Definition 3.2 (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

$$s_i(a_i) = \frac{\exp[\lambda \cdot u_i(a_i, s_{-i})]}{\sum_{a_i' \in A_i} \exp[\lambda \cdot u_i(a_i', s_{-i})]}. \tag{1}$$

Definition 3.3 (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 $i \in N$.

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 3.3, below).

Definition 3.4 (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

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

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 Poisson distribution is commonly used [e.g. 7, 12]. We do the same and estimate a mean parameter τ on a truncated Poisson distribution where the max level of an agent is 3.

Definition 3.5 (Poisson quantal cognitive hierarchy). Poisson quantal cognitive hierarchy is a specification of QCH in which the level

distribution L is specified by a Poisson distribution with the mean parameter τ:

$$L_{\tau;0:k} = \sum_{\ell=0}^{k=3} \frac{\text{Poisson}(\ell; \tau)}{\sum_{\ell'=0}^{k} \text{Poisson}(\ell'; \tau)}$$

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

3.3 Non-strategic Models

It is standard in the literature to assume that non-strategic agents randomize uniformly over their actions. Recently, Wright and Leyton-Brown [28] 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.[2] 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-linear4, each decision rule computes its prediction as a quantal response to the different actions' criterion values. In contrast, the predictions for linear4 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 [27]. 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 3.6 (Quantal-linear4). A quantal-linear4 (QL4) strategy for a player i in a game G with precision $\lambda_0 > 0$ and weights $w_{\max}, w_{\min}, w_{\text{eff}}, w_{\text{fair}}, w_{\text{unif}}$ 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-magnitude 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 across players that is possible by playing a given action. Formally,

$$f_i^{\max}(G)(a_i) = \frac{\exp[\lambda_0 \max_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i})]}{\sum_{a_i' \in A_i} \exp[\lambda_0 \max_{a_{-i} \in A_{-i}} u_i(a_i', a_{-i})]}$$

$$f_i^{\min}(G)(a_i) = \frac{\exp[\lambda_0 \min_{a_{-i} \in A_{-i}} u_i(a_i, a_{-i})]}{\sum_{a_i' \in A_i} \exp[\lambda_0 \min_{a_{-i} \in A_{-i}} u_i(a_i', a_{-i})]}$$

$$f_i^{\text{fair}}(G)(a_i) = \frac{\exp[\lambda_0 \max_{a_{-i} \in A_{-i}} \min_{j,j' \in N} d_{i,j}]}{\sum_{a_i' \in A_i} \exp[\lambda_0 \max_{a_{-i} \in A_{-i}} \min_{j,j' \in N} d_{i,j}']}$$

where

$$d_{i,j} = (u_j(a_i, a_{-i}) - u_{j'}(a_i, a_{-i}))$$

and

$$d_{i,j,A_i}' = (u_j(a_i', a_{-i}) - u_{j'}(a_i', a_{-i}))$$

$$f_i^{\text{eff}}(G)(a_i) = \frac{\exp[\lambda_0 \max_{a_{-i} \in A_{-i}} \sum_{j \in N} u_j(a_i, a_{-i})]}{\sum_{a_i' \in A_i} \exp[\lambda_0 \max_{a_{-i} \in A_{-i}} \sum_{j \in N} u_j(a_i', a_{-i})]}$$

$$f_i^{\text{unif}}(G)(a_i) = \frac{1}{|A_i|}.$$

3.4 Separating Models into Strategic and Non-strategic Components

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. 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 responds 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 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 $1 - \beta$. 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.

The parameters for each model are thus $\theta = (\theta_S, \theta_{NS})$, according to Table 1.

3.5 Estimation Methods

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 [5, 30]. 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.

---

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

3.5.1 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 $\hat{s}_{-i}^g$ defined by

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

For equilibrium models we maximize the following likelihood:

$$\log \mathcal{L}(\lambda, v, \beta, \lambda_0, w_{L0}) =$$
$$\sum_i \sum_{g \in G(i)} \log \left[\beta f^{L0}(G_g(v); \lambda_0, w_{L0})(a_i^g) \right.$$
$$\left. + (1 - \beta) QBR_i(s_{-i}^g \mid G_g(v), \lambda)(a_i^g) \right] \quad (2)$$

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; $\hat{s}_{-i}^g$ is the empirical distribution of play in allocation game g; and a_i^g is the action taken by participant i in game g.[34]

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. 4, 9, 15, 23]. 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 first estimating v and then $\lambda \mid v$ 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 incorrect by up to an order of magnitude (refer to Table 8 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,[5] and select the value that maximizes equation (2). 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 ensures 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.

3.5.2 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:

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

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.[6]

3.5.3 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 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.[7][8] 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. (4) 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 $\Pr(a_i^g \mid \ell_i = \ell, \theta)$.

$$\log \Pr(D \mid \theta, v) = \sum_i \sum_{g \in G(i)} \log \left[\sum_\ell \Pr(\ell_i = \ell \mid \theta) \Pr\left(a_i^{G_g(v)} \mid \ell_i = \ell, \theta \right) \right]$$
$$(4)$$

---

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

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

[5]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 5 in the appendix for details.

[6]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 9 in the appendix.

[7]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 $\ell_i = 0$.

[8]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.

In contrast, the likelihood for an otherwise-identical model in which each agent's level can vary between games is given by (5).

$$\log \Pr(D \mid \theta, v) = \sum_i \log \left[\sum_\ell \Pr(\ell_i = \ell \mid \theta) \prod_{g \in G(i)} \Pr\left(a_i^{G_g(v)} \mid \ell_i = \ell, \theta \right) \right] \tag{5}$$

When running our analysis on synthetic data we find that the likelihood of (5) is more numerically stable than that of (4), while returning a similar value estimate. We therefore report the parameters estimated using (5) in this paper. Our dataset is available for future research questions or models that require panel data.

3.6 Game Construction

In our setting, n subjects play a set of bimatrix payoff games $\mathcal{G}$. To simulate our subjects 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 our subjects 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 reuse of the same dataset for multiple sets of allocation games, as well as providing a ground truth value with which to evaluate model performance. A more detailed pseudocode explanation can be found in Section A.4 of the appendix.

3.7 Experimental Details

We tested our approach on experimental data collected from participants on Amazon Mechanical Turk (MTurk). 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. All participants had at least 95% HIT approval and had completed at least 100 HITs. 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.[9]

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

The payoff games were generated by randomly sampling payoffs from a uniform distribution on $[0, 100]$, and 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. Our qualitative results in these additional treatments were unchanged from those of the main treatment.

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 (USD), with the goal being to have all participants achieve an equivalent wage of at least \$10 per hour between the bonus and base payment if they had uniformly randomized and taken the maximum allowed time of 30 minutes.

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 assume that non-strategic agents uniformly randomize. 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.
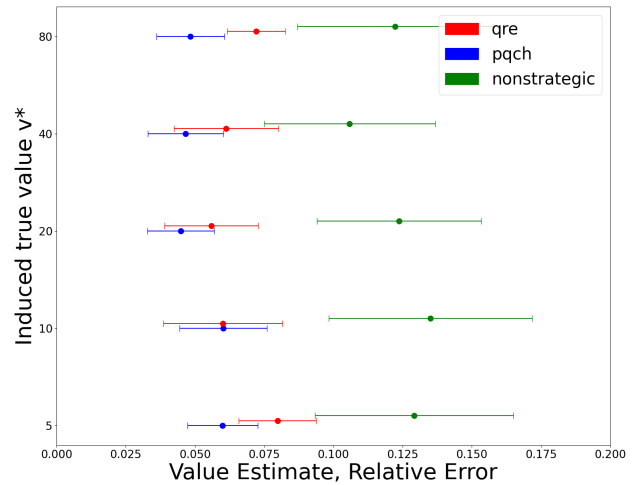

Figure 1: Summary plot showing values of v^* vs. the relative error for models with quantal-linear4 as the non-strategic componen The two left-most points for all values of v^* are models containing both a strategic component and quantal-linear4 as the non-strategic component; the green point indicates a model that assumes all agents are non-strategic in a quantal-linear4 manner with no strategic behavior.

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 $\mathcal{A}_G$ 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. 26]. We say that a model outperforms another in value estimation when the 95% confidence intervals do not overlap.

Figure 1 and Table 2 show the performance in value estimation across models, with Figure 1 being a visualization of the data in Table 2. 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 demonstrate the importance of obtaining accurate value estimates in Table 3. 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 for 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 an arbitrary level-0 model is not sufficient to improve performance, but a rich level-0 model is required.

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 Section 3.1 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 include linear4 from [27] 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 resulting model resulting from 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 [10]. Doing this allows us 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.

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 provides a possible explanation for why quantal-linear4 outperforms differentiable-linear4.

5 CONCLUSION

This paper 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 estimations from initial play than value estimates 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.

5.1 Future Work

There are a number of directions in which this work could be extended. We made a simplifying assumption that all agents shared

Table 2: 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 3: 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

Table 4: Percentage of the time that a model's relative error falls within 10% of v^*, across all values 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	0.1200 (0.0640 0.1760)	0.0960 (0.0480 0.1520)	0.00	0.0480 (0.0160 0.0880)	0.6880 (0.5040 0.8400)
QRE	0.4960 (0.4080 0.5680)	0.3280 (0.2560 0.4080)	0.00	0.7120 (0.6160 0.8160)	0.7040 (0.6000 0.8160)
PQCH	n/a	0.5200 (0.4160 0.6400)	0.00	0.6240 (0.4960 0.7440)	0.8440 (0.7200 0.9360)
None	n/a	n/a	0.00	0.00	0.4480 (0.3040 0.5840)

a homogeneous value, but an important future direction would be estimating heterogeneous values. Further to this direction, we could extend this work to estimating individual behavioral parameters of agents. This would allow us to model differences in behavior based on heterogeneous beliefs about the values of others, which could lead to better individual welfare. As mentioned in Section 3.5, we could also extend the approach of responding to empirically observed distributions in models of iterated reasoning, which would allow us to move beyond specifying a distribution over levels.

Using our framework of separating models into a strategic and non-strategic component, we could examine other models of non-strategic behavior (for example, using the model from Fudenberg and Liang [14] as a non-strategic model) beyond those discussed in this paper. Finally, further extending non-strategic behavioral models is an important direction for future work. This could take the form of extending QL4 to be more predictive, or evaluating domain-specific models of nonstrategic behavior.

ACKNOWLEDGEMENTS

REFERENCES

[1] Nikhil Agarwal and Paulo Somaini. 2018. Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica* 86, 2 (2018), 391–444.

[2] Susan Athey and Philip A Haile. 2007. Nonparametric approaches to auctions. *Handbook of econometrics* 6 (2007), 3847–3965.

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

[4] Patrick Bajari and Ali Hortacsu. 2005. Are structural estimates of auction models reasonable? Evidence from experimental data. *Journal of Political economy* 113, 4 (2005), 703–741.

[5] Richard H Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on scientific computing* 16, 5 (1995), 1190–1208.

[6] Caterina Calsamiglia, Chao Fu, and Maia Güell. 2020. Structural estimation of a model of school choices: The boston mechanism versus its alternatives. *Journal of Political Economy* 128, 2 (2020), 642–680.

[7] Colin F Camerer, Teck-Hua Ho, and Juin-Kuan Chong. 2004. A cognitive hierarchy model of games. *The Quarterly Journal of Economics* 119, 3 (2004), 861–898.

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

[9] Yefen Chen, Xuanming Su, and Xiaobo Zhao. 2012. Modeling bounded rationality in capacity allocation games with the quantal response equilibrium. *Management Science* 58, 10 (2012), 1952–1962.

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

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

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

[13] Drew Fudenberg, Jon Kleinberg, Annie Liang, and Sendhil Mullainathan. 2019. Measuring the completeness of theories. (2019).

[14] Drew Fudenberg and Annie Liang. 2019. Predicting and understanding initial play. *American Economic Review* 109, 12 (2019), 4112–41.

[15] Jacob K Goeree, Charles A Holt, and Thomas R Palfrey. 2002. Quantal response equilibrium and overbidding in private-value auctions. *Journal of Economic Theory* 104, 1 (2002), 247–272.

[16] Emmanuel Guerre, Isabelle Perrigne, and Quang Vuong. 2000. Optimal nonparametric estimation of first-price auctions. *Econometrica* 68, 3 (2000), 525–574.

[17] Yinghua He. 2015. Gaming the Boston school choice mechanism in Beijing. (2015).

[18] Il Myoung Hwang. 2015. *A robust redesign of high school match*. Ph.D. Dissertation. The University of Chicago.

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

[20] Daniel McFadden. 1981. Econometric models of probabilistic choice. *Structural analysis of discrete data with econometric applications* 198272 (1981).

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

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

[23] Gali Noti. 2021. 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. 5637–5646.

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

[25] Dale O Stahl and Paul W Wilson. 1994. Experimental evidence on players' models of other players. *Journal of economic behavior & organization* 25, 3 (1994), 309–327.

[26] Ian H Witten and Eibe Frank. 2002. Data mining: practical machine learning tools and techniques with Java implementations. *Acm Sigmod Record* 31, 1 (2002), 76–77.

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

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

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

[30] Ciyou Zhu, Richard H Byrd, Peihuang Lu, and Jorge Nocedal. 1997. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on mathematical software (TOMS)* 23, 4 (1997), 550–560.