A Bayesian Level-k Model in n-Person Games

Teck-Hua Ho, So-Eun Park, and Xuanming Su *

January 9, 2020

In standard models of iterative thinking, players choose a fixed rule level from a fixed rule hierarchy. Non-equilibrium behavior emerges when players do not perform enough thinking steps. Existing approaches, however, are inherently static. This paper introduces a Bayesian level-k model, in which level-0 players adjust their actions in response to historical game play, while higher level thinkers update their beliefs on opponents' rule levels and best-respond with different rule levels over time. As a consequence, players choose a dynamic rule level (i.e., sophisticated learning) from a varying rule hierarchy (i.e., adaptive learning). We apply our model to existing experimental data on three distinct games: p-beauty contest, Cournot oligopoly, and private-value auction. We find that both types of learning are significant in p-beauty contest games, but only adaptive learning is significant in the Cournot oligopoly, and only sophisticated learning is significant in the private-value auction. We conclude that it is useful to have a unified framework that incorporates both types of learning to explain dynamic choice behavior across different settings.

1 Introduction

Over the past few decades, evidence accumulated from laboratory and field experiments has shown that players' actual behavior may significantly deviate from equilibrium in a systematic manner. That is, despite non-trivial financial incentives, subjects frequently do not play equilibrium. For example, the first round offer in sequential bargaining games often neither corresponds to the equilibrium offer nor is accepted immediately (Binmore, Shaked and

*Ho: National University of Singapore, University Hall, 21 Lower Kent Ridge Road, Singapore 119077, teck@nus.edu.sg. Park: University of British Columbia, 2053 Main Mall, Vancouver, British Columbia V6T1Z2 Canada, soeun.park@sauder.ubc.ca. Su: University of Pennsylvania, 3730 Walnut Street, Philadelphia, PA 19104, xuanming@wharton.upenn.edu. Authors are listed in alphabetical order. Direct correspondence to any of the authors. We thank Colin Camerer, Vince Crawford, David Levine, and seminar participants at the National University of Singapore, the University of British Columbia and the University of California, Berkeley for their helpful comments. We thank Rehan Ali for his excellent editorial help and Hang Wu and Zhang Xing for their research assistance in conducting laboratory experiments.

Sutton, 1985; Ochs and Roth, 1989); centipede games often proceed to intermediate stages rather than end immediately (McKelvey and Palfrey, 1992); the unique Nash equilibrium is not instantly chosen in private-value auctions (Goeree et al., 2002); many players do not choose the unique iterative dominance solution in dominance solvable games such as the p-beauty contest or Cournot oligopoly games (Fouraker and Siegel, 1963; Nagel, 1995; Stahl and Wilson, 1994, 1995; Ho et al., 1998; Costa-Gomes et al., 2001; Bosch-Doménech et al., 2002; Costa-Gomes and Crawford, 2006; Costa-Gomes and Weizsäcker, 2008).

These games have been extensively studied because they capture the essence of various settings in the real world. The p-beauty contest game, for example, parallels the stock market in that investors make their investment decisions not based on what they believe the company's fundamental value is but rather on what everybody else in the market assesses it to be. The Cournot oligopoly game closely resembles the competitive behavior of firms in commodities markets such as petroleum, electricity, chemicals, cement, sugar and rice. Auctions are widely used in a variety of settings such as government procurement auctions, mineral rights sales, U.S. treasury bills and online advertising platforms (McAfee and McMillan, 1987). As a consequence, studying the choice dynamics and interactions among players in these games can shed light on how decision makers actually behave in real life.

To explain non-equilibrium behavior in these games, cognitive hierarchy (Camerer et al., 2004) and level-k (Nagel, 1995; Stahl and Wilson, 1994, 1995; Costa-Gomes et al., 2001; Costa-Gomes et al., 2009; Costa-Gomes and Crawford, 2006; Crawford, 2003; Crawford and Iriberri, 2007b) models have been developed; for a comprehensive review, see Crawford et al. (2013). These models assume that each player chooses a fixed rule from a fixed rule hierarchy. The rule hierarchy is defined iteratively by assuming that a level-k rule best-responds to a lower level rule (e.g., level-$(k-1)$ where $k \geq 1$) and by specifying a priori a fixed level-0 rule (e.g., uniform randomization among all possible strategies). Non-equilibrium behavior emerges when players do not perform enough thinking steps.[1]

Let us illustrate how to apply these models using the p-beauty contest game. In the p-

[1] The level-k models have also been used to explain non-equilibrium behavior in games with asymmetric information (e.g., Camerer et al., 2004; Crawford and Iriberri, 2007a.; Crawford et al., 2009; Brocas et al., 2010; Östling et al., 2011; Brown et al., 2012). Goldfarb and Yang (2009) and Goldfarb and Xiao (2011) apply the cognitive hierarchy model to capture heterogeneity in the abilities of firms or managers, and use it to predict a firm's long-term success in oligopolistic markets.

beauty contest game, n players simultaneously choose numbers ranging from 0 to 100. The winner is the player whose number is the closest to the *target number*, defined as p times the average of all chosen numbers, where $0 < p < 1$. A fixed reward goes to the winner; in the case of a tie, it is divided evenly among winners. This game is dominance-solvable and the unique number that survives the iterative elimination process is 0. To apply the standard level-k model, assume that the level-0 player randomizes over all possible choices between 0 and 100, and thus chooses 50 on average. Best-responding to level-0 opponents, the level-1 player seeks to hit the target number in order to maximize the payoff. Specifically, the level-1 player chooses $x^* = p \cdot \frac{x^*+(n-1)\cdot 50}{n}$ as a best response to level-0 opponents, taking into account the player's own influence on the group mean.[2] Proceeding iteratively, a level-k player will choose $\left(\frac{p\cdot(n-1)}{n-p}\right)^k \cdot 50$, which converges to 0 as k increases. In other words, higher level players choose smaller numbers and these choices converge to 0, the unique Nash equilibrium.

However, standard level-k models were not designed to account for dynamic trends in observed data in repeated games. In laboratory experiments, subjects' choices are initially far from equilibrium but move closer to it over time (Ho et al., 1998; Capra et al., 1999; Goeree and Holt, 2001). Existing models are static, so a level-k player will always choose rule k. As a result, these static models imply a *fixed* distribution of choices as the proportion of players choosing each rule remains unchanged over time. Figure 1 compares data from the first and last rounds of the p-beauty contest game from Ho et al. (1998). As shown, subjects' choices move closer to equilibrium over time, from 2% choosing zero in the first round to 13% in the last round. A Kolmogorov-Smirnov test rejects the null hypothesis that the distributions of choices in the first and last rounds are identical ($p < 0.001$). As a consequence, any model seeking to capture this shift in behavior over time needs to be dynamic in nature.

[INSERT FIGURE 1 HERE.]

One approach to capture dynamic behavior in games is to incorporate adaptive learning (Fudenberg and Levine, 1998; Erev and Roth, 1998; Camerer and Ho, 1999; Hück, Normann,

[2]Some prior research (e.g., Nagel, 1995) ignores the player's own influence on the group average by simply assuming that the level-1 player chooses $p \cdot 50$ and the level-k player chooses $p^k \cdot 50$. This approximation is good only when n is large. For example, when $n = 100$ and $p = 0.7$ the level-1 player's exact best response is 34.89, which is close to 35.

and Oechssler, 1999), allowing players to adjust their behavior in response to observed history. Specifically, consider level-0 players who change their choices over time according to some adaptive learning model (e.g., Cournot dynamics where players best-respond to their most recent observation). As long as observed actions approach equilibrium, the choices made by these adaptive level-0 players, as well as by higher level players, will also converge to equilibrium. Thus, by incorporating adaptive learning, level-k models can incorporate a varying rule hierarchy and predict convergence towards equilibrium, explaining the main feature of the data.

However, the adaptive learning approach still assumes that each player chooses a fixed rule over time. That is, each level-k player always thinks k steps and their choices shift towards equilibrium as the level-0 rule converges to equilibrium. In other words, equilibration arises from shifts in the level-0 rule (and the iterated rule hierarchy), but not from advances in the number of levels players think.

Another way to model the equilibration process is to incorporate sophisticated learning, i.e., allowing players to choose higher rule levels over time (Stahl, 1996, 1999, 2000, 2001; Stahl and Haruvy, 2002; Weber, 2013; Alaoui and Penta, 2016). For example, a player choosing a level-k rule may advance to a level-$(k+1)$ rule in subsequent rounds. Since the rule hierarchy is defined by iterative best responses, choices shift closer to equilibrium as players advance in rule levels. Thus, sophisticated learning, like adaptive learning, can also explain the dynamic behavior observed in the data.

In this paper, we develop a structural Bayesian level-k model to incorporate both adaptive and sophisticated learning. We refer to level-0 players as adaptive learners and higher level players as sophisticated learners. The following sequence of events takes place in every round. First, adaptive learners choose their level-0 actions by responding to observed game history. Starting from the level-0 rule, we construct the rule hierarchy using iterative best responses. Second, each sophisticated player forms a conjecture about an opponent's rule level by sampling from a belief distribution. Then, each sophisticated player best responds accordingly, i.e., chooses the level-(k+1) action if the opponent's rule level is conjectured to be level-k. At this point, all players have chosen their actions. Next, at the end of each round, upon observing other players' actions, sophisticated players infer their opponents' actual rule levels (which may differ from their conjectures). Finally, sophisticated learners use these inferred opponents' rule levels to update their own belief distribution via Bayesian updating. This concludes the chronology for this round.

In summary, level-0 players exhibit adaptive learning by changing their choices in response to observed game history, while higher level players exhibit sophisticated learning by updating their beliefs on opponents' rule levels after each round in a Bayesian manner. Since the level-0 rule varies over time, the rule hierarchy varies over time. Additionally, since beliefs update over time, sophisticated players' rule levels vary over time as their sampling belief distribution changes. In this way, players choose varying rule levels from a varying rule hierarchy.

Our model is general and applies to all games where rule hierarchies are iteratively defined. The Bayesian level-k model nests existing non-equilibrium models such as adaptive learning models and static level-k models as special cases. This nested structure allows us to disentangle adaptive learning and sophisticated learning.

To examine the practical applicability of our model, we fit it to three different classes of games: the classic p-beauty contest, the Cournot oligopoly and the private-value auction. Estimation results from these games show that our Bayesian level-k model describes subjects' dynamic behavior better than its special cases but in different ways across the three games. Specifically, both types of learning are significant in p-beauty contest games, while only one type of learning is dominant in the other two games: adaptive learning in the Cournot oligopoly, and sophisticated learning in the private-value auction. Overall, we find that it is crucial to have a unified framework that incorporates both types of learning.

The rest of the paper is organized as follows. Section 2 sets up the preliminaries. Section 3 formulates the Bayesian level-k model and its special cases. Sections 4, 5 and 6 illustrate the empirical results from applying our model to the existing experimental data on the p-beauty contest, the Cournot oligopoly and the private-value auction, respectively. Section 7 is the conclusion.

2 Preliminaries

2.1 Notations

We begin with notations. We consider the class of n-player games. Players are indexed by i ($i = 1, 2, \ldots, n$). Player i's strategy space is denoted by S_i and the set of mixed strategies is denoted by $\triangle S_i$. Players play the game repeatedly for a total of T rounds. Player i's best response correspondence, given how she anticipates opponents acting, is denoted by $BR_i(\cdot) : (\triangle S_1 \times \cdots \times \triangle S_{i-1} \times \triangle S_{i+1} \times \cdots \times \triangle S_n) \rightarrow \triangle S_i$. Player i's choice observed at time t is denoted by $x_i(t) \in S_i$ and the vector of all players' choices at time t, excluding player i's choice, is denoted by $\mathbf{x}_{-i}(t) = (x_1(t), \ldots, x_{i-1}(t), x_{i+1}(t), \ldots, x_n(t))$.

2.2 Rule Hierarchy

Players choose rules $k \in \mathbb{Z}_+^0 = \{0, 1, 2, \ldots\}$ from iteratively-defined discrete rule hierarchies, and choose actions corresponding to these rules. Each player has a rule hierarchy, and each rule hierarchy is iteratively defined such that a level-$(k+1)$ rule is a best response to the opponents' level-k rules (Stahl and Wilson, 1994; Nagel, 1995; Ho et al., 1998; Crawford et al., 2013). In symmetric n-player games, all players' rule hierarchies are identical, but they may not be in general.

Let $k_i(t)$ be player i's chosen rule and $a_i^{k_i(t)}(t) \in \triangle S_i$ be the action corresponding to rule $k_i(t)$ under player i's rule hierarchy at time t. At each time t, $a_i^{k_i(t)}(t)$ is iteratively defined from the level-0 rule's action, $a_i^0(t)$, which will be described in detail below. We further take a Bayesian approach to model how player i chooses $k_i(t)$ and how $a_i^0(t)$ varies over time as the game evolves.

2.3 Iterative Best Response

Each player i has a conjecture about which rule level opponents will choose at time t (we elaborate on how this conjecture is formed in the next section). We denote player i's conjectured rule at time t by $d_i(t)$. Hence, player i expects all opponents j ($\forall j \neq i$) to choose action $a_j^{d_i(t)}(t)$ that corresponds to rule $d_i(t)$ at time t. Player i best-responds by choosing the strategy yielding the highest payoff as:

$$a_i^{d_i(t)+1}(t) \quad = \quad BR_i(a_1^{d_i(t)}(t), \ldots, a_{i-1}^{d_i(t)}(t), a_{i+1}^{d_i(t)}(t), \ldots, a_n^{d_i(t)}(t))$$

If such best response is unique, $a_i^{d_i(t)+1}(t) \in \triangle S_i$ is a pure strategy and player i chooses $x_i(t) = a_i^{d_i(t)+1}(t)$. However, if there are multiple strategies with the highest expected payoff, $a_i^{d_i(t)+1}(t) \in \triangle S_i$ is defined as a mixed strategy with equal probability weights on these multiple strategies. That is, player i randomly chooses $x_i(t)$ among her indifferent best responses with equal probability.

In a symmetric game, we have $a_i^0(t) = a^0(t)$ for $\forall i$. This implies that all players who best respond with the same rule level have identical choices, i.e., $a^{k+1}(t) = a_i^{k+1}(t) = BR_i(a^k(t), a^k(t), \ldots, a^k(t))$. As a consequence, in a symmetric game, all players' rule hierarchies are identical.

3 Bayesian Level-k Model

This section develops a structural model with both sophisticated learning and adaptive learning. We allow players to be heterogeneous (adaptive vs. sophisticated). Specifically, a fraction α of the players are level-0, adaptive learning players, and the remaining $1 - \alpha$ fraction are higher level, sophisticated players. Below we describe the choice dynamics of these two types of players in our model.

3.1 Adaptive, Level-0 Players

With probability α, players are level-0 and learn adaptively. We incorporate adaptive learning by allowing this α fraction of level-0 players to change their choices over time. That is, adaptive, level-0 player i chooses $a_i^0(t)$ in round t. The functional form of $a_i^0(t)$ depends on the type of adaptive learning dynamics we use and can be specified independently of our model. As $a_i^0(t)$ varies over time, since a rule hierarchy is iteratively defined, the choices of all higher-level players will also change even if players choose fixed rule levels over time.

Incorporating non-sophisticated, adaptive players into the Bayesian level-k model this way allows us to nest any adaptive learning model as a special case. In addition, it empirically controls for the potential existence of adaptive learning and hence determines the respective roles of adaptive and sophisticated learning in describing players' choice dynamics.

3.2 Sophisticated Player's Conjecture of Opponents' Rule

With probability $1-\alpha$, players are higher level players (level 1 or above) and learn sophisticatedly. Each sophisticated player has a belief distribution over rule levels opponents will choose. We denote player i's belief distribution at the end of time t by $\mathbf{B}_i(t) = (B_i^0(t), B_i^1(t), B_i^2(t), \ldots)$, where $B_i^k(t)$ is player i's conjectured probability of opponents choosing rule k. From $\mathbf{B}_i(t)$, player i samples a *single* rule $d_i(t+1) \in \{0, 1, 2, \ldots\}$ with probability $B_i^{d_i(t+1)}(t)$ at the beginning of time $t+1$.[3] Player i then conjectures that *all* opponents will choose rule $d_i(t+1)$ and that opponent j $(\forall j)$ will choose action $a_j^{d_i(t+1)}(t+1)$ at time $t+1$.

Player i then best-responds to this prediction by choosing $x_i(t+1)$ according to $a_i^{d_i(t+1)+1}(t+1)$ as defined in Section 2.3. Hence, player i's chosen rule is exactly one level higher than the sampled rule from the belief distribution on the opponents' rule level. Since the lowest rule level that players can sample is 0, sophisticated players always choose rules of level 1 or higher.

We highlight a key simplifying assumption in our model. We assume that given the belief distribution (of opponent rule levels), sophisticated players sample a rule from the distribution and best respond to the sampled rule, instead of directly best responding to the belief distribution. This assumption ensures that the best response belongs to the rule hierarchy (since it is exactly one level above the sampled rule level). In contrast, best responding to a distribution of rule levels may yield an action that falls outside the rule hierarchy.[4] Hence, our assumption confers both internal consistency and analytical tractability.

[3] In a two-person game, the single-rule sampling assumption is reasonable. In an n-player game, players might sample multiple rules, e.g., one sampling for each opponent. If this is indeed the case, a player samples for each opponent from a separate distribution (updated differently for each opponent). Prior research suggests that this is unlikely to be true. In fact, opponents' choices are often found to be perfectly correlated and players tend to treat opponents as a single aggregate player (Ho et al., 1998).

[4] In the cognitive hierarchy (CH) model (Camerer et al., 2004), rules are iteratively defined as best responses to a belief distribution over lower rules. However, a similar definition is not feasible in the Bayesian level-k model since the belief distribution dynamically changes for each player following the player's own learning path, while it is static in the CH model for players of all levels.

3.3 Sophisticated Player's Bayesian Belief Updating

In the Bayesian level-k model, sophisticated players update their beliefs over time. Specifically, at the end of round t ($t \geq 1$), player i observes her opponents' choices, $\mathbf{x}_{-i}(t)$, infers what *common* rule level corresponds to these choices in round t, and updates her belief accordingly.

Each player maintains a probability distribution over *discrete* rule levels that opponents may choose and updates it every round in a Bayesian manner. Let player i's initial prior belief about opponents' rule levels be $\mathbf{B}_i(0) = (B_i^0(0),\ B_i^1(0), B_i^2(0), \ldots)$ before the game begins. We operationalize this initial prior using a set of rule counts $\mathbf{N}_i(0) = (N_i^0(0), N_i^1(0), N_i^2(0), \ldots)$ such that $B_i^k(0) = \frac{N_i^k(0)}{\sum_k N_i^k(0)}$. Belief updating is then equivalent to updating these rule counts based on observed chosen rules. Let player i's rule counts at the end of round t be $\mathbf{N}_i(t) = (N_i^0(t), N_i^1(t), N_i^2(t), \ldots)$ where $N_i^k(t)$ denotes player i's cumulative rule count for rule k. Then, $B_i^k(t) = \frac{N_i^k(t)}{\sum_k N_i^k(t)}$. In a symmetric game, players have common priors, i.e., $\mathbf{B}_i(0) = \mathbf{B}(0)$ and $\mathbf{N}_i(0) = \mathbf{N}(0),\ \forall i$.

After the game play in time t, upon observing all other players' choices $\mathbf{x}_{-i}(t)$, player i infers the probability each rule level could give rise to these observed choices $\mathbf{x}_{-i}(t)$ and updates her rule counts accordingly. Specifically, the probability that all opponents' choices $\mathbf{x}_{-i}(t)$ result from a rule k, $P_{-i}^k(t)$, is given by:

$$P_{-i}^k(t) = \frac{B_i^k(t-1) \cdot \prod_{j \neq i} f_j(x_j(t) \,|\, a_j^k(t))}{\sum_k B_i^k(t-1) \cdot \prod_{j \neq i} f_j(x_j(t) \,|\, a_j^k(t))}$$

where $f_j(x_j(t) \,|\, a_j^k(t))$ represents the probability density of opponent j choosing $x_j(t)$ given that the opponent is conjectured to choose rule k and action $a_j^k(t)$. This probability density has a straightforward definition whether $a_j^k(t)$ is a pure strategy or a mixed strategy over multiple best responses as in Section 2.3. For example, if $a_j^k(t)$ is a pure strategy and $f_j(x_j(t) \,|\, a_j^k(t)) = \mathcal{N}(x_j(t) \,|\, a_j^k(t), \sigma)$ is a normal distribution with mean $a_j^k(t)$ and standard deviation σ, the normal distribution evaluated at $x_j(t)$, i.e., $\mathcal{N}(x_j(t) \,|\, a_j^k(t), \sigma)$, denotes the probability that opponent j chooses $x_j(t)$ when conjectured to choose $a_j^k(t)$; if $a_j^k(t)$ is a mixed strategy with equal probabilities over multiple best responses, $f_j(x_j(t) \,|\, a_j^k(t))$ simply becomes the expected value of multiple conditional normal densities, each with mean at each of these best responses and with standard deviation σ. In essence, the farther $x_j(t)$ is from

$a_j^k(t)$, the smaller the probability. In a symmetric game, $f_j(\cdot) = f(\cdot), \forall j$.

Then, player i updates her cumulative rule count for a rule k at the end of time t by:

$$N_i^k(t) = \rho \cdot N_i^k(t-1) + (n-1) \cdot P_{-i}^k(t)$$

where ρ is a memory decay factor. These updated rule counts, $\mathbf{N}_i(t)$, yield a posterior belief $B_i^k(t) = \frac{N_i^k(t)}{\sum_k N_i^k(t)}$. Note that this updating process is consistent with Bayesian updating involving a multinomial distribution with a Dirichlet prior (Fudenberg and Levine, 1998; Camerer and Ho, 1999; Ho and Su, 2013).

Note that the updating process requires a feedback on opponents' choices after each round. If this feedback is in the aggregate (e.g., only the average choice of opponents is given), then one approach researchers can adopt is that each player i treats all opponents as symmetric, infers $x_j(t)$, which is identical for all $j \neq i$ from the aggregate feedback provided, and uses $x_j(t)$ $(\forall j \neq i)$ in her Bayesian updating.

Let us emphasize how the memory decay factor (ρ) and rule counts ($\mathbf{N}_i(t)$) interplay with each other. Let $N_i(0)$ be the sum of player i's rule counts before the game starts. In other words, $N_i(0) = \sum_k N_i^k(0)$ and hence $N_i^k(0) = N_i(0) \cdot B_i^k(0)$. $N_i(0)$ captures the strength of prior belief $\mathbf{B}_i(0)$. That is, $N_i(0) = \infty$ and $\rho = 1$ imply that the prior belief is completely sticky and subsequent observations have no influence at all, i.e., $N_i^k(t) = N_i^k(0),\ \forall t$. On the contrary, $N_i(0) < \infty$ and $\rho = 0$ implies a weak prior and that the posterior belief is driven entirely by observed rule levels.

Note that players' chosen rules are not fixed and can vary over time. At each time t, player i's rule level depends on sampled rule $d_i(t)$, which is moderated by belief $\mathbf{B}_i(t-1)$; this belief is then updated to $\mathbf{B}_i(t)$ using the observed choices $\mathbf{x}_{-i}(t)$ after time t. In this regard, players learn sophisticatedly and change rule levels in response to opponents' previously chosen rules.

3.4 Model Summary and Special Cases

The Bayesian level-k model ($BL_k(t)$) has the following key features:

1. (Adaptive Level-0 Players) Players can be either adaptive or sophisticated. Level-0 adaptive players choose $a_i^0(t)$ at time t.

2. (Sophisticated Player's Conjecture of Opponents' Rule) At the beginning of round t, player i samples a rule $d_i(t)$ from $\{0, 1, 2, \ldots\}$ with probability $B_i^{d_i(t)}(t-1)$. Player i then conjectures that all opponents will choose rule $d_i(t)$. Consequently, player i best responds in round t by choosing $a_i^{d_i(t)+1}(t) = BR_i(a_1^{d_i(t)}(t), \ldots, a_{i-1}^{d_i(t)}(t),\, a_{i+1}^{d_i(t)}(t), \ldots,\, a_n^{d_i(t)}(t))$.
3. (Sophisticated Player's Belief Updating) At the end of round t, player i's belief about her opponents' rule levels may change. Specifically, upon observing choices $\mathbf{x}_{-i}(t)$, player i updates her rule counts to: $N_i^k(t) = \rho \cdot N_i^k(t-1) + (n-1) \cdot P_{-i}^k(t)$ where $P_{-i}^k(t) = B_i^k(t-1) \cdot \prod_{j \neq i} f_j(x_j(t) \,|\, a_j^k(t)) / \left(\sum_k B_i^k(t-1) \cdot \prod_{j \neq i} f_j(x_j(t) \,|\, a_j^k(t)) \right)$.

Our $BL_k(t)$ model, which incorporates dynamic rule levels over a dynamic rule hierarchy, nests three classes of non-equilibrium structural models described below.

3.4.1 Static Level-k with Adaptive Level-0 Players ($L_k(t)$)

We first describe a special case with static rule levels over a dynamic rule hierarchy. By setting $N_i(0) = \infty$, $\forall i$ and $\rho = 1$, we suppress sophisticated learning, so players do not adjust their rule levels. However, the level-0 rule $a_i^0(t)$ adapts over time, inducing an adaptive rule hierarchy. We denote this model by $L_k(t)$. This model allows for an α fraction of adaptive players and a $1-\alpha$ fraction of higher level players who iteratively best respond to level-0 learning dynamics.

By further restricting $\alpha = 1$, $L_k(t)$ reduces to adaptive learning models because all players are adaptive level-0 players. We describe two possible specifications for $a_i^0(t)$ below.

- Ho et al. (1998) propose a model of adaptive level-0 players for p-beauty contests. In the model, $a_i^0(t)$ is the weighted sum of *ex post* best responses from previous rounds. Similar imitation learning models, which borrow ideas from evolutionary dynamics, have been applied to two- and three-player Cournot quantity games, in which varying information is given to players (Hück, Normann, and Oechssler, 1999; Bosch-Domenech and Vriend, 2003; Stahl 2000). Imitation learning has been shown to work well in low-information learning environments.
- Another candidate model for adaptive level-0 players is EWA learning. The EWA learning model (Camerer and Ho, 1999; Ho et al., 2007) unifies two separate classes

of adaptive learning models: 1) reinforcement learning (Erev and Roth, 1998) and 2) belief learning (Fudenberg and Levine, 1998) (including weighted fictitious play and Cournot best-response dynamics). The weighted fictitious play model has been used to explain learning behavior in games (Brown, 1951; Ellison and Fudenberg, 2000; Fudenberg and Kreps, 1993; Fudenberg and Levine, 1993).

Note that the adaptive learning models described above do not allow for higher level thinkers, who perform iterative best responses to level-0 learning dynamics. As a result, $L_k(t)$ is more general and should describe learning behavior better than any adaptive learning model.

3.4.2 Bayesian Level-k with Stationary Level-0 Players (BL_k)

Next, we describe a special case with dynamic rule levels over a static rule hierarchy. By setting $a_i^0(t) = a_i^0(1),\ \forall i, t,\ BL_k(t)$ reduces to BL_k. In BL_k, players update their beliefs and sophisticatedly learn about opponents' rule levels, while keeping rule 0's choices fixed over time. As a result, this class of models only captures sophisticated learning.

Ho and Su (2013) adopt this approach and generalize the standard level-k model to allow players to dynamically adjust their rule levels. Their model is shown to explain dynamic behavior in centipede and sequential bargaining games quite well. Nonetheless, the model works only for two-player games and for games where the strategy space and rule hierarchy have a one-to-one mapping (note that this one-to-one mapping does not exist for games with a continuous strategy space, including the p-beauty contest and Cournot oligopoly games). BL_k does not have these limitations.

3.4.3 Static Level-k Models (L_k)

The standard level-k model, in which players use static rule levels over a static rule hierarchy, has been used to explain non-equilibrium behavior in applications such as auctions (Crawford and Iriberri, 2007b), matrix games (Costa-Gomes et al., 2001), and signaling games (Brown et al., 2012). In our model, if $a_i^0(t) = a_i^0(1),\ \forall i, t,\ N_i(0) = \infty, \forall i$ and $\rho = 1$, then $BL_k(t)$ reduces to L_k, where both the level-0 players' choices and higher level thinkers' beliefs are fixed over time. As a result, our model will naturally capture the empirical regularities in the above games.

There is a subtle difference between L_k and the standard level-k model. The standard level-k model posits that each player is endowed with a rule at the beginning of the game and the player *always* plays that rule over time. Thus, any choice that does not correspond to her prescribed rule is attributed to error. Our L_k model prescribes that each player is endowed with a static *belief distribution* of opponents' levels and players sample independently in each round from this static distribution. Since a player may sample a different rule each time, the best response may change. As a consequence, a player's choice dynamics over time are interpreted as rational best responses to different sampled rules for opponents' levels, instead of choice errors. Note that if each player i's initial prior is restricted to a point mass, L_k is identical to the standard level-k model. Moreover, L_k also differs from the cognitive hierarchy model (Camerer et al., 2004). In the cognitive hierarchy model, each rule is defined as the best response to a distribution of lower rules. In L_k, however, each rule is the best response of the rule one level below.

Note that L_k further reduces to the iterative dominance solution in dominance-solvable games if we restrict $\alpha = 0$, $B_i^{\infty}(0) = 1$ and $B_i^k(0) = 0$, $\forall k < \infty$, $\forall i$ (i.e., all players have a unit mass of belief on rule ∞, which corresponds to the iterative dominance solution in dominance solvable games). Since $BL_k(t)$ nests the iterative dominance solution as a special case, we can use our model to empirically evaluate the fit of the iterative dominance solution.

The proposed Bayesian level-k model is quite general; it applies to any n-person game, irrespective of its symmetry and uniqueness of best responses. In our empirical estimation below, we study three symmetric games: p-beauty contest game, Cournot quantity game and private-value auction. Since all games are symmetric, we assume that the prior belief distribution, the density function used in belief updating, and the rule-0 action are common across all players (i.e., $\mathbf{B}_i(0) = \mathbf{B}(0)$, $f_i(\cdot\,|\,\cdot) = f(\cdot\,|\,\cdot)$, and $a_i^0(t) = a^0(t)$, $\forall i$).

4 p-Beauty Contest Game

In this section, we estimate the Bayesian level-k model and its special cases using data on p-beauty contest games. Below, we describe our experimental data, estimation methods, and results.

4.1 Data

We use experimental data from p-beauty contest games collected by Ho et al. (1998). We focused on games where players chose numbers from $[0, 100]$ with $p = 0.7$ or 0.9 and $n = 3$ or 7. The prize in each round was \$1.50 for groups of size 3 and \$3.50 for groups of size 7, keeping the average prize at \$0.50 per player per round. A total of 277 subjects participated in the experiment. Each subject played the same game for 10 rounds in a fixed matching protocol.[5] After each round, only the group average and the target number were publicly announced. For $p = 0.7$, there were 14 groups of size 7 and 14 groups of size 3; for $p = 0.9$, there were 14 groups of size 7 and 13 groups of size 3. Thus, we have a total of 55 groups and 2,770 observations. The first and last round of data and the convergence pattern of the data are presented in Figure 1.

4.2 Estimation Methods

Similar to Ho et al. (1998), we normalized the data with each round's sample mean and standard deviation so that each round's data has equal influence on the estimation.[6] Since players were told target numbers, it was natural to model the level-0 player as one who simply mimics past periods' target numbers. Note that these target numbers may or may not be chosen by a player and they simply represent the ideal choices for winning a prize *ex post*. Ignoring one's own influence on the average, target numbers are also the *ex post* best responses.

Let $\omega(t)$ be the target number in time t. We define the rule 0's choice at time t, $a^0(t)$, to be the weighted sum of her choice at time $t-1$, $a^0(t-1)$, and the most recently observed

[5]Reputation building is unlikely because the p-beauty contest game is a constant-sum game where players' interests are strictly opposed.

[6]See Ho et al. (1998) for the exact procedure for normalization.

target number at time $t-1$, $\omega(t-1)$. Formally, $a^0(t)$ is defined as:

$$a^0(t) \;=\; \delta \cdot a^0(t-1) + (1-\delta) \cdot \omega(t-1) \quad (t \geq 2) \tag{4.1}$$

where δ captures the rate at which a level-0 player decays past choices. To initialize the above updating equation, we set $a^0(1) = \omega(0)$, a parameter that we will estimate from the data. Note that when $\delta = 0$, level-0 players decay history completely after one round and simply repeat the most recently observed target number. Hence, level-0 players capture Cournot best-response dynamics in this case. On the other hand, when $\delta = 1$, level-0 players never incorporate observed game plays and their choices remain unchanged, i.e., $a^0(t) = a^0(1) = \omega(0),\ \forall t$.

To initialize the belief updating process, the prior belief distribution was assumed to be a simple one-parameter Poisson distribution with mean τ. Note that in Ho et al. (1998), only the opponents' average choice is provided as an aggregate feedback after each round. Thus, as described in Section 3.3, each player i infers the symmetric $x_j(t)$ as the same as the opponents' average choice for all $j \neq i$ and updates her belief based on this information. In the belief updating process, as in Camerer and Ho (1999), we assumed $N(t-1) \leq N(t)$, which is equivalent to $N(0) \leq (n-1)/(1-\rho)$, so that experience always accumulates.

The error structure of a player's choice x around her predicted choice μ was assumed to follow a normal distribution with mean at the predicted choice μ and standard deviation σ, i.e., $f(x \mid \mu) = \mathcal{N}(x \mid \mu, \sigma)$, where $\mathcal{N}(x \mid \mu, \sigma)$ denotes the normal density with mean μ and standard deviation σ evaluated at x. In order to better capture heterogeneity among players, we allowed a separate variation, σ_0, for level-0 players, i.e., $f(x \mid \mu) = \mathcal{N}(x \mid \mu, \sigma_0)$. These standard deviations, σ and σ_0, were empirically estimated.

We used the maximum likelihood estimation method with the MATLAB optimization toolbox. To ensure that the maximized likelihood is indeed the global maximum, we utilized multiple starting points in the optimization routine. The standard errors of parameter estimates were estimated using the jackknife method.

4.3 Likelihood Function

Each player i's probability of choosing $x_i(t)$ at time t is based on her actual interactions with her matched opponents up to time $t-1$. Let $\Theta = (\alpha, \omega(0), \delta, \sigma_0, \tau, N(0), \rho, \sigma)$ and let $H(t)$ be the history of actions chosen by *all* players up to time t, i.e., $H(t) = \{x_i(t') \,|\, \forall t' \leq t \text{ and } \forall i\}$.

Further, let $L_i^0(x_i(t) \,|\, \Theta,\, H(t-1))$ be the probability that a *level-0* player i chooses $x_i(t)$ at time t, and $L_i^h(x_i(t) \,|\, \Theta,\, H(t-1))$ be the probability that a *higher level* player i chooses $x_i(t)$ at time t, given Θ and $H(t-1)$. Since each player i's type (i.e., whether level 0 or higher level) is fixed throughout the game, the likelihood of observing player i's entire data, $L_i(x_i(1), \ldots, x_i(T) \,|\, \Theta, H(T-1))$, is given by:[7]

$$
\begin{aligned}
&L_i(x_i(1), \ldots, x_i(T) \,|\, \Theta, H(T-1)) \\
&\quad= \alpha \cdot \prod_{t=1}^{T} L_i^0(x_i(t) \,|\, \Theta, H(t-1)) + (1-\alpha) \cdot \prod_{t=1}^{T} L_i^h(x_i(t) \,|\, \Theta, H(t-1)) \\
&\quad= \alpha \cdot \prod_{t=1}^{T} \mathcal{N}(x_i(t), a^0(t), \sigma_0 \,|\, \Theta, H(t-1)) \\
&\qquad + (1-\alpha) \cdot \prod_{t=1}^{T} \left(\sum_{k=0}^{\infty} B_i^k(t-1 \,|\, \Theta, H(t-1)) \cdot \mathcal{N}(x_i(t), a^{k+1}(t), \sigma \,|\, \Theta, H(t-1)) \right)
\end{aligned}
$$

That is, player i's likelihood function is the weighted sum of conditional likelihoods depending on whether the player is level 0 (i.e., with probability α) or higher (i.e., with probability $1-\alpha$). The log-likelihood of observing the entire data, $LL(H(T) \,|\, \Theta)$, is then given by:

$$
LL(H(T) \,|\, \Theta) \;=\; \sum_{i=1}^{n} \log\left(L_i(x_i(1), \ldots, x_i(T) \,|\, \Theta, H(T-1))\right)
$$

In sum, the parametric space of the estimation is $\Theta = (\alpha, w(0), \delta, \sigma_0, \tau, N(0), \rho, \sigma)$: 1). For level-0 players, we estimated the probability that a player is only a level-0, adaptive-learning player throughout the game (α), the initial conditions ($\omega(0)$), the decay rate of past periods' ex-post best responses (δ) and the error term for adaptive level-0 players (σ_0); 2). For higher level thinkers, we estimated the prior Poisson belief distribution (τ), the strength of the initial prior ($N(0)$), the memory decay factor (ρ) and the error term for sophisticated, belief-updating players (σ).

[7] Note that we empirically assume that players may make errors when they best respond, and that their actual choices are normally distributed around the predicted best-responses. For the sake of parsimony, the standard deviation of these errors is empirically assumed to be σ, the same as that of $f(\cdot \,|\, \cdot)$. The errors could be assumed to be different in general, i.e., what a player believes the error rates of opponents' choices to be, may be different from their actual error rates.

4.4 Estimated Models

We estimated a total of four models, which can be classified into two groups. The first group includes two models that incorporate adaptive level-0 players (i.e., $a^0(t)$) while the second group includes two models that do not incorporate adaptive level-0 players (i.e., $a^0(t) = a^0(1),\ \forall t$).

The first group includes the following 2 models:

1. Bayesian level-k model with adaptive level-0 ($BL_k(t)$): This is the full model with the parametric space $(\alpha, \omega(0), \delta, \sigma_0, \tau, N(0), \rho, \sigma)$.
2. Static level-k model with adaptive level-0 ($L_k(t)$): This model allows level-0 players to adaptively learn and $a^0(t)$ to change over time. However, the rule levels of higher level thinkers remain fixed over time. This model is a nested case of $BL_k(t)$ with $N(0) = \infty$ and $\rho = 1$, i.e., the parametric space is $(\alpha, \omega(0), \delta, \sigma_0, \tau, \infty, 1, \sigma)$.

The second group includes the following 2 models:

3. Bayesian level-k model with stationary level-0 (BL_k): This model ignores level-0's adaptive learning and only captures sophisticated learning. Specifically, it allows higher level thinkers to update their beliefs after each round about what rules others are likely to choose. As a result, these higher level thinkers can choose a rule with different probabilities over time. This model is obtained by setting $\delta = 1$ in $BL_k(t)$, i.e., $a^0(t) = a^0(1), \forall t$. The resulting parametric space is $(\alpha, \omega(0), 1, \sigma_0, \tau, N(0), \rho, \sigma)$.
4. Static level-k with stationary level-0 (L_k): This captures neither adaptive learning nor sophisticated learning. It is obtained by setting $N(0) = \infty$ and $\rho = \delta = 1$. The resulting parametric space is $(\alpha, \omega(0), 1, \sigma_0, \tau, \infty, 1, \sigma)$. This model is similar to the standard level-k model (as discussed in Section 3).

The full model $BL_k(t)$ nests the other 3 models as special cases. By comparing the fit of our model with those of nested cases, we can determine the relative importance of adaptive and sophisticated learning.

4.5 Estimation Results

Table 1 shows the estimation results. The first column lists the names of parameters and test statistics. The next four columns report results from each of the four estimated models. The top panel reports parameter estimates and the bottom panel reports maximized log-likelihood, χ^2, the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and the two-fold cross validation result.

The bottom panel of Table 1 shows that $BL_k(t)$ fits the data better than the three special cases. In fact, the test statistics χ^2 and AIC indicate that the three special cases are rejected in favor of $BL_k(t)$. For instance, the χ^2 test statistics of L_k, BL_k, and $L_k(t)$ against $BL_k(t)$ are 11.04, 846.46 and 1,630.06, respectively. These results suggest that having adaptive level-0 players (i.e., adaptive learning) and allowing Bayesian updating by higher level thinkers (i.e., sophisticated learning) can help to explain subjects' dynamic choices over time.[8]

In order to check the out-of-sample validity of the model, we ran the two-fold cross validation. Specifically, we randomly split the data into two groups, calibrated the model using one group of data and validated using the other; we repeated this process 50 times and reported the average likelihood value obtained from these repetitions. On each run, the same random split was used across the four models (the full model and three nested models) for the cross validation in order to control for the sampling error. As shown, we see that the results from the likelihood on the whole sample are mostly preserved in the two-fold cross validation except that the likelihood difference between the full model and the second best fitting adaptive learning model, $L_k(t)$, is halved since each fold consists half the data. The χ^2 statistics of $L_k(t)$, BL_k and L_k against the full model in the two-fold cross validation has a p-value of 0.13, 0.00 and 0.00 respectively.

[INSERT TABLE 1 HERE.]

The estimated parameters of the best-fit model, $BL_k(t)$, appear reasonable. The fraction

[8] We also estimate the self-tuning EWA model (Ho et al., 2007). Since the EWA model uses a likelihood function over discrete choices, we discretize our probability density function so that we can compare the two log-likelihoods. We obtain LL = -11038.77, AIC = 22093.54, and BIC = 22140.95 for the $BL_k(t)$ model, and LL = -11733.83, AIC = 23469.66, and BIC = 23475.59 for the self-tuning EWA model. These results show that sophisticated learning is important in capturing the choice dynamics in the data.

of adaptive level-0 players ($\hat{\alpha}$) is estimated to be 47%. These adaptive level-0 players weigh the most recently observed ex-post best response by 0.89 (i.e., $1-\hat{\delta}=0.89$), suggesting that players notice that ex-post best responses are converging towards equilibrium. The level-0 players' choices have a standard deviation of $\hat{\sigma_0}=1.32$, and their choices are more variable than those of higher level thinkers ($\hat{\sigma}=0.46$). Since we normalized the choice, as in Ho et al (1998), by sample standard deviation (described in Section 4.2), the σ value of 1 means that higher-level players' choices are as variable as those made by the entire population observed in the data. These higher level thinkers ($k \geq 1$) give their initial priors the weight $\hat{N}(0)=2.00$ and decay this prior weight at a rate of $\hat{\rho}=0.00$. That is, the influence of the initial prior immediately approaches near-zero after the first round. Comparing $\hat{N}(0)$ and $\hat{\rho}$ between the $BL_k(t)$ and $L_k(t)$ models, we note that our models are not very sensitive to these two parameters.

We note the following patterns in the parameter estimates:

1. The estimated fraction of level-0 players does not vary much across the four models, ranging from 0.32 to 0.48.
2. Across all models, the choices of level-0 players are more variable than those of higher level players ($\hat{\sigma_0}$ ranges from 1.31 to 10.00, while $\hat{\sigma}$ ranges from 0.43 to 0.46).
3. Introducing adaptive learning decreases the average of the initial belief distribution over opponents' rule levels of higher level players ($\hat{\tau}$ is 0.00 and 0.07 compared to 3.84 and 18.06).

Figure 2 presents a three-dimensional bar graph of actual choice frequency of the experimental data. To study the prediction power of the model for each individual choice, we look at the entire distribution of the absolute deviation of the model prediction from each data point. That is, for each player's choice in each round, we calculate how distant the actual choice is from the expected model prediction given the observed history of game play. The three-dimensional bar graph in Figure 3 plots the distribution of these absolute deviations. (Note that the "Choices" axis is reversed compared to Figure 2 so taller bars do not block shorter ones.) We find that the deviation is as large as 60 in the first round, but most deviations are below 10 from the 5th round. We also compute the relative magnitude of the absolute deviation compared to the actual data. The relative absolute deviation (absolute deviation divided by the the actual data value) is on average 59% for the full model, while for the nested models, ($L_k(t)$, BL_k and L_k) it is 62%, 317% and 122%, respectively. These statistics, consistent with those shown in Table 1, demonstrates that the full model predicts better than the nested models.

[INSERT FIGURES 2 AND 3 HERE.]

We further demonstrate how players in the data appear to adjust their rule levels depending on the feedback they receive. If players change their rules in subsequent rounds to beat their opponents by taking into account the opponents' observed rules, players' rules must change towards opponents' observed rules. Specifically, let OWNDIFF denote a player's own rule in time $t+1$ minus her own rule in time t, and OPPDIFF denote the rule opponents choose in time t minus her own rule in time t. Then, OWNDIFF and OPPDIFF must be positively correlated if players are changing rules in response to the feedback they receive. Figure 4 plots the pairwise correlation of the OWNDIFF and OPPDIFF values of all players. It is clearly shown that the two values are positively correlated with the correlation coefficient of 0.45.

[INSERT FIGURE 4 HERE.]

5 Cournot Quantity Game

In this section, we analyze another classical game, the Cournot oligopoly game, using the Bayesian level-k model. In the classical symmetric Cournot game, n players compete in quantities and simultaneously choose what quantities of a homogeneous product to produce. The cost structure of players is common knowledge and is the same for all players. The market price is determined by the total quantity produced by all players, which is normally assumed to be a linear function of the total quantity produced. The payoff to each player is determined by the unit margin (market price less the variable cost) times the quantity produced by the player. This game is known to have a unique interior symmetric equilibrium characterized by the intersection of each player's best response function given the other players' chosen quantities.

We find that the classical Cournot game is interesting for applying our Bayesian level-k model for three reasons. First, the class of Cournot games bears resemblance to the class of p-beauty contest games. That is, it is dominance solvable with a clearly defined rule hierarchy where the rules converge to the unique equilibrium. Second, despite these similarities, the Cournot quantity game has an interior equilibrium solution while the p-beauty contest game has a boundary equilibrium solution. This implies that the rule levels alternate between choices greater and smaller than the equilibrium until they finally converge to the interior equilibrium point as rules increase. This characteristic may help disentangle adaptive learning from sophisticated learning. Third, from the practical application standpoint, Cournot games resemble a commodities market in which firms compete in quantity. Thus, examining the Cournot game will broaden our perspective on how different types of learning occur in these different types of games.

5.1 Data

We use experimental data on classical Cournot oligopoly quantity games, collected by Fouraker and Siegel (1963). Fouraker and Seigel ran experiments on the Cournot quantity game with group size two or three.[9] Since players were told their own as well as the total sum of opponents' choices in all conditions, we fit our model to the entire dataset. All games followed a fixed matching protocol. There were a total of 22 rounds. The price and profit function of the Cournot quantity game was given by: $p = A - B \cdot \sum_{i=1}^{n} q_i$, where $A = 2.4$, $B = 0.04$, p is the market price and q_i is the quantity chosen by player i. Player i's profit in each round is defined by: $\pi_i = p \cdot q_i = \left(A - B \cdot \sum_{j=1}^{n} q_j\right) \cdot q_i$. The unique interior equilibrium of duopoly and triopoly settings was 20 and 15, respectively. There were a total of 130 subjects, of which 64 were under duopoly, and 66 were under triopoly conditions. Subjects were rewarded in

[9] There were separate treatments for two information states, termed "complete" and "incomplete", with entirely different meanings from the ones normally used nowadays. In Fouraker and Siegel, subjects were told their own as well as opponents' total profits in the complete information state, while they were told only their own profits in the incomplete information state. In all states, however, players were told their opponents' choices and knew that the payoff function was symmetric. Therefore, we fit our model to the entire dataset.

the amount of the total profit obtained across all rounds times \$0.01.

Figure 5 compares the first round data to the last round data in duopoly and triopoly settings. As shown, although the number of equilibrium choices has increased from about 15% to 30% in both settings, the rest of the choices maintain a similar spread. Figure 6 plots the average and standard deviation of choices over time. The average stays quite static around the equilibrium. The standard deviation also remains relatively constant, reducing only from about 5 to 4 over the 22 periods. The standard deviation is quite low given that the strategy space spans from 0 to 60. Figure 5 and 6 together demonstrate that the pattern of data is fixed quite instantly at the onset of the game, and most choices are quite close to the equilibrium point from period 1. The data pattern in earlier rounds remains relatively unchanged over time except for a 15% increase in equilibrium choices.

5.2 Rule Hierarchy

Player i's best response that maximizes her profit function given other players' chosen quantities can be obtained as $q_i = (A - B \cdot \sum_{j \neq i} q_j)/(2B)$ and the symmetric equilibrium quantity is $\frac{A}{(n+1)B}$.

The rule hierarchy is iteratively defined from the level-0 rule. Believing that all her opponents are level 0, the level-1 player best responds by choosing $a^1(t)$ that maximizes her profit. Thus, $a^1(t) = \frac{A}{2B} - \frac{n-1}{2} \cdot a^0(t)$. Iteratively,

$$a^k(t) = \frac{A}{(n+1)B} - \frac{(n-1)}{2} \cdot \left(a^{k-1}(t) - \frac{A}{(n+1)B} \right) .$$

5.3 Estimation Results

Table 2 shows the estimation results of the Cournot quantity game. Estimation methods and assumptions were exactly the same as in the p-beauty contest game and level-0 players are assumed to be imitation learners choosing a weighted sum of *ex post* best responses. As opposed to the p-beauty contest game, the χ^2 test statistics indicate that the $BL_k(t)$ fits the data significantly better than BL_k and L_k ($\chi^2 = 68.81$ and 72.36 respectively) but it is not statistically different from $L_k(t)$ ($\chi^2 = 3.15$). This suggests that adaptive learning is the dominant type of learning and the role of sophisticated learning is negligible in the Cournot quantity game.[10]

The same results hold in the two-fold cross validation as well. The validation fit of $BL_k(t)$ is actually 5 points less than the second-best fitting $L_k(t)$ implying the possibility of over-fitting compared to the $L_k(t)$. Against the other two nested models (BL_k and L_k), $BL_k(t)$ still fits significantly better in the cross validation.

[INSERT TABLE 2 HERE.]

The fraction of adaptive level-0 players ($\hat{\alpha}$) is estimated to be 11%. These adaptive level-0 players weigh the most recently observed ex-post best response by 0.23 (i.e., $1 - \hat{\delta} = 0.77$). The level-0 players' choices have a standard deviation of $\hat{\sigma_0} = 9.25$, and their choices are more variable than those of higher level thinkers ($\hat{\sigma} = 4.02$). These higher level thinkers ($k \geq 1$) give their initial priors the weight $\hat{N}(0) = 52.73$ and decay this prior weight at a rate of $\hat{\rho} = 0.45$. That is, the influence of the initial prior drops to $52.73 \cdot 0.45^t$ after t rounds.

We note that allowing for adaptive learning reduces the fraction of level-0 players from around 0.47 to 0.11 and increases the standard deviation for their choices from around 4.34 to 9.25. The choices of level-0 players are about twice as variable as the choices of higher level players ($\hat{\sigma_0}$ ranges from 4.34 to 9.54 while $\hat{\sigma}$ ranges from 3.98 to 4.34).

6 Private-Value Auction

To demonstrate how the Bayesian level-k model explains different types of games, we further fit our model to another completely different game: the first-price, private-value auction. The private-value auction is distinct from both the p-beauty contest and Cournot oligopoly games because a player's type, i.e., her private value, is not known to opponents, who have to make best responses based on their expectations. As a consequence, studying the private-value auction will help us understand how the Bayesian level-k model behaves in games with

[10] The self-tuning EWA model yields LL = −9582.83, AIC = 19169.66, and BIC = 19173.62. The $BL_K(t)$ model fits better than the self-tuning EWA model because it has a more flexible error structure (spreading probability mass more evenly across choices). This "more forgiving" error structure is important in the Cournot game where choices remain spread out in the final rounds.

incomplete information.

6.1 Data

We utilize the existing experimental first-price, private-value auction data from Goeree et al. (2002). In their experiments, each subject was randomly matched to another in each round. Each bidder received a private value (v) for a prize in a first-price, sealed-bid auction. The prize went to the highest bidder and each bidder's private value was exogenously determined by a random device (rolling a die), among \$0, \$2, \$4, \$6, \$8, and \$11 for the low-values treatment condition, and among \$0, \$3, \$5, \$7, \$9, and \$12 for the high-values treatment condition. Upon being randomly assigned their private values, subjects placed a sealed bid. Bids were constrained to integer dollar amounts and ties were decided with a coin flip. Profits were realized only when an auction was won and the net surplus (private value minus the winning bid) was given. The equilibrium prediction for both treatment conditions is \$0, \$1, \$2, \$3, \$4, and \$5 for the six private values in both treatment conditions.

Eight sessions were conducted in total, among which four sessions were under the low-value treatment conditions and four other sessions were under the high-value treatment conditions. Each session ran for 15 rounds and had 10 subjects who were randomly matched to each other in each round. After each round, subjects were privately told their own earnings and their opponent's bid of that round. At the end of the session, subjects were paid one-half of the total profits accumulated across all rounds. While Goeree et al. (2002) fit data from rounds 6 to 15 after data has stabilized, we fit the entire 15 rounds of data to fully examine the learning dynamics.

Figure 7 shows average bids by period (Figure 2 of Goeree et al. (2002)). The data is erratic, particularly in the earlier rounds and even in the later rounds except for some lowest private values. Unlike the previous two games, there is not much indication of convergence in the private-value auction. Persistent overbidding beyond the equilibrium prediction (i.e., \$0, \$1, \$2, \$3, \$4, and \$5 for six values) is observed.

6.2 Estimation Method

We first define the rule hierarchy in the private-value auction. Since rule hierarchy depends on one's private value, we define each rule level conditional on one's private value. We denote the choice of rule k by $a^k(t\,|\,v)$, with a private value of v at time t. Then, $a^k(t\,|\,v)$ is the best-response to the rule-$(k-1)$ opponent who has an equal probability of having one of the 6 possible private values and is thus defined as:

$$\begin{aligned} a^k(t\,|\,v) &= \arg\max_{s \in S}\ \mathrm{E}_{v'}[\pi(s, a^{k-1}(t\,|\,v'))] \\ &= \arg\max_{s \in S}\ \frac{1}{6} \sum_{v'} \pi(s, a^{k-1}(t\,|\,v')) \end{aligned}$$

Note again that if multiple best responses exist, $a^k(t\,|\,v)$ is defined as a mixed strategy placing equal probability weights over these multiple best responses.

Next, in order to define the error structure of the model, i.e., $f(\cdot)$, we note that the first-price, sealed-bid auction prescribes that one shall place a bid lower than one's value. This is because a bid lower than one's value (except when the value is zero) will yield a positive expected payoff whereas the expected payoff from bidding one's value is zero and that from bidding higher than one's value is negative. This indeed manifested in the dataset. Subjects who were randomly assigned the private value of zero bid zero 96% of the time. Subjects who were randomly assigned a private value higher than zero bid a number strictly smaller than their private value 99% of the time. Therefore, we concluded that $f(\cdot)$ must capture these characteristics of the data and assumed $f(x\,|\,a^k(t\,|\,v)) \propto \exp(-|x - a^k(t\,|\,v)|)$ only if $x < v$ for $v > 0$; and only if $x = v = 0$ for $v = 0$. Hence, $f(x\,|\,a^k(t\,|\,v)) = 0$ if $x > v$ always holds. Note that the farther x is from $a^k(t\,|\,v)$, the greater $f(x\,|\,a^k(t\,|\,v))$ is, i.e., this error structure is of similar shape to a normal distribution.

6.3 Estimation Results

Table 3 shows the estimation results of the private-value auction game. In this game, we find that the χ^2 test statistics indicate that $BL_k(t)$ fits the data significantly better than $L_k(t)$ and L_k ($\chi^2 = 31.10$ for both cases) but is not statistically different from BL_k ($\chi^2 = 0.00$).[11]

[11]The $BL_k(t)$ model also fits the private-value auction better than the self-tuning EWA model. The self-tuning EWA model yields LL = -2912.93, AIC = 5827.87, and BIC = 5832.96. The self-tuning EWA model

This suggests that sophisticated learning is the dominant type of learning in the private-value auction, as opposed to the Cournot oligopoly game where adaptive learning was dominant. The same conclusion holds for the two-fold cross validation result too.

[INSERT TABLE 3 HERE.]

The fraction of sophisticated players is estimated to be 87% since $\hat{\alpha} = 0.13$. These sophisticated players weigh their initial priors by $\hat{N}(0) = 0.11$ and decay this prior weight at a rate of $\hat{\rho} = 0.99$. That is, the influence of the initial prior drops to $0.11 \cdot 0.99^t$ after t rounds.

Incorporating sophisticated learning reduces the prior belief $\hat{\tau}$ of higher-level players from 0.13 to 0.02, and also reduces the standard deviation of their choices from 0.62 and 1.34 to around 0.58. However, the estimated parameters for adaptive level-0 players do not vary appreciably. Across all four models, the fraction of adaptive players ($\hat{\alpha}$) is estimated to be 13%. Their initial choices hardly decay at all (since $\hat{\delta} = 1.00$), and the standard deviation of their choices $\hat{\sigma}_0 = 0.86$ is higher than the corresponding estimate for sophisticated players in three out of the four models.

Note that the fraction of level-0 players is estimated to be similar to that in the Cournot oligopoly games, where $\hat{\alpha} = 0.11$ but for quite different reasons. In the Cournot oligopoly games, α is estimated to be small because the convergence to equilibrium occurs instantly and the level-0 rule becomes indistinguishable from higher-level rules when it is very close to equilibrium. On the contrary, in the private-value auction, the noise in the data persists in all rounds with no sign of convergence (see Fig 2 of Goeree et al., 2002) and the estimated level-0 rule remains distinct from higher-level rules. In this case, the fraction of level-0 players is estimated to be small because adaptive learning (represented by level-0 players) cannot explain the choice dynamics and variation that persists in the data.

In sum, in a private-value auction, which is an incomplete information game, the adaptive

performs worse in the auction because its choice attraction does not have a natural way to take players' private values into account in the updating process. In contrast, the L_k model captures a player's private value via its higher rule level players. The choices of higher rule players are determined as the iterated best responses based on their private values. As a consequence, the level-k thinking aspect of the $BL_k(t)$ model is crucial in capturing choice dynamics in this game.

learning-only model (i.e., where players form best responses based solely on the observed action, without inferring opponents' rules) is strictly worse than the sophisticated learning-only model. This result is quite intuitive because in sophisticated learning, players infer opponents' rules and update their beliefs accordingly. This process necessarily requires players to deduce the conditional probability of a choice occurring given a player's type (i.e., private-value). As a consequence, sophisticated learning can generate more accurate predictions than adaptive learning.

7 Conclusion

The level-k model is an appealing structural model for explaining non-equilibrium behavior because it is based on an intuitive, iterative reasoning process. Each iteration in the level-k model is akin to a thinking step. Players start at level 0 and iteratively advance one step by best-responding to their counterparts who reason one step less. Nonequilibrium behavior emerges when players do not perform enough thinking steps. While the level-k model has been successfully applied to both laboratory and field data, it is inherently static and has not been designed to capture choice dynamics. A wide body of empirical evidence shows that choice behavior is inherently dynamic and converges to equilibrium over time.

To capture dynamic choice behavior, this paper develops a structural Bayesian level-k model that distinguishes between two possible types of learning dynamics: adaptive and sophisticated learning. The model generalizes the standard level-k model in two significant ways:

1. We allow the level-0 rule to adaptively respond to historical game plays. To the extent that observed game plays move towards equilibrium, players may approach equilibrium without advancing in rule levels. In this way, players learn adaptively.
2. We allow players to change their rule levels by Bayesian updating, where players update beliefs about opponents' rule levels based on observations. Thus, players perform more thinking steps when they anticipate similar behavior from opponents. In this way, players exhibit sophisticated learning.

Interestingly, the extended model provides a novel unification of two separate streams of nonequilibrium models: level-k and adaptive learning models. On one hand, level-k models are "static" and have been used to predict behavior in one-shot games. On the other hand, adaptive learning models have been used to describe choice dynamics in repeated games. While these two seemingly distinct classes of models have been treated separately in the literature, our model brings them together and provides a sensible way to model behavior in both one-shot and repeated games in a common, tractable framework.

We apply the Bayesian level-k model to the existing experimental data on three games that much resemble practical settings: the p-beauty contest (Ho et al. 1999), the Cournot quantity game (Fouraker and Siegel 1963) and the private-value auction (Goeree et al. 2002). We find that our generalized model explains the data well for all three classes of games. In contrast, nested models that capture either adaptive or sophisticated learning explains behavior in some games but not others.

Our structural estimates allow us to identify whether adaptive learning or sophisticated learning is dominant in each class of games. We find that although subjects' dynamic behavior is readily observable in the experimental data for all games, the underlying learning dynamics are quite different across them: in p-beauty contest games, players' dynamic behavior is driven by both types of learning, whereas in Cournot quantity games it is driven by adaptive learning alone, and in private-value auction games it is driven by sophisticated learning alone. We attribute such differences to the convergence pattern (instant vs. sluggish) and information setting (complete vs. incomplete). Hence, we show that a more restrictive model that allows only one type of learning may be mis-specified, and would not explain behavior well across different classes of games.

In summary, by incorporating both adaptive and sophisticated learning, our model may be viewed as an empirical characterization of the equilibration process. This view bears a similar spirit to Harsanyi's "tracing procedure" (Harsanyi and Selten, 1988), in which players' successive choices, as they react to new information in strategic environments, trace a path towards the eventual equilibrium outcome.

References

Alaoui, L. and A. Penta. 2016. "Endogenous Depth of Reasoning." Rev. of Econ. Stud. 83: 1297-1333.

Binmore, K., A. Shaked and J. Sutton. 1985. "Testing Noncooperative Bargaining Theory: A Preliminary Study." Am. Econ. Rev. 75: 1178-1180.

Bosch-Doménech, A., J. G. Montalvo, R. Nagel and A. Satorra. 2002. "One, Two, (Three), Infinity, ... : Newspaper and Lab Beauty-Contest Experiments." Am. Econ. Rev. 92(5): 1687-1701.

Bosch-Doénech, A. and N. J. Vriend. 2003. "Imitation of Successful Behaviour in the Markets." The Economic Journal 113: 495-524.

Brocas, I., J. D. Carrillo, S. W. Wang, and C. F. Camerer. 2014. "Imperfect Choice or Imperfect Attention? Understanding Strategic Thinking in Private Information Games." The Review of Economic Studies 81(3) 944-970.

Brown, A., C. F. Camerer and D. Lovallo. 2012. "To Review or Not To Review? Limited Strategic Thinking at the Movie Box Office." American Economic Journal: Microeconomics 4(2): 1-28.

Brown, George W. 1951. "Iterative solution of games by fictitious play." Activity Analysis of Production and Allocation 13(1): 374-376.

Capra, C. M., J. K. Goeree, R. Gomez and C. A. Holt. 1999. "Anomalous Behavior in a Traveler's Dilemma?" Am. Econ. Rev. 89(3) 678-690.

Camerer, C. and T-H. Ho. 1999. "Experience-Weighted Attraction Learning in Normal Form Games." Econometrica 67: 837-874.

Camerer, C. Ho, T-H. and Chong, J-K. 2002. "Sophisticated Learning and Strategic Teaching." J. Econ. Theory 104: 137-188.

Camerer, C. F., T-H. Ho and J-K. Chong. 2004. "A Cognitive Hierarchy Model of Games." Quarterly Journal of Economics 119(3): 861-898.

Costa-Gomes, M., V. Crawford and B. Broseta. 2001. "Cognition and Behavior in Normal-Form Games: An Experimental Study." Econometrica 69: 1193-1235.

Costa-Gomes, M. and V. Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." Am. Econ. Rev. 96: 1737-1768.

Costa-Gomes, M. and G. Weizsäcker. 2008. "Stated Beliefs and Play in Normal-Form Games." Review of Economic Studies 75(3): 729-762.

Costa-Gomes, M., V. Crawford and N. Iriberri. 2009. "Comparing Models of Strategic Thinking in Van Huyck, Battalio, and Beil's Coordination Games." J. Eur. Econ. Assoc. 7: 377-387.

Crawford, V. P. 2003. "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions." Am. Econ. Rev. 93: 133-149.

Crawford, V. P. and N. Iriberri. 2007a. "Fatal Attraction: Salience, Naivete, and Sophistication in Experimental Hide-and-Seek Games." Am. Econ. Rev. 97: 1731-1750.

Crawford, V. P. and N. Iriberri. 2007b. "Level-k Auctions: Can a Non-Equilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?." Econometrica 75: 1721-1770.

Crawford, V. P., T. Kugler, Z. Neeman and A. Pauzner. 2009. "Behaviorally Optimal Auction Design: An Example and Some Observations." J. Eur. Econ. Assoc. 7(2-3): 377-387.

Crawford, V. P., M. A. Costa-Gomes and N. Iriberri. 2013. "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications." J. Econ. Lit. 51(1): 5-62.

Ellison, G. and D. Fudenberg. 2000. "Learning Purified Mixed Equilibria." J. Econ. Theory 90(1): 84-115.

Erev, I. and A. E. Roth. 1998. "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria." Am. Econ. Rev. 88(4): 848-881.

Fouraker, L. and S. Siegel. 1963. "Bargaining Behavior" New York: McGraw-Hill.

Fudenberg, D. and D. Kreps. 1993. "Learning mixed equilibria." Games and Economic Behavior 5: 320-367

Fudenberg, D. and D. Levine. 1993. "Steady State Learning and Nash Equilibrium." Econometrica 61(3): 547-573.

Fudenberg, D. and D. Levine. 1998. "Theory of Learning in Games." Cambridge, MIT Press.

Gill, D. and V. L. Prowse. 2016. “Cognitive Ability, Character Skills, and Learning to Play Equilibrium: A Level-k Analysis.” J. Polit. Econ. 124(6): 1619-1676.

Goldfarb, A. and B. Yang. 2009. “Are All Managers Created Equal?.” J. Marketing Res. 46(5): 612-622.

Goldfarb, A. and M. Xiao. 2011. “Who Thinks About the Competition? Managerial Ability and Strategic Entry in US Local Telephone Markets.” Am. Econ. Rev. 101(7): 3130-3161.

Goeree, J. K. and C. A. Holt. 2001. “Ten Little Treasures of Game Theory and Ten Intuitive Contradictions.” Am. Econ. Rev. 91(5):1402-1422.

Goeree, J. K., C. A. Holt and T. R. Palfrey 2002. “Quantal Response Equilibrium and Overbidding in Private-Value Auctions.” Journal of Economic Theory. 104(1): 247-272.

Harsanyi, J. C. and R. Selten. 1988. “A General Theory of Equilibrium Selection in Games.” Cambridge, MIT Press.

Ho, T-H., C. Camerer and K. Weigelt. 1998. “Iterated Dominance and Iterated Best Response in Experimental p-Beauty Contests.” Am. Econ. Rev. 88: 947-969.

Ho, T-H., C. Camerer, C. and J-K. Chong. 2007. “Self-tuning Experience-Weighted Attraction Learning in Games.” J. Econ. Theory 133(1): 177-198.

Ho, T-H. and X. Su. 2013. “A Dynamic Level-k Model in Sequential Games.” Manage. Sci. 59(2): 452-469.

Hück, S., H. Normann and J. Oechssler. 1999. “Learning in Cournot Oligopoly–An Experiment.” Economic Journal 109: C80-C95.

McAfee, R. P., literature, J. M. J. O. E., 1987. “Auctions and bidding.” Journal of Economic Literature 25(2): 699-738.

McKelvey, R. D. and T. R. Palfrey. 1992. “An Experimental Study of the Centipede Game.” Econometrica 60: 803-836.

Nagel, R. 1995. “Unraveling in Guessing Games: A Experimental Study.” Am. Econ. Rev. 85(5): 1313-26.

Ochs, J. and A. E. Roth. 1989. “An Experimental Study of Sequential Bargaining.” Am. Econ. Rev. 79: 355-384.

Östling, R., J. T-Y. Wang, E. Chou and C. Camerer. 2011. "Testing Game Theory in the Field: Swedish LUPI Lottery Games." American Economic Journal: Microeconomics 3(3): 1-33.

Stahl, D. O. 1996. "Boundedly rational rule learning in a guessing game." Games and Economic Behavior 16(2), 303-330.

Stahl, D. O. 1999. "Evidence based rules and learning in symmetric normal-form games." International Journal of Game Theory 28(1), 111-130.

Stahl, D. O. 2000. "Rule learning in symmetric normal-form games: theory and evidence." Games and Economic Behavior 32(1), 105-138.

Stahl, D. O. 2001. "Population rule learning in symmetric normal-form games: theory and evidence." Journal of Economic Behavior Organization, 45(1), 19-35.

Stahl, D. O. and E. Haruvy. 2002. "Aspiration-based and reciprocity-based rules in learning dynamics for symmetric normal-form games." Journal of Mathematical Psychology 46(5), 531- 553.

Stahl, D. and P. Wilson. 1994. "Experimental Evidence on Players' Models of Other Players." J. Econ. Behav. Organ. 25 (3): 309-327.

Stahl, D. and P. Wilson. 1995. "On Players' Models of Other Players - Theory and Experimental Evidence." Games and Economic Behavior 10: 213-254.

Weber, R. A. 2003. "'Learning' with no feedback in a competitive guessing game." Games and Economic Behavior, 44 (1), 134-144.

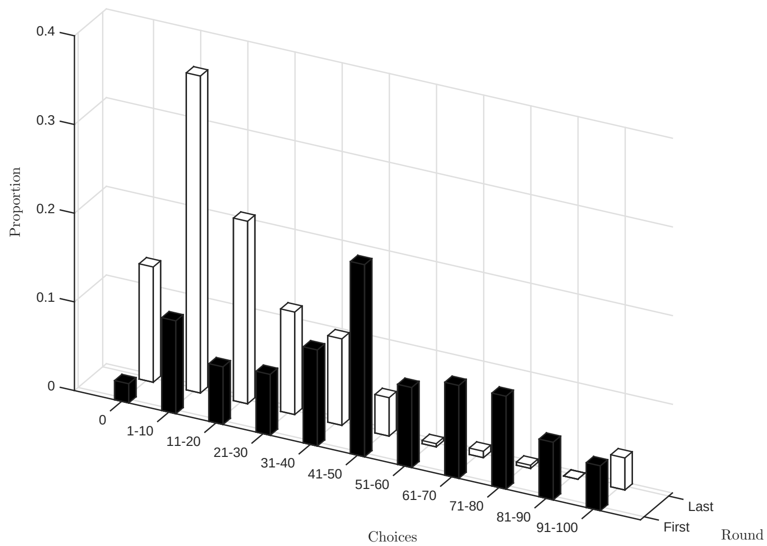

Figure 1: p-beauty Contest: First and Last Round Data

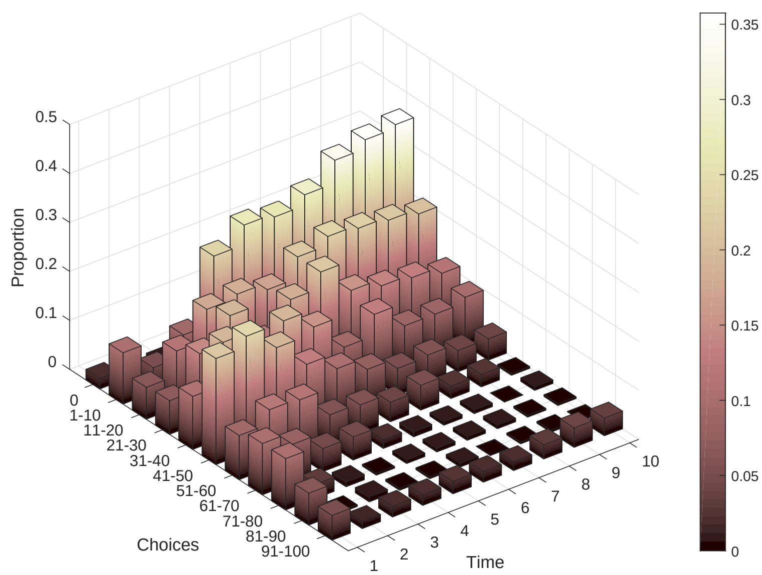

Figure 2: p-beauty Contest: Actual Data Frequency

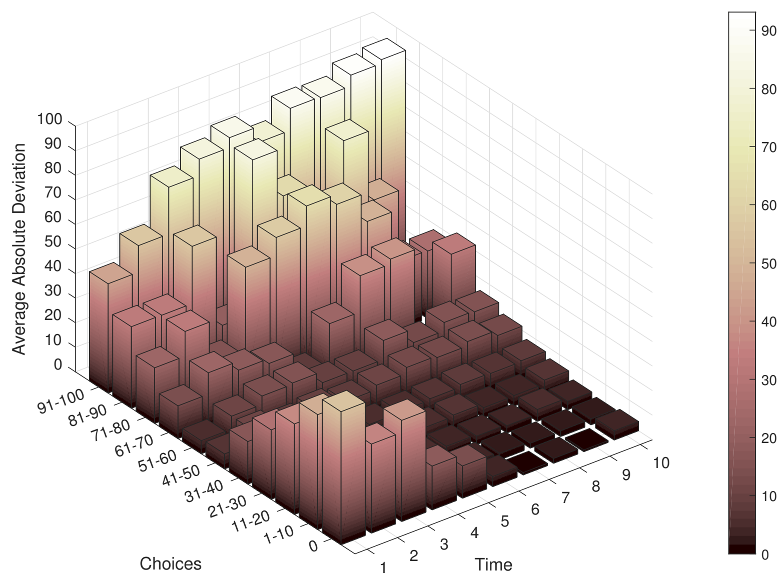

Figure 3: p-beauty Contest: Absolute Deviation of $BL_k(t)$ from Data

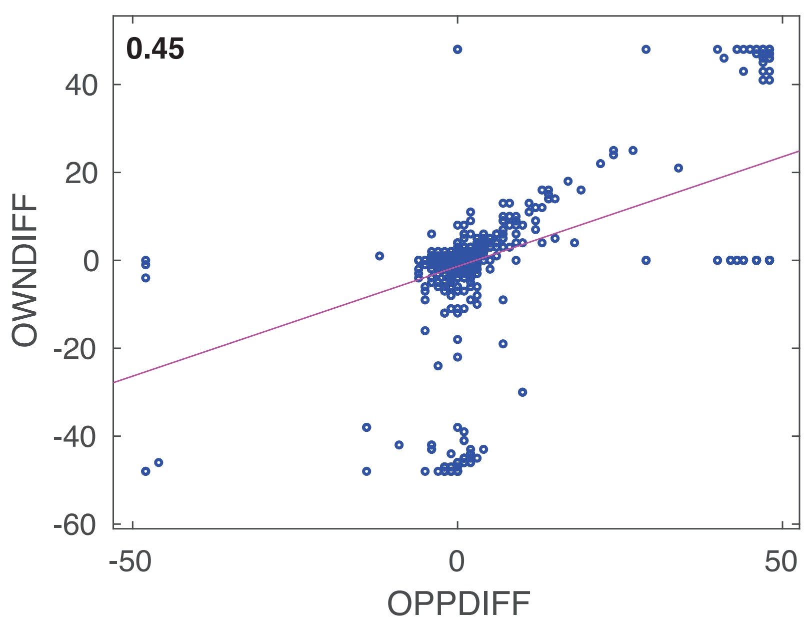

Figure 4: Correlation between Own Rule Changes (OWNDIFF) and Difference from Opponents' Rule (OPPDIFF)

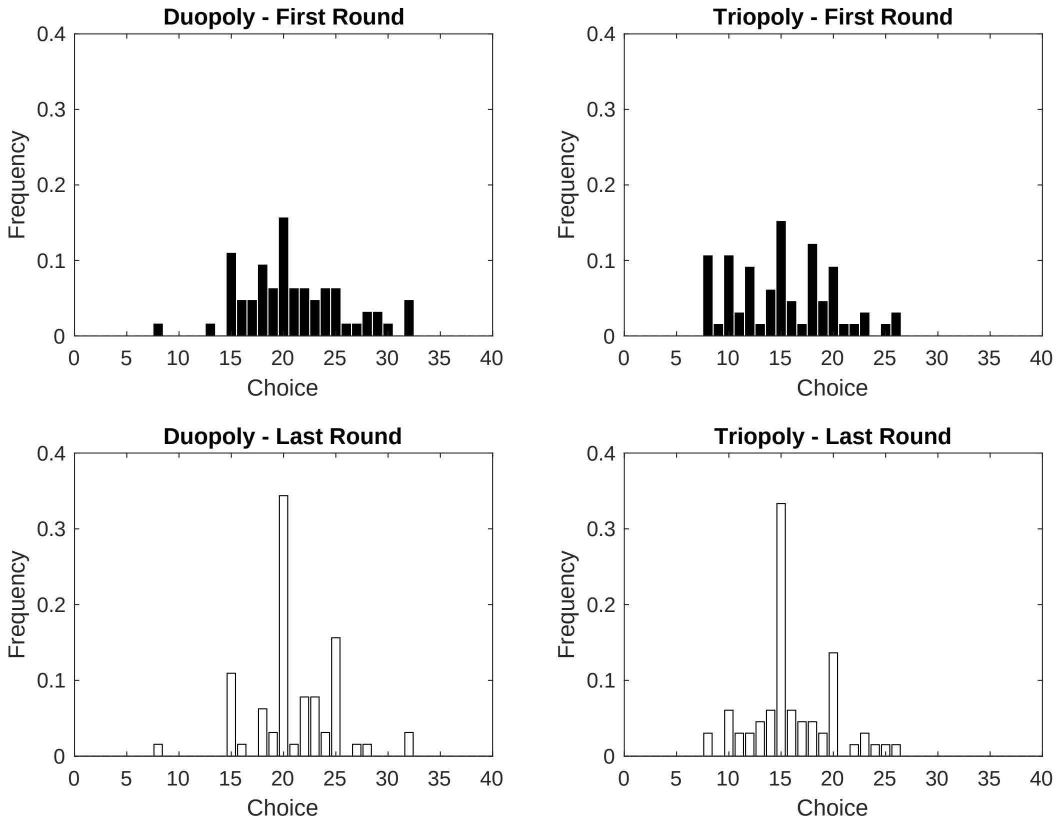

Figure 5: Cournot Quantity Game: First and Last Round Data

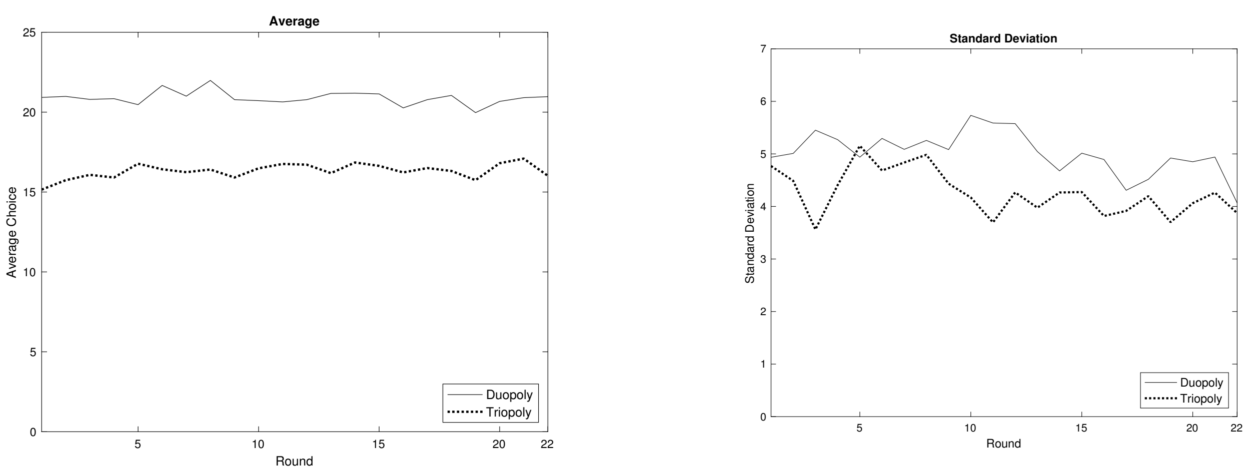

Figure 6: Cournot Quantity Game: Average and Standard Deviation of Choices by Round

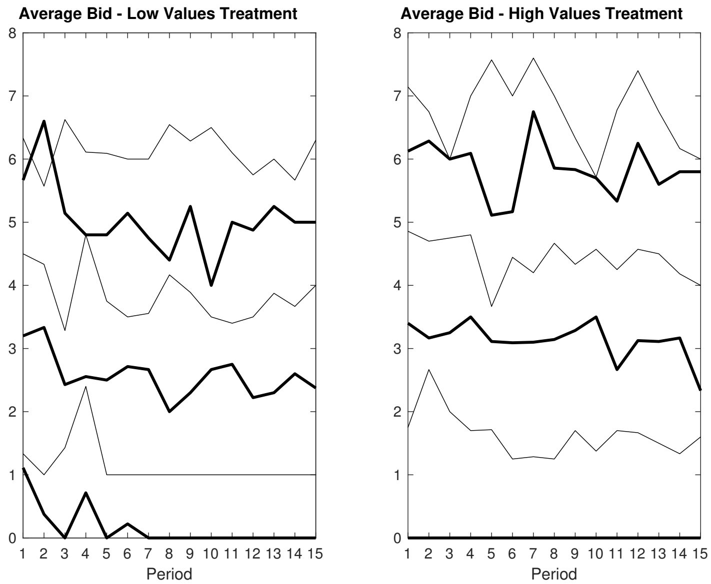

Figure 7: First-Price, Private-Value Auction Data (Fig 2 of Goeree et al. (2002)): Average bids by period. Key: Dark lines are averages for the first, third, and fifth lowest values in each treatment.

	Models			
	Dynamic $a_i^0(t)$		Static $a_i^0(t)$	
	Bayesian	Static Level-k	Bayesian	Static Level-k
	$BL_k(t)$	$L_k(t)$	BL_k	L_k
Level-0 Players				
$\hat{\alpha}$	0.47	0.48	0.27	0.32
	(0.03)	(0.01)	(0.00)	(0.01)
$\hat{\omega}(0)$	61.97	63.74	100.00	39.71
	(6.77)	(2.21)	(0.00)	(0.54)
$\hat{\delta}$	0.11	0.12	$\underline{1}$	$\underline{1}$
	(0.05)	(0.01)		
$\hat{\sigma_0}$	1.32	1.31	10.00	1.34
	(0.02)	(0.01)	(0.00)	(0.01)
Bayesian Updating				
$\hat{\tau}$	0.00	0.07	3.84	18.06
	(0.16)	(0.04)	(0.00)	(0.00)
$\hat{N}(0)$	2.00	$\underline{\infty}$	2.87	$\underline{\infty}$
	(21.35)		(0.02)	
$\hat{\rho}$	0.00	$\underline{1}$	0.30	$\underline{1}$
	(0.30)		(0.00)	
$\hat{\sigma}$	0.46	0.45	0.43	1.34
	(0.05)	(0.02)	(0.00)	(0.01)
LL	-2534.22	-2539.74	-2957.45	-3349.25
χ^2	–	11.04	846.46	1630.06
(p-value, dof)	–	(0.00, 2)	(0.00, 1)	(0.00, 3)
AIC	5084.44	5091.48	5928.90	6708.50
BIC	5131.85	5127.04	5970.39	6738.13
Two-fold Cross Validation	-1260.12	-1262.13	-1472.57	-1673.39

Table 1: p-beauty Contest: Parameter Estimates

	Models			
	Dynamic $a_i^0(t)$		Static $a_i^0(t)$	
	Bayesian	Static Level-k	Bayesian	Static Level-k
	$BL_k(t)$	$L_k(t)$	BL_k	L_k
Level-0 Players				
$\hat{\alpha}$	0.11	0.12	0.45	0.47
	(0.03)	(0.02)	(0.01)	(0.02)
$\hat{\omega}(0)$	14.87	14.87	16.58	16.73
	(0.54)	(0.46)	(0.06)	(0.10)
$\hat{\delta}$	0.77	0.75	$\underline{1}$	$\underline{1}$
	(0.09)	(0.08)		
$\hat{\sigma_0}$	9.25	9.54	4.67	4.34
	(1.82)	(1.59)	(0.11)	(0.17)
Bayesian Updating				
$\hat{\tau}$	0.00	0.03	0.00	0.05
	(0.01)	(0.03)	(0.01)	(0.03)
$\hat{N}(0)$	52.73	$\underline{\infty}$	100.00	$\underline{\infty}$
	(30.41)		(42.00)	
$\hat{\rho}$	0.45	$\underline{1}$	0.02	$\underline{1}$
	(0.31)		(0.34)	
$\hat{\sigma}$	4.02	3.98	4.13	4.34
	(0.10)	(0.08)	(0.07)	(0.12)
LL	-8308.34	-8309.92	-8342.75	-8344.52
χ^2	–	3.15	68.81	72.36
(p-value, dof)	–	(0.21, 2)	(0.00, 1)	(0.00, 3)
AIC	16632.69	16631.84	16699.50	16699.05
BIC	16680.36	16667.59	16741.21	16728.84
Two-fold Cross Validation	-4038.40	-4033.82	-4058.13	-4052.43

Table 2: Cournot Quantity Game: Parameter Estimates

	Models			
	Dynamic $a_i^0(t)$		Static $a_i^0(t)$	
	Bayesian	Static Level-k	Bayesian	Static Level-k
	$BL_k(t)$	$L_k(t)$	BL_k	L_k
Level-0 Players				
$\hat{\alpha}$	0.13	0.13	0.13	0.13
	(0.01)	(0.01)	(0.01)	(0.01)
$\hat{\omega}(0)$	5.00	5.00	5.00	5.00
	(0.00)	(0.31)	(0.00)	(0.00)
$\hat{\delta}$	1.00	1.00	$\underline{1}$	$\underline{1}$
	(0.02)	(0.02)		
$\hat{\sigma_0}$	0.86	0.86	0.86	0.86
	(0.02)	(0.02)	(0.00)	(0.01)
Bayesian Updating				
$\hat{\tau}$	0.02	0.13	0.02	0.13
	(0.01)	(0.02)	(0.00)	(0.01)
$\hat{N}(0)$	0.11	$\underline{\infty}$	0.10	$\underline{\infty}$
	(9.17)		(0.00)	
$\hat{\rho}$	0.99	$\underline{1}$	0.95	$\underline{1}$
	(0.25)		(0.03)	
$\hat{\sigma}$	0.58	0.62	0.59	1.34
	(0.01)	(0.01)	(0.01)	(0.01)
LL	-1620.70	-1636.25	-1620.70	-1636.25
χ^2	–	31.10	0.00	31.10
(p-value, dof)	–	(0.00, 2)	(1.00, 1)	(0.00, 3)
AIC	3257.41	3284.50	3255.41	3282.50
BIC	3298.13	3315.04	3291.04	3307.95
Two-fold Cross Validation	-827.41	-844.69	-823.10	-841.18

Table 3: Private Value Auction: Parameter Estimates