Opponent's Foresight and Optimal Choices

Jeevant Rampal[*,†]

November 25, 2016

Abstract

This experimental study investigates *how* and *why* the behavior of experienced players, who understand the "sure-win" strategy in a "winner-take-all" sequential move game, varies systematically based on two types of information about the opponent's expertise. Treatment (1): experienced subjects are told their opponent's experience-level in the game. Treatment (2): a different set of experienced subjects are only shown their opponent's play against a computer. We find that both (i) exogenous information, and (ii) endogenous inference about the opponent's inexperience increase the probability with which experienced players abandon the "sure-win" strategy and try for a higher payoff attainable only by winning from a losing position, i.e., a position from which one wins only if the opponent makes a mistake. A maximum likelihood analysis shows that a model of limited foresight and uncertainty about the opponent's foresight (Rampal (2016)) explains the data better than the Dynamic Level-k (Ho and Su (2013)) and AQRE (McKelvey and Palfrey (1998)) models.

Keywords: Race Game, Experience-level, Limited Foresight, Dynamic Level-k, Agent Quantal Response

1 Introduction

There are countless examples where economic players participate in multi-stage interactions with each other. We focus on finite perfect information games where each player is aware of all prior decisions made in the interaction. The standard game-theoretic method for making outcome predictions about such

[*]Email: rampal.5@osu.edu. This study is funded by the Decision Sciences Collaborative, OSU and the Journal of Money Credit and Banking, Department of Economics, OSU.

[†]I would like to thank James Peck, John Kagel and P.J. Healy for their valuable advice. I also want to thank Dan Levin for his inputs. All errors are my own.

interactions is to solve for the Subgame Perfect Nash Equilibrium (henceforth SPNE). The experimental literature has studied two key aspects of the SPNE model: first, do real players perform backward induction in solving their strategies (see Johnson et al (2002) and Binmore et al (2002))?[1]Second, how does the belief about the opponent's expertise in backward induction affect strategies? In two studies that use expert chess players, Palacios-Huerta and Volij (2009), and Levitt, List and Sadoff (2009) found contrary results in the context of the Centipede game. The former found that expert chess players were more likely to play SPNE strategies against other expert chess players, but not against student subjects. The latter did not find such a systematic difference in behavior. However in a zero-sum "race" game they found high incidence of SPNE play by expert chess players.[2]

Our paper contributes to the second aspect of the debate. We investigate *how* and *why* the optimal strategy of expert players, in a perfect information "winner take all" dynamic game, is affected by the opponent's given or perceived experience-level in the game. We *add* to the debate by investigating the question of whether subjects change behavior based on their *inference* about their opponent's experience-level, where the inference is drawn by observing the opponent's previous play during the game, without being explicitly told the opponent's level of experience/expertise. We induce different levels of expertise by providing subjects with different levels of experience. In particular, the sequential move game we use is a "winner-take-all" game which has a "sure-win" strategy, similar to the "race" game. The data confirms that the experienced subjects learn the "sure-win" strategy of the game in their training session. Our key findings are: (i) experienced subjects are more likely to risk a loss and attempt to attain a higher payoff (higher than the payoff from winning using the "sure-win" strategy), which is attainable only if the opponent makes a mistake (plays a dominated strategy), when they are informed that their

[1]Not necessarily, according to Johnson et al (2002), who monitored subjects' "look-ahead" in a bargaining game, and found that the amount of stages subjects look forward was different across subjects. Binmore et al (2002) also study bargaining and find evidence showing that subjects' behavior violates subgame and truncation consistency.

[2]Palacios-Huerta and Volij (2009) found that expert chess players stop a one-shot Centipede game at the first node (the SPNE prediction) more often when matched with other chess players as opposed to when they are matched with a student subject. They attribute this result to common knowledge of rationality among capable players, and thus conclude that it is the level of rationality and information about opponent's rationality that determines outcomes rather than altruism or social preference. Levitt, List and Sadoff (2011) found that contrary to Palacios-Huerta and Volij (2009), expert chess players, playing with each other, play like student subjects in the Centipede game and cooperate in the beginning of the game. Levitt et al (2011) also test a zero-sum game called the "race to 100 game," where each player has a dominant strategy. The "race to 100 game" is a zero-sum game, and there are no benefits from cooperation. They find a high incidence of SPNE play by expert chess players. Their results suggest that "failure to stop at the first node in centipede has little to do with an inability to reason backward." They find that the "best inductors" in the race game had low stoppage rates in the centipede game at any node.

opponent is inexperienced. In a different treatment with different subjects we find that: (ii) experienced subjects' accuracy of beliefs about the opponent's experience level improves upon observing more moves of their opponent, and (iii) experienced subjects use the inference about the opponent's inexperience, and increase the likelihood of risking a loss to try to attain the higher payoff against the opponents who play a dominated strategy in the earlier part of the interaction.

Therefore, finding (i) confirms the Palacios-Huerta and Volij (2009) result in the context similar to the "race to 100" game used by Levitt, List and Sadoff (2011). That is, our findings point to a systematic failure of SPNE within the paradigm of *selfish and rational* behavior. We can be sure that the behavior is indeed selfish because the game is "winner-take-all" and thus there is no possibility of a "fair" split of the payoffs. There is also no possibility of co-operation among players to increase joint payoffs like in the Centipede game. We know that experienced players are acting rationally in choosing to deviate from the SPNE strategy, and risking losing, because we can observe that they converge to SPNE behavior and play the "sure-win" strategy when playing another experienced opponent. In all previous literature and in our first treatment, each player is informed about his and the opponent's expertise. One can argue that subjects being told their opponent's experience-level/expertise reduces the relevance of these findings to dynamic games in the real world. But in treatment 2, our finding is that experienced subjects infer the experience-level of the opponent by observing his/her past behavior, and then act on this endogenous inference to take advantage of the opponent's perceived inexperience. We also *add* to the literature by exploring the "why," i.e., by trying to distinguish among which model explains such data the best. A model with limited foresight and uncertainty about the opponent's foresight (Rampal (2016)) fits the data better than the dynamic Level-k model (Ho and Su (2013)) and the quantal response equilibrium for extensive games (McKelvey and Palfrey (1998)). The Ho and Su (2013) Level-k model relies on subjective beliefs about the opponent's cognitive "level" to explain the data. The reason that the Ho and Su (2013) Level-k model does worse is that it doesn't allow for a failure to play weakly dominant strategies. But this is often the case in our data because a "high" level of foresight is required for understanding the weakly dominant strategy in our sequential move game.

This experimental study uses two modified versions of a "race game." The race game is a two player, alternate move, perfect information, zero sum game. One version of the race game we use has the following specifications: two players move alternately, choosing numbers of "items to remove" from a box containing 9 items. At every move one can remove 1, 2 or 3 items. The player who removes the last item, *loses*, while his opponent wins. We call this game the Avoid 9 game. Note that if the second mover plays (4 minus the opponent's previous choice) at each of his moves, he is guaranteed to win. In effect, it is a race to remove the 8^{th} item. Similar race games have been extensively tested in the lab; c.f. Gneezy et al (2010), Dufwenberg et al (2010), Levitt et al (2011), Mantovani (2014).

Some of the reasons why the literature studies race games are: (a) they are zero sum games, which gets rid of explanations of the data originating from other-regarding preferences; (b) the SPNE of a race game is an equilibrium in weakly dominant strategies, and therefore the SPNE strategy remains a best response to the other player's strategy no matter what one might believe about the opponent (c) the SPNE is easy to understand (once known) and apply once understood, and therefore deviations from SPNE strategies are easy to spot; (d) a player needs to reason backwards to understand the optimal strategies, and therefore the game serves as a good test of foresight.

The Avoid 9 game we use is different from the standard race game in the following ways. First, to decide the first and second mover in the game, we ask both players their first mover/second mover choices. After both players make their simultaneous choices, one of the choices is selected with 50 percent chance each. Second, we make the prize for winning as the first mover equal to 500 experimental currency units (henceforth, ECUs), and the prize for winning as the second mover as 200 ECUs. The first/second mover choices of the experienced subjects tell us if they want to play the "sure-win" strategy or if they want to try for the higher payoff, attainable only if the opponent makes a mistake as the second mover.

In treatment 1, we provide experience in the Avoid 9 game to half the subjects. Like Gneezy et al (2010), the experience leads to convergence towards SPNE play. That is, the experienced subjects learn the second mover advantage as they play 12 repetitions of the Avoid 9 game. Then we run a combined sub-session of experienced and inexperienced subjects stranger matched with each other. In the combined sub-session each player is told if the opponent is "experienced" or not. We find that experienced subjects are significantly more likely to choose to be the first mover, and put themselves in a losing position to try and win the bigger prize, if their opponent is inexperienced than if their opponent is experienced. We also observe that inexperienced subjects display no such difference in behavior when playing against inexperienced or experienced subjects.

In another treatment, treatment 2, we endogenize the process of learning about the opponent's experience-level. Each subject in each round goes through two parts of the round. First, he plays an "avoid removing the 13^{th} item" game with a perfectly playing computer. Label this game as C13. In C13: (i) the box contains 13 items and the player who avoids removing the 13^{th} item wins; (ii) the human subject decides who the first mover is, him or the computer; (iii) there is only one prize, 500ECUs to win. In each round, each subject first finishes C13 and then plays his matched human opponent in an Avoid 13 game, with rules exactly like treatment 1, except that the box contains 13 items. That is, the Avoid 13 game played with a human subject contains the extra incentive to win as the first mover like the Avoid 9 game, and the second mover advantage is also the same. The earnings from the round are a sum of the earnings from the two parts of the round: C13 with a computer and Avoid 13 with a human subject. Similar to the first treatment, we provide experience to half the subjects, and make them do 8 repetitions of this two part round. In each two-part round, the experienced subject is shown his opponent's moves versus the computer. After half the subjects are experienced, we mix experienced and inexperienced subjects together who are then randomly re-matched every round. We observe that experienced subjects are more likely to choose to be the first mover in their Avoid 13 interaction with a human opponent, if that human opponent lost his C13 game, compared to the case where the opponent won his C13 game. That is, replacing explicit information about the opponent's experienced level with information about the opponent's behavior in the first part of the round also produced a similar effect to actually telling the experienced player about the level of experience of his opponent. Further, we ask the experienced subject about his opponent's level of experience at two points of his opponent's play in C13, and find that the accuracy of the experienced subjects' answer improves after observing more moves of his opponent.

2 Related Literature

Other studies of race games include Dufwenberg et al (2010) and Gneezy et al (2010) who study different aspects of learning in the context of these games. A modified version of a race game is used by Mantovani (2014) to also show that subjects indeed display limited foresight. Mantovani (2014) observes evidence of a "sophisticated type" who puts himself in a losing position to attain a higher prize. But, based on their own data analysis, they rule out such behavior. The Mantovani (2014) model deals with limited foresight, but doesn't model uncertainty about the opponent's foresight. Thus, in their model, exogenous/endogenous information about the opponent's foresight would not make a difference to the model's prediction, contrary to our data results. Our design has two new elements to the race games studied in the literature: (a) in treatment 1, the simultaneous first/second mover decision stage to explore the effect of the belief about the opponent's foresight on the experienced player's optimal strategy; and (b) in treatment 2, the opportunity for endogenous learning of the opponent's foresight/level, based on observation of the opponent's prior play.

The effect of mixing experienced and inexperienced subjects in dynamic games has been studied in different contexts in the literature. Coq and Sturluson (2012) study the effect of exogenous information about the opponent's experience in the context of the quantity precommitment dynamic game. They show that when experienced subjects play against an inexperienced opponent, the former choose higher capacities than when playing against an experienced opponent. They explain their findings using the AQRE model of McKelvey and Palfrey (1998). Dufwenberg et al (2005) study the effect of mixing inexperienced subjects among experienced subjects in multi-stage asset pricing games. There are some important differences in these studies compared to our paper. First, the games are not "winner take all" and therefore the loss from a suboptimal strategy, if the opponent plays optimally, is not as stark; and second, they have no treatment with endogenous learning of the opponent's experience level.

The effects of combining experienced and inexperienced subjects in simultaneous move games have been studied extensively.[3] We refer the interested reader to Agranov et al (2012), Aloui and Penta (2016), Gill and Prowse (2014) and Slonim (2005).

We compare the Rampal (2016) model with the dynamic Level-k model of Ho and Su (2013), and the AQRE model of McKelvey and Palfrey (1998), to explain the data. Kawagoe and Takizawa (2012) also

[3] Agranov et al (2012) study the effect of manipulating their subjects' beliefs about their opponent's cognitive levels in a simultaneous move 2/3 guessing game. Aloui and Penta (2016) endogenize the choice of level in a Level-k framework by modeling and studying the incentives and costs of choosing a certain cognitive level. In a simultaneous move setting, these incentives and costs are shown to depend on the payoffs and the opponent's level. They disentangle the effect of one's own cognitive limitation from the effect of one's beliefs about the opponent's cognition. Gill and Prowse (2014) study convergence of play towards equilibrium across repetitions of a simultaneous move p-beauty contest based on their measure of cognitive ability and character skills. Slonim (2005) tested the effects of varying the experience levels of the players in simultaneous move games similar to the beauty contest game. They find that only experienced subjects, and not inexperienced subjects, condition their behavior on the opponent's experience level. The also find that introduction of new players interrupted the convergence towards equilibrium.

have a sequential level-k model that they apply to data on the Centipede game. We test their model with respect to our data as a robustness check of our results.[4] Levin and Zhang (2016) is a working paper which brings together the Nash Equilibrium and the Level-k model. They extend their model to sequential move games. An application of their model to our data is left for future work. Also see Crawford , Costa-Gomes and Iriberri (2013) for a thorough survey of the work using the Level-k theory, most of which has been in simultaneous move games, unlike the sequential move game we study here.

3 Experimental Design: Treatment 1

The experiment for treatment 1 was conducted at The Ohio State University's experimental economics laboratory using the laboratory's subject pool.

3.1 Exogenous Information Treatment

Treatment 1 comprised of 13 sessions, with a total of 154 subjects. The sessions were conducted using zTree (Fischbacher (2007)). Each session contained between 8 and 18 subjects. A session lasted 62 minutes on average with an average payment of USD 14.45. Each session used two games. First, a game called *Avoid Removing the* 9^{th} *Item.* We refer to this game as the Avoid 9 game. The other game used was a three period sequential bargaining game (Rubinstein (1982)) with the common discount factor = 0.6. The bargaining game and its results are not relevant to this paper. We now explain the rules of the Avoid 9 game.

The Avoid 9 game is played by 2 players. There are 9 items in a box. At every move, each player can choose to remove 1, 2, or 3 items from the box. If, at a move, the number of items left in the box is 2 or 1 then the maximum number of items that a player can remove at that move is 2 or 1, respectively. Before the game begins, both players simultaneously choose between "First Mover" and "Second Mover", i.e., if they want to be the first mover/ second mover in the game. One of the two players' choices is implemented

[4] Kawagoe and Takizawa (2012) use a model of level-k where each level plays best responses with logit errors like the AQRE model of McKelvey and Palfrey (1998). We use their model as a robustness check of our results. However, we argue that a comparison with a logit Level-k model like theirs is not a reasonable model to compare. This is because, unlike the logit level-k model, when we compare the Ho and Su (2013) sequential level-k model to the LFLE model, both bounded rationality models have error-less *strategies* (although both of them need an error term on the *outcomes* to keep the likelihood function finite), and therefore the comparison is among similar, though non-nested, models.

with 50 percent chance each. After it has been decided who the first mover is, the two players choose alternately. All prior choices are displayed to both players. If a player removes the 9th item, he loses, and his partner wins. If a player wins as the first mover, his payoff for the round is 500 ECUs (Experimental Currency Units). If a player wins as the second mover, his payoff for the round is 200 ECUs. If a player loses, he gets 50ECUs. The conversion rate used was 60ECUs = 1USD.

This game is a "race" game (a term used by Gneezy et al (2010), Levitt et al (2011), Mantovani (2014); Dufwenberg et al (2010) refer to their "race" game as "game to 21") with the first mover/ second mover decision stage (henceforth F/S decision stage) added to it. The game is called a "race" game because a player has to remove the 8th item to win, i.e., it is a race to 8. Define a "position" n of the Avoid 9 game as the set of nodes such that for any node in that position, the sum of items removed at all nodes preceding that node is n. A particular position can contain several nodes. A node is in a winning position if the player moving at that node can choose a strategy (in the subgame with that node as the root) that guarantees a win, regardless of the opponent's strategy. A node is in a losing position if the opponent of the player moving at that node can choose a strategy (in the subgame with that node as the root) that guarantees a win to the opponent, regardless of the player's strategy. The winning positions are $W_9 = \{1, 2, 3, 5, 6, 7\}$, and losing positions are $L_9 = \{0, 4, 8\}$. A position of 0 is the set of decision nodes of the selected first mover immediately succeeding the F/S decision stage. In the Avoid 9 game, the second mover can always win. As the second mover, one can choose (4 minus the opponent's previous choice) at each move to remove the 8th item. However, if the second mover fails to put the opponent in L_9 after one of his moves, then the first mover is guaranteed a win by playing a strategy which puts the second mover at a losing position at all of the second mover's subsequent moves. Note that in the Avoid 9 game, the SPNE is for both players to choose "Second Mover" (henceforth S) and then choose 3, 2, and 1 from positions in {1, 5}, {2, 6}, and {3, 7}, respectively. The SPNE strategy places no restriction on actions from a position in L_9. We call this SPNE strategy as the "perfect" strategy.

The *perfect strategy* in Avoid 9:

$$perfect\ action\ = \begin{cases} S & at\,\{F/S\}\ decision\ stage \\ 3 & if\ Position \in \{1,5\} \\ 2 & if\ Position \in \{2,6\} \\ 1 & if\ Position \in \{3,7\} \\ Arbitrary & if\ Position \in \{0,4\} \end{cases} \tag{1}$$

Note that if one's F/S decision is implemented, the perfect strategy is a "sure-win" strategy. However, due to the extra incentive to win as the first mover, the perfect strategy is *not* a weakly dominant strategy. If a risk neutral rational player believes with probability at least $\frac{1}{3}$ that his opponent will make a mistake, then choosing "First Mover" (henceforth F) is optimal. Let any subgame with its root at a node in position 0 be labeled as $A9_{sub}$. Note that in $A9_{sub}$, the perfect strategy *is* weakly dominant because the payoff from winning/losing is already decided, and the perfect strategy is a "sure-win" strategy from a winning position.

The design of treatment 1 is as follows:

1. In each session each subject went through 2 sub-sessions. First, the subjects were split into two types: **Experienced (Exp)** and **Inexperienced (Inexp)**, with Inexp subjects being at least 50 percent of the total subjects in any session. In the first sub-session (training sub-session):

   (a) **Exp subjects** (74 total subjects) **trained**. **Exp** subjects played 12 rounds of the Avoid 9 game among themselves. One round was randomly drawn as the round determining earning from the first sub-session.

   (b) **Inexp subjects** (80 total subjects) were not told about the Avoid 9 Game. They played between 5-8 rounds of a three period bargaining game (Rubinstein (1982)) among themselves.[5] One round was randomly drawn as the round determining earning from the first sub-session.

2. In the second sub-session (combined sub-session) Exp and Inexp subjects were mixed. They played 8 rounds (round numbers 13-20 of the session) of the Avoid 9 Game together. The subjects were

[5] The bargaining data will be the subject of another paper.

stranger matched into pairs before the beginning of each round. At all points during the Avoid 9 Game, each player was told if their opponent was Exp (if the other was Exp, we told the subject that the other was of *type S,* i.e., a subject who had played the *same* Avoid 9 game in the first sub-session) or Inexp (if the other was Inexp, we told the subject that the other was of *type D,* i.e., a subject who had played a *different* game in the first sub-session).

3. One round from the second sub-session was drawn at random for payment and added to the payment due from the first sub-session. All payments were made at the end of the session.

Subjects moved sequentially, with each subject, at each move being given a clock with 45 seconds on it to remind them to move. The clock could only flash if the time taken was more than 45 seconds, and the game did not proceed without the subject's choice.

4 Data Results for Treatment 1

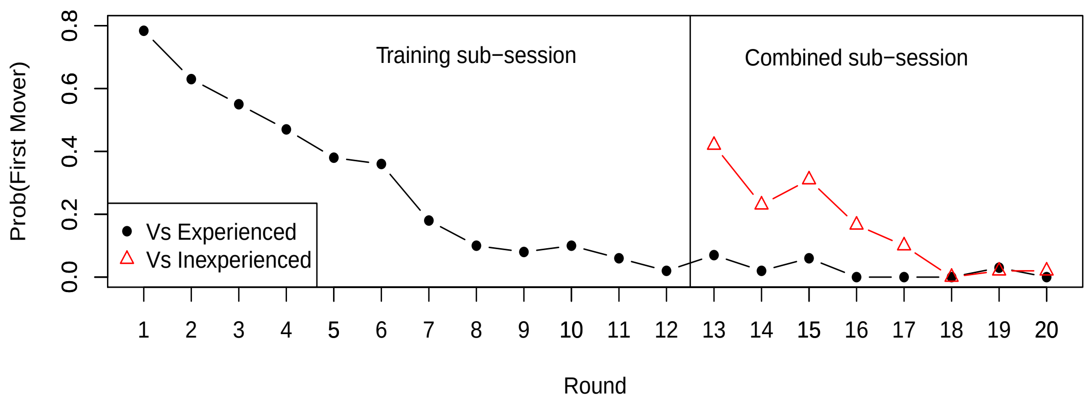
In our analysis of the data, the following aspects of the subjects' behavior will be important:

1. How did the Exp subjects' "First/Second Mover" choices vary based on the opponent's experience level.

2. Was there any systematic pattern of imperfect play?

The main results of treatment 1 can be observed in Figure 1. Figure 1 plots the average proportion of the Exp subjects choosing "First Mover" by round and opponent type. Figure 1 shows the choices of experienced subjects who had understood the "sure-win" perfect strategy in the training sub-session. We can see that in the combined sub-session (rounds 13-20), when the experienced subjects were playing against an inexperienced opponent (red triangles), as opposed to an experienced opponent (black dots), they were significantly more likely to choose "First Mover" and put themselves in a losing position in order to attain a higher winning prize.

Figure 1: Experienced Subjects' Behavior by Opponent's Experience



Notes. The figure depicts the round-wise proportion of experienced subjects matched with an experienced opponent who chose "First Mover" (black dots) and the proportion of experienced subjects matched with an inexperienced opponent who chose "First Mover" (red triangles). The first mover loses if the second mover plays perfectly. The latter proportion is significantly higher, with a p-value < 0.05, for each of the rounds 13 through 16. This shows that the experienced subjects were more likely to risk losing to try for the higher prize when matched with an inexperienced opponent. The training sub-session was the first 12 rounds, and the combined sub-session was rounds 13-20.

Result 1(a): *Training Successful*: in the training sub-session (rounds 1-12) there was a significant increase in the proportion of Exp subjects playing the Avoid 9 game perfectly. The proportion of Exp subjects playing the imperfect strategy almost converged to zero.

Recall that the selected second mover cannot lose the Avoid 9 game if he plays perfectly. The reward for winning as the first mover (second mover) is 500 ECUs (200ECUs). This tempts the subjects to choose F until they gain understanding of the perfect strategy of the game and until they are sufficiently convinced about their opponent's understanding of the perfect strategy. In round 1, 78.4 percent of Exp subjects had chosen F. This percentage declined steadily (see Figure 1) through the rounds, and in round 12, only 2.7 percent of the Exp subjects chose F. This decline in proportions is highly significant with a p-value of approximately 0. The 2.7 percent of Exp subjects choosing F in round 12 is a positive but insignificant. Recall that once the first and second movers are decided, i.e., in the subgame $A9_{sub}$, it is a weakly dominant strategy to play perfectly. The percentage of pairs of subjects who displayed imperfect play in $A9_{sub}$, reduced from 36.5 percent in the first two rounds, to 5.4 percent in the last two rounds of the training sub-session. Although 5.4 percent is positive and significant, the reduction in imperfect play in $A9_{sub}$ is highly significant with a p-value of approximately 0.[6]

Result 2(a): *Opponent's experience level has a significant effect:* in the combined sub-session (rounds 13-20), experienced subjects were more likely to choose F ("First mover") against an inexperienced opponent than an experienced opponent.

Figure 1 illustrates result 2(a). Round-wise tests for difference in proportions show that for each of rounds 13-16, the Exp subjects matched with an Inexp opponent chose F at a significantly higher rate (p-value ≤ 0.015 for each round) than the Exp subjects matched with an Inexp opponent (see Table 8 in the Appendix).

Next, consider the probit results for 74 Exp subjects and 80 Inexp subjects' choice data from rounds 13-20, the combined sub-session. Table 1, model (1), marked with Exp, reports the results from a probit estimation with the Exp subjects' choice of F/S in the Avoid 9 game as the dependent variable. The dependent variable took value 1 if an Exp subject chose F, and 0 if he chose S. The independent variables are: (i) a dummy for if the opponent was Inexp (value 1) or Exp (value 0); (ii) Round variable; (iii) Constant term; (iv) Session dummies. Table 1, column 3, marked with Inexp, reports the results from the same probit estimation done with the Inexp subjects' choice of F/S in the Avoid 9 game as the dependent variable. The definition of the dependent variables implies that a positive coefficient on an independent variable means that a higher value of the independent variable increased the probability of the player choosing F.

[6] We are looking at group level data because in $A9_{sub}$, if the selected second mover plays perfectly, any action sequence of the selected first mover is "perfect." Thus, it is more meaningful to look at the group level data than individual level data.

Table 1

Factors Influencing Probability of choosing "First Mover"

Dependent Variable: Prob(First Mover)

	Exp	Inexp
	(1)	(2)
Opponent is Inexp	1.18***(0.33)	0.12(0.13)
Round	-0.31***(0.07)	-0.43***(0.04)
Constant	3.25***(0.87)	6.2(0.67)
No. of Obs.	544	640
Pseudo R^2	0.3342	0.3273
Session Dummies	Yes	Yes

* p <0.10, ** p <0.05, *** p <0.01. Numbers in parenthesis are robust standard errors clustered at subject-level. For model (1), the dependent variable is Exp players' choice of first (takes value 1) or second mover (takes value 0). The sample size is 68 Exp subjects (after dropping session 8) across rounds 13-20. For model (2), the dependent variable is Inexp players' probability of choice of first (takes value 1) or second mover (takes value 0). The sample size is 80 Inexp subjects across rounds 13-20.

Focusing on model (1), which reports the probit results for the Exp subjects,[7] we observe the following. The highly significant coefficient on the dummy variable for "opponent is Inexp" (p-value < 0.001) suggests that Exp subjects were significantly more likely to choose F when playing against an Inexp opponent than an Exp opponent. The highly significant negative coefficient of the round variable (reflected in the downward sloping lines in Figure 1) can be caused by a combination of two factors. As subjects play more rounds, two things happen: (a) more of them learn the perfect strategy; (b) there is an increase in the subjects' belief that the opponent (Exp or Inexp) understands the "sure-win" perfect strategy in $A9_{sub}$. As almost all the Exp subjects had learnt the perfect strategy after the training sub-session (result 1(a)), (b) appears to be the reason for the significant reduction in the proportion of Exp subjects choosing F as the rounds progressed.

[7] In addition to the results in Table 1, for the probit estimates with the Exp subjects' F/S choices, session numbers 2^{***}, 4^{***}, 5^{**}, 7^{***}, 9^{**}, 11^{**} and 12^{***} had a significantly negative coefficient, while session 8 perfectly predicted the choice S by Exp players in rounds 13-20, and thus was omitted without affecting the likelihood or the estimates of the other coefficients. For the probit estimation with the Inexp subjects' F/S choices, only sessions 5^{*} and 9^{*} had a significantly positive effect (p-value < 0.1) on the probability to choose F.

The positive but statistically insignificant coefficient of the "Opponent is Inexp" dummy in model (2) shows that while Inexp subjects also increased the likelihood of choosing F when they faced an Inexp opponent, this increase was statistically insignificant. The round variable's highly significant negative coefficient (p-value < 0.001) implies that as the Inexp subjects played more rounds of the Avoid 9 game, there was a significant increase in the proportion of Inexp subjects who understood the perfect strategy and believed that their opponent also understood the perfect strategy.

Result 3(a): *Faster learning speed of Inexp subjects*: Inexp subjects, playing their first 8 rounds of the Avoid 9 game in the combined sub-session, learned the "sure-win" perfect strategy of the second mover significantly faster than Exp subjects in their first 8 rounds of playing the Avoid 9 game.

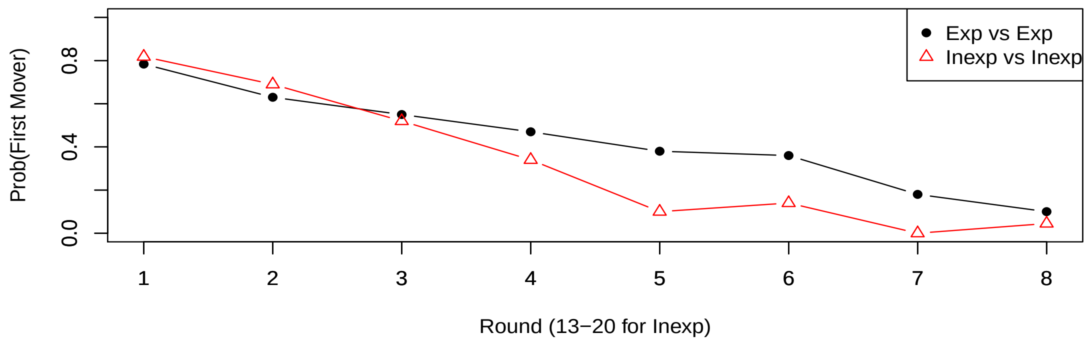
Learning can be observed in two aspects of the data: (i) the F/S decision of the players, and (ii) perfect/imperfect play as the selected second mover in $A9_{sub}$.[8] Note that while for (i) the belief about the opponent's understanding of the perfect strategy matters, for (ii), there is no possible belief than incentivizes imperfect play. Both aspects of the data lead us to conclude result 3(a).

Figure 2 depicts the learning speed difference based on the F/S choices. Figure 2 depicts the Exp players' average probability of choosing F against other Exp players in their first 8 rounds (black dots), i.e., when the Exp subjects were gaining experience in the training sub-session. It also depicts the Inexp players' average probability of choosing F against another Inexp player in rounds 13-20 (red triangles), which are the Inexp subjects' first 8 rounds.[9] Note how both Inexp and Exp start out with the same rate of choosing F, but after their respective 4 rounds, the Inexp subjects play F at a significantly lower rate. In their first three rounds, both Exp and Inexp subjects behave similarly, but for their rounds 5, 6, and 7 (session round numbers 17,18 and 19), Inexp subjects choose F in a significantly smaller percentage of cases against other Inexp players, as compared to the Exp players in their rounds 5, 6 and 7 of playing the Avoid 9 game. These differences in proportions have p-values ≤ 0.011 for each of the three rounds. These results are also confirmed by a probit estimation (details in Tables 9 and 10 in the Appendix).

[8] Looking at all possible positions where one can distinguish perfect play from imperfect play of a player does not change our results. For simplicity, we just look at the selected second mover's perfect/imperfect play at his first decision.

[9] Note that in restricting ourselves to comparing among the data where the own-type and opponent-type was the same, we are trying to maintain the belief about the opponent's understanding constant across the comparison. This cannot be truly successful because by their second round, some Inexp subjects had already played Exp subjects in a previous round. Therefore the comparison of aspect (ii) is cleaner.

Figure 2: Learning Speed Comparison



Notes. The figure depicts the behavior of experienced subjects in rounds 1-8 of the session and the inexperienced subjects' behavior in rounds 13-20 of the session (which are the inexperienced subjects' first 8 rounds). The figure depicts the round-wise proportion of experienced subjects matched with an experienced opponent who chose "First Mover" (black dots) and the proportion of inexperienced subjects matched with an inexperienced opponent who chose "First Mover" (red triangles). The first mover loses if the second mover plays perfectly. The latter proportion is significantly lower, with a p-value < 0.011, for each of the rounds 5, 6 and 7. This shows the faster learning speed of the inexperienced subjects.

As noted above, findings on the F/S decision could just be a function of the subjective beliefs of the players. However, the data on aspect (ii) (perfect/imperfect play as the selected second mover in $A9_{sub}$) also supports the result 3(a): a significantly higher proportion of Inexp subjects (87.3 percent) played perfectly as the selected second mover in their first 8 rounds compared to Exp subjects as the selected second mover in their first 8 rounds (80.7 percent). The difference has a two-tailed p-value of 0.031. This is also confirmed by a probit estimation (Table 10 in the Appendix).

The reason for result 3(a) is not clear from the data. One would expect the reason to be that the Inexp subjects were playing with a combination of Inexp and Exp subjects in their first 8 rounds. This is in contrast to Exp subjects in their first 8 rounds, who just played other subjects new to the Avoid 9 game. A test of this rationale is not confirmed by the data.[10]

[10] We divide Inexp subjects into two categories: first, those who were matched with Exp opponents in at most 2 rounds out of the first four rounds of the combined sub-session. Second, those who were matched with Exp opponents in at least 3 rounds out of the first four rounds of the combined sub-session. Contrary to intuition, we find that the proportion of Inexp subjects who choose F against an Inexp opponent in rounds 5-8 of the combined sub-session is lesser in the first category (6.1%) compared to the second category (7.5%). Also, the proportion of Inexp subjects who play perfectly as the second mover in rounds 5-8 of the combined sub-session is more in the first category (96%) compared to the second category (92%). But both these differences are statistically insignificant. This finding is robust to other such divisions of the Inexp subjects based on how often they encountered an Exp subject in the first four rounds of the combined sub-session.

The next result shows the presence of limited foresight among subjects.[11] Result 4(a) reports that as subjects got closer to the end of the game, the rate of imperfect play declined significantly. Limited foresight would imply this because if subjects have limited foresight, then as subjects get closer to the end of the game, even the lower foresight subjects have enough foresight to understand the weakly dominant strategy and play perfectly. Thus, the rate of imperfect play declines as subjects get closer to the end of the game. This result is in line with the proof of limited foresight given by Mantovani (2014). Why limited foresight implies this pattern of imperfect play is discussed more explicitly in the model comparison section.

Result 4(a): *The rate of imperfect play declined significantly closer to the end of the game.* That is, as the position increased, or as the number of items left reduced, the rate of imperfect actions reduced. In particular, a player lost from a winning position with 4 or lesser items left in only 13 (1.13%) of the 1060 Avoid 9 games played in treatment 1.

Table 2

Rate of imperfect play by position

	Position					
	1	2	3	5	6	7
Exp: First 4 Rounds	35.1	35.4	20.8	2.1	0	0
Inexp: First 4 Rounds	20.7	24.4	14.3	5.6	4.2	0
Session	10.8	12.4	8	1.9	0.4	0.4

Notes. The figures are in percentage. The rate of perfect play at each position is 100 minus the rate of imperfect play given. The figures are reported from positions from which a perfect action is distinguishable from an imperfect action. A perfect action makes the position 4 (8) from a position 1, 2, or 3 (5, 6, or 7).

Table 2 shows the observed rates of imperfect play by position. We show the rates separately for three parts of the data: (i) the first four rounds of the Exp players; (ii) first four rounds of the Inexp players; and (iii) session (training and combined sub-sessions together). Recall that a position is a particular sum of items removed. From position 1-3 (5-7), the perfect action makes the subsequent node's position as 4 (8). We don't report imperfect play rates from positions 0 and 4 because any action at those positions is

[11] The foresight of a player is defined as the number of subsequent stages that the player can observe from a given decision node (Rampal (2016)).

a perfect action. The rate of imperfect play is significantly more (p-values ≈ 0) from a position in 1-3 (6-8 items remaining) as compared to 5-7 (2-4 items remaining) in each of (i), (ii), and (iii).

The next result shows that at losing positions, i.e., at positions where SPNE is silent on which action should be chosen, there was a systematic pattern of the actions chosen: subjects were significantly more likely to remove 1 item as compared to the next most chosen alternative. This was also true for experienced subjects. This indicates that subjects were uncertain about the foresight of their opponent, because if there is some chance that the opponent can have limited foresight, then that provides an extra incentive to remove just 1 item to keep the opponent as far away from the end of the game as possible and maximize the probability that he makes a mistake. This is discussed more explicitly in the model comparison sections.[12]

Result 5(a): *Choice proportions at losing positions.* At position 0 and 4, subjects were significantly more likely (with p-values of 0.001 and approximately 0, respectively) to choose to remove 1 item than next most likely alternative. This was also true for Exp subjects in the combined sub-session (see Table 3 below).

Table 3

Proportion of choices from losing positions

	Session					Exp in Combined Sub-session				
	Items Removed					Items Removed				
Position	1	2	3	p-value	Sample	1	2	3	p-value	Sample
0	41.1	29.3	29.6	0	1060	46.4	28.1	25.5	0	267
4	53.5	24.2	22.3	0	1011	49	28	23	0	253

Notes. The figures are in percentages. The p-values are two tailed p-values comparing the proportion of times 1 was chosen, as compared to the next most chosen alternative.

The next result, 6(a), tells us that it was ex-post weakly optimal for an Experienced subject to choose "First Mover" against an Inexperienced opponent when the Inexperienced subjects were "truly" inexperienced, i.e., in the first two rounds of the combined sub-session.

[12] The model with limited foresight and uncertainty about the opponent's foresight (Rampal (2016)) partially captures this: the prediction for the full foresight player's choice at position 0 is "remove 1 item." This is because removing 1 item keeps an opponent with a certain level of limited foresight far enough from the end of the game to make a mistake. But at position 4, any action brings the opponent with any foresight level too close to the end of the game to make a mistake, hence there is no such prediction. The Dynamic Level-k model predicts that the Level-1 player, who believes that her opponent is Level-0 (who uniformly randomizes among all available actions at each decision node), would be indifferent among removing 1 and 2 items only at position 4. However, there is no predicted bias towards removing 1 item.

Result 6(a): *Ex-post optimality of being "First Mover" against a "truly" Inexp opponent.*

In the first two rounds of the combined sub-session, when the Inexp opponent was "truly" inexperienced it was weakly ex-post optimal for an Exp player to choose F. This can be seen in Table 4 below. In rounds 13-14, the first two rounds of the combined sub-session, if an Exp subject was the selected first mover, his average earnings were weakly more than his average earnings if he were the selected second mover. However, for all later rounds, enough of the Inexp subjects had understood the perfect strategy so that it was *not* ex-post optimal for an Exp subject to choose F against an Inexp subject: for each round after round 14, the average earnings of the Exp subjects selected as the first mover against an Inexp opponent was significantly lesser that the average earnings of the Exp subjects selected as the second mover against an Inexp opponent. This finding agrees with the faster learning speed of the Inexp subjects (result 3(a)).

Table 4

Ex-post average earnings of the Exp subject against Inexp opponent, by round.

Round	Earning as Second Mover	Earning as First Mover	p-value
13-14	194 (16.5)	216 (31.6)	0.5
15-16	193 (15.8)	107 (23.2)	0
17-18	193 (11.5)	85 (17)	0
19-20	200 (7.3)	64 (11.4)	0

5 Experimental Design: Treatment 2

5.1 Within Game Learning Treatment

Treatment 2 comprised of 8 sessions. Each session contained between 8 and 18 subjects. A session lasted 64 minutes on average. The average payment made to the subjects was USD14.80. Treatment 2 was designed to ask the following question."Do Exp subjects endogenously learn about and respond to their opponent's experience-level by observing the opponent's prior play in the round, without being explicitly informed about the opponent's Exp/Inexp type?" Each session used three games. First, a game called the *Computer 13* game, henceforth shortened to C13. Second, the *Avoid Removing the* 13^{th} *Item* game, which we refer to as the *Human 13* game (henceforth, H13 for short). Last, a three period sequential bargaining game (Rubinstein (1982)) with a common discount factor of 0.6.

The rules of the H13 game are the same as the Avoid 9 game except that to begin, the total number of items in the box is 13, so the player who removes the 13th item loses, while his opponent wins. The H13 game also begins with both players choosing F/S, and either player's choice being selected with 50 percent chance each. In the H13 game, like in the Avoid 9 game, all prior choices in the game are displayed to both players, except at the simultaneous F/S decision stage. The payoffs of the H13 game were also the same as the Avoid 9 game: 500 ECUs for winning as the first mover 200 ECUs for winning as the second mover, and 50 ECUs for losing. We use 13 items instead of 9 to further widen the gap between subjects who understood the perfect strategy and those who did not understand it. Using 13 items also mean that this gap persisted for more rounds. These features help in distinguishing among competing theoretical models described later.

The winning positions in H13 are $W_{13} = \{1, 2, 3, 5, 6, 7, 9, 10, 11\}$, and losing positions are $L_{13} = \{0, 4, 8, 12\}$. In the H13 game, the second mover can always win. As the second mover, one can choose (4 minus the opponent's previous choice) at each move to remove the 12th item. However, if the second mover fails to put the opponent in L_{13} after one of his moves, then the first mover is guaranteed a win by playing a strategy which puts the second mover at a losing position at all of the second mover's subsequent moves. Note that in the H13 game, the SPNE is for both players to choose "Second Mover" (henceforth S) and then choose 3, 2, and 1 from positions {1, 5, 9} , {2, 6, 10}, and {3, 7, 11}, respectively. The SPNE strategy places no restriction on actions from a position in L_{13}. We call this SPNE strategy as the "perfect" strategy in H13.

The *perfect strategy* in H13:

$$perfect\ action\ = \begin{cases} S & at\, \{F/S\}\ decision\ stage \\ 3 & if\ \ Position \in \{1, 5, 9\} \\ 2 & if\ \ Position \in \{2, 6, 10\} \\ 1 & if\ \ Position \in \{3, 7, 11\} \\ Arbitrary & if\ \ Position \in \{0, 4, 8\} \end{cases} \tag{2}$$

Note that if one's F/S decision is implemented, the perfect strategy is a "sure-win" strategy in H13. However, due to the extra incentive to win as the first mover, the perfect strategy is *not* a weakly dominant strategy in H13. If a risk neutral rational player believes with probability at least $\frac{1}{3}$ that his opponent will make a mistake, then choosing "First Mover" (henceforth F) is optimal. Let any subgame with its root at a node in position 0 be labeled as $H13_{sub}$. Note that in $H13_{sub}$, the perfect strategy *is* weakly dominant because the payoff from winning/losing is already decided, and the perfect strategy is a "sure-win" strategy from a winning position.

The C13 (Computer 13) game was played by a human subject, individually, against a perfectly playing computer. In treatment 2, the opponent's performance in C13 is shown to the Exp player before the H13 game begins. The information about the opponent's performance in the C13 game is used to replace treatment 1's method of informing the Exp player if the opponent is Exp/Inexp. In the C13 game, the subjects were told that "the computer plays perfectly to win."[13]There are two key differences in the C13 game compared to the H13 game. First, in the C13 game, the human subject decides who the first mover will be: him or the computer. After the first mover is decided by the human subject, he and the computer move alternately, choosing 1, 2 or 3 items to be removed from a box containing 13 items. The player who removes the last item loses, and his opponent wins. The second key difference in the C13 game is that a player earns 50 ECUs for a loss, and 500 ECUs for a win. That is, there is no extra-incentive to win as the first mover (which is also not possible as the computer plays perfectly). Thus, in C13, it is weakly dominant and an SPNE strategy to play the perfect strategy for H13 described above.

The design of treatment 2 was as follows:

1. In each session each subject went through 2 sub-sessions. First, the subjects were split into two types: **Exp** and **Inexp**, with Inexp subjects being at least 50 percent of the total subjects in the session. In the first sub-session (training sub-session):

   (a) **Exp subjects trained**. Each round of this sub-session comprised of two parts. At the start of each round, each Exp player, for example X, was randomly matched with another Exp player, say Y. In the first part of the round, both X and Y separately played C13 against their respective

[13]The subjects were told that *"The computer plays perfectly to win. That is, if there is a choice or a sequence of choices that the computer can make to win against you, the computer will make that choice or sequence of choices without making any mistake."*

perfectly playing computer. In the second part of the round, X and Y played the H13 game with each other. In the training sub-session, Exp subjects played 8 rounds of C13-H13 among themselves, where each round had two parts: play the computer and then play the human opponent. The earning from a round was the sum of the earnings from its two parts. For C13, a subject was given 500ECUs for a win, and 50ECUs for a loss. For H13, a subject was given 500ECUs for winning as the first mover, 200ECUs for winning as the second mover, and 50ECUs for a loss. One round was randomly drawn as the round determining earning from the first sub-session.

- At two junctures in the C13 part of each round, X and Y were asked about the type (Exp or Inexp ?) of their opponent. First, after X made his F/S decision in C13, he was shown Y's F/S decision vs the computer and then (this is the first juncture) X was asked about the type of his opponent. Second, X was shown Y's actions and outcome versus the computer in Y's C13 and then (this is the second juncture) X was asked about the type of his opponent. For each correct answer about the opponent's type, 100 ECUs were added to the total earnings of the round for X. Similarly, Y was also shown X's history of moves versus the computer in C13 and asked about the X's type at two junctures. Note that in this training sub session, with only Exp subjects playing each other, the correct answer about the opponent's type was always Exp. The Exp subjects were informed that the Inexp subjects were playing a different game, so in the training sub-session these answers were a source of "free money".[14]

(b) **Inexp** subjects were not told about the H13 or C13 games. They played between 5-8 rounds of a three period bargaining game with the common discount factor = 0.6 (Rubinstein (1982)) among themselves. One round was randomly drawn as the round determining earning from the first sub-session.

[14]Despite this information, the answer about the opponent's type was wrong in 12.76% of the cases at the first juncture, and 7.55% of the cases at the second juncture. These error percentages increased if the opponent lost his C13 part, which might have led a subject to believe that his opponent is an Inexp type. This confusion may be stemming from the fact that the question about opponent's type seemed counter-intuitive given the information provided that the opponent cannot be Inexp. Further, the "free money" being offered to the subjects may also have made them suspicious. Several subjects asked the experimenter "how can my opponent be Inexp?" We stuck to this design to keep the rounds the same for Exp subjects in both sub-sessions of the experiment.

2. In the second sub-session (combined sub-session) Exp and Inexp subjects were mixed and were "absolute stranger" matched into pairs before every round. They played 6 rounds as described in 1(a), i.e., play the C13 with one's respective perfectly playing computer and then play the H13 game with the human opponent (C13-H13), with two modifications. First, any pair of types, i.e., (Inexp, Inexp), (Exp, Exp) or (Exp, Inexp) was possible. Second, Inexp subjects, unlike the Exp subjects, were *not shown* the history of their opponent's moves versus the computer in C13 or the outcome of the opponent in the opponent's C13 game versus the computer. Notably the Exp subjects were told the fact that the Inexp subjects would not be shown the opponent's history or outcome of play in the C13 part of any round.[15] The Inexp subjects were also not asked about the type of their opponent. Exp subjects *were* asked about their opponent's type like in 1(a), however now the answer could be either Exp or Inexp. The reason for not showing the opponent's history from C13 to the Inexp subjects was that we wanted the gap of understanding of the perfect strategy between the Inexp and Exp subjects to persist for as many rounds as possible in the combined sub-session.

3. One round from the second sub-session was drawn at random for payment and added to the payment due from the first sub-session. The conversion rate used in the training and combined sub-sessions was 120ECUs=1USD. The payment due to the Inexp subjects at the end of the two sub-sessions was paid to the Inexp subjects who then exited the experiment.

4. After the end of the two sub-sessions, Exp subjects participated in a risky choice study using the DOSE procedure (Wang, Filiba, and Camerer (2010)). Exp subjects made 20 binary choices (see figure in the attached subject instructions), where each choice was between a lottery and a deterministic payment. The earnings/losses from the DOSE procedure were scaled by a factor of $\frac{1}{5}$. One of the 20 choices was selected at random and implemented. The resultant earning/loss was added/subtracted to the earning due to the Exp subjects from the training and combined sub-sessions. The Exp subjects were then paid, which concluded the experimental session.

Subjects moved sequentially in the training and combined sub-sessions. Each subject was given a clock with 30 seconds at each move to remind them to decide. The clock could only flash if the time taken was

[15]In fact all aspects of the design were common knowledge to all subjects with only one exception: the Inexp and Exp subjects did not know what games the Exp and Inexp, respectively, were playing in the training sub-session. They did know that these games were unrelated.

more than 30 seconds, and the game did not proceed without the subject's choice. Note that the SPNE prediction of the C13-H13 game (combination of the two games) is that outcomes in C13 should not affect the outcomes in H13. In particular, both players should play the perfect strategy in C13 and then again in H13.

The key difference in treatment 2 as opposed to treatment 1 is that in treatment 1, in the combined sub-session, where Exp and Inexp subjects were mixed, each player was told if their opponent is Exp or Inexp. In treatment 2, only the Exp subjects are shown the history of their opponent's moves versus the computer, but not told their opponent's type.

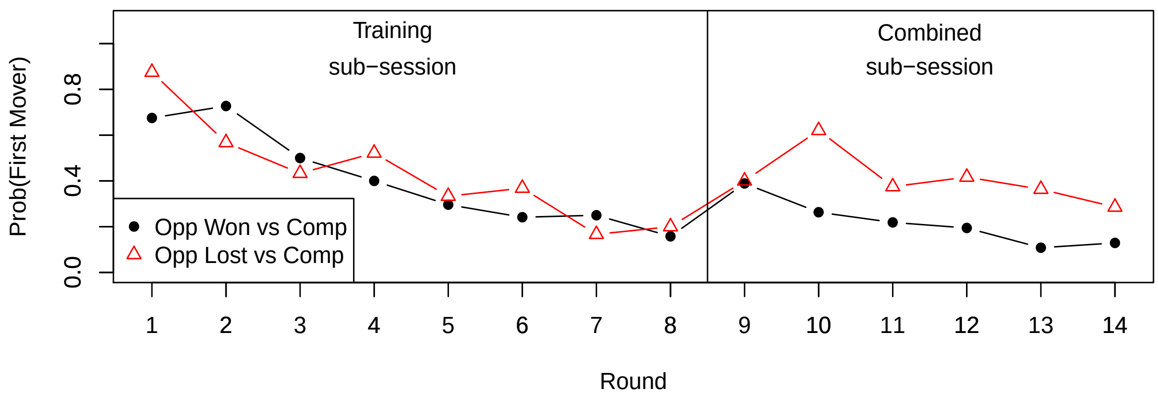
6 Data Results for Treatment 2

As per the analysis for treatment 1, we shall again focus on Exp subjects' "First/Second Mover" choices and patterns of imperfect play.

Result 1(b): *Training successful:* In the training sub-session (rounds 1-8) the proportion of Exp subjects winning the C13 game against the computer increased significantly. The round 1 proportion was 17 percent, while the round 8 proportion was 79 percent, which is a highly significant increase with a p-value of approximately 0. Recall that the computer plays perfectly, and any deviation from the perfect strategy by the human player leads to a loss. Further, in $H13_{sub}$, the proportion of Exp subjects playing perfectly as the selected second movers increased significantly in round 8 (75%) compared to round 1 (25%). This increase is also highly significant with a p-value of approximately 0.[16]

[16] To study the success/failure of training, we focus on the play of the Exp subjects in (i) the C13 part and (ii) $H13_{sub}$. This is because in both (i) and (ii), playing perfectly is a weakly dominant strategy. In contrast, the optimal choice among F/S in the H13 game is also affected by the belief about the opponent's probability of playing imperfectly in the subsequent stages of the H13 game. Despite this interpretation problem, if we look at the F/S choice data by opponent's outcome we have the following results. In rounds 1 and 2, the Exp subjects chose S in 34.4% out of 96 cases, while the same for rounds 7 and 8 was 80.2%. The difference in proportion of S choices is highly significant with a p-value of approximately 0. Further, in the training sub-session, F/S choice in the H13 game weren't affected by the opponent's outcome in C13 (discussed below).

Figure 3: Experienced Subjects' Behavior by Opponent's Performance



Notes. The figure depicts the round-wise proportion of experienced subjects who chose "First Mover" when matched with an opponent who won C13 (black dots) and the proportion of experienced subjects who chose "First Mover" when matched with an opponent who lost C13 (red triangles). The first mover loses if the second mover plays perfectly. The latter proportion is significantly higher, with a p-value < 0.05, for rounds 10 and 13. This shows that the experienced subjects were more likely to risk losing to try for the higher prize when matched with an opponent who lost C13. The training sub-session was the first 8 rounds, and the combined sub-session was rounds 9-14.

Result 2(b): *Opponent's loss versus the computer has a significant effect on the Exp subjects' behavior in the combined sub-session* : In the combined sub-sessions (rounds 9-14), in the H13 game, Exp subjects were more likely to choose F against an opponent who lost his C13 game against the computer, as opposed to an opponent who won his C13 game.[17]

Figure 3 shows that the Exp subjects whose opponent lost in C13 (red triangles) chose F at a significantly higher rate than the Exp subjects whose opponent won C13 (black dots). The p-values of this difference in proportions is less than 0.05 for rounds 10, and 13, and the one-tailed p-value of this difference is 0.062 for round 12 (see Table 11 in the Appendix).

Next, consider Table 5 which reports the probit estimation results for the data from the H13 games in the combined sub-session. The dependent variable is the probability that a subject chooses F in the F/S decision stage of H13 in the combined sub-session. The definition of the dependent variable implies that a positive coefficient on an independent variable means that a higher value of the independent variable increased the probability of the player choosing F. Estimations (1)-(3) have the Exp player's F/S choice

[17]We don't observe any such significant difference in the training sub-session. The knowledge that one's opponent is similarly experienced, may have led the Exp subjects to not put too much weight on the opponent's performance in the C13 part of one round that the subject observes. Further, learning the perfect strategy is one skill, learning how to benefit from the opponent's lack of understanding of the game is another skill. The latter might have taken additional practice to acquire.

in H13 as the dependent variable. The highly significant coefficient of "Opponent Lost C13" dummy in the probit estimations (1) and (2) implies that, in the combined sub-session, if the average Exp subject observed that his opponent in H13 had lost against the computer in the C13 part of that round, then he was significantly more likely to choose F in the F/S decision stage of H13.[18] The coefficient of the same variable is statistically insignificant for the Inexp subjects (estimation (4)), which is expected because Inexp subjects were not shown the moves or outcome of the opponent in the opponent's C13 part of the round. The round variable is negative for all the models, (1)-(4), because an increase in the round variable captures the learning about how to play the perfect strategy, and the increase in the belief that the opponent understands the perfect strategy, which makes choosing S a worse option. In model (2) we also add the loss aversion and risk aversion parameters as measured using the DOSE (Wang et al (2010)) estimation technique.[19]

In model (2) we find that both risk aversion and loss aversion have a negative effect on the probability that an Exp player chooses F. The coefficient of risk aversion is statistically insignificant, but the coefficient of loss aversion is significant with a p-value of 0.02. The sign of these coefficients is as expected. For the Exp subject, conditional on his F/S choice being selected, choosing F entails a gamble between winning (500ECUs) or losing (50ECUs) while choosing S guarantees a sure win worth 200ECUs.[20]

[18] The two tailed p-value of this coefficient in model (1) and (2) is 0.005 and 0.004 respectively. For the model (1), if we replace the dummy for "Opponent lost C13" by a dummy for the "Opponent is Inexp", the coefficient for that latter equals to 0.304 (robust standard error 0.15), which is significant with a p-value < 0.05. Recall that the opponent's experience level is not known to any subject, and this result reflects that Exp subjects were able to spot Inexp opponents just by their play against the computer. We discuss this further below.

[19] Recall from the design that only Exp subjects participated in the risky choice study. Only Exp subjects participated in the risky choice study because of budgetary constraints and because we are more interested in how and why players who understand the perfect strategy (most of the Exp players) respond to information about the opponent's performance in the C13 part of the round.

[20] Assuming the expected utility framework, the sign of the risk aversion is expected to be negative because in choosing F an Exp subject is choosing a lottery where, conditional on his choice being implemented, and according to his subjective belief, there is some probability that his opponent will play imperfectly as the selected second mover in $H13_{sub}$, and give him 500ECUs for the H13 part of the round, and with remaining probability his opponent will play perfectly, and give him 50ECUs for the H13 part of the round. Thus the greater a subject's risk aversion, the lesser likelihood of him not taking the 200ECUs for sure (conditional on his choice being selected), and choosing a lottery with an expected payment dependent on his subjective belief. The loss aversion term captures how much more a subject values losses compared to gains. In our design even if a player loses in H13, he still receives 50ECUs. But if an Exp player uses the 200ECUs he could have received by winning as a reference point (Koszegi and Rabin (2006) for example use rational expectations based reference points), and considers the reduction in earnings by 150ECUs due to choosing F as a "loss", one would expect the significant decrease in the propensity to choose F due to a higher loss aversion term, as observed in model (2) and (3). The sign of the risk aversion coefficient is the opposite of expected in model (3), but the coefficient is again statistically insignificant.

Table 5

Factors Influencing Probability of choosing “First Mover”

Dependent Variable: Prob(First Mover)

	Exp	Exp	Exp	Inexp
	(1)	(2)	(3)	(4)
Opponent Lost C13	0.58***(0.2)	0.59***(0.21)	0.15(0.26)	-0.19(0.18)
Round	-0.14***(0.05)	-0.17***(0.05)	-0.16***(0.05)	-0.40***(0.05)
Risk Aversion		-0.14(0.53)	0.17(0.54)	
Loss Aversion		-0.5***(0.22)	-0.67***(0.25)	
High			-1.32***(0.37)	
High*Opponent Lost C13			1.04**(0.46)	
Constant	0.53(0.67)	2.04(1.32)	3.13***(1.41)	4.54***(0.78)
No. of Obs.	278	272	272	312
Pseudo R^2	0.1368	0.1940	0.2586	0.1793
Session Dummies	Yes	Yes	Yes	Yes

Notes: These results are for rounds 9-14 of 8 sessions. The errors were clustered by subject. Probit (1) reports the results for 48 Exp subjects, while probits (2) and (3) report the results for 47 Exp subjects (risk aversion and loss aversion data for 1 Exp subject was lost). Probit (4) reports the results for 54 Inexp subjects.

The result that Exp players were more likely to choose F in the H13 part of the round when they observed that the opponent lost in C13 is driven by the more “skilled” Exp players. To verify this, we create a dummy variable, $High$, which takes value 1 if an Exp subject won the C13 game in at least 3 of the last 4 rounds of the training sub session (rounds 5-8). The $High$ variable indicates that the Exp subject has a high level of understanding of the perfect strategy for the H13 and C13 games. We add the $High$ variable and an interaction term of *High*Opponent Lost in C13* to the model (2) and estimate model (3). The coefficient of $High$ is negative and significant (p-value of 0.003), which means that given that the opponent won the C13 part of the round, being the *High* Exp player significantly reduced the probability of choosing F, which agrees with the definition of $High$. The coefficient of *High*Opponent Lost C13* is positive and significant (p-value of 0.029). This means that being a *High* Exp type who faced an opponent who lost in the C13 part of the average round of the combined sub-session, significantly increased a subject’s likelihood of choosing F in that round.

Results 3(b), 4(b), 5(b) and 6(b) approximately replicate the findings from the treatment 1 results 3(a), 4(a), 5(a), and 6(a), respectively. That is, in treatment 2, we find the following: (i) the learning speed of Inexp subjects was faster; (ii) the rate of imperfect play declined significantly closer to the end of the H13 game; (iii) subjects, including trained Exp subjects, were more likely to remove 1 item from losing positions in the H13 game; and (iv) in the combined sub-session, it was ex-post optimal for an Exp subject to be the first mover in H13 if the opponent lost C13. Therefore, these results are described in the Appendix.[21]

Result 7: *Updating opponent's type:* In the combined sub-session (rounds 9-14), Exp subjects' accuracy while answering the questions about their opponent's Exp/Inexp type improved after they observed more actions of their opponent. When asked the opponent's type at the first juncture, i.e., immediately after observing the opponent's first/second mover decision against the computer in C13, the accuracy was 45 percent. At the second juncture, immediately after observing the opponent's history of play and the consequent outcome versus the computer in the C13 part, the accuracy increased to 62 percent. The difference in proportions is highly significant with a p-value of approximately 0.

7 Model Comparison

In this section, we use the Maximum Likelihood Estimation method to distinguish among which theoretical explanation fits our data the best. The first of the three theoretical models we consider is "Limited Foresight and Learning Equilibrium" (Rampal (2016)), henceforth LFLE. The LFLE model allows for players to have different levels of limited foresight (foresight implies the number of subsequent stages one can observe from a given decision stage) and uncertainty about the opponent's foresight level. The uncertainty about the opponent's foresight level implies that the LFLE also entails belief updating about the opponent's foresight level within the play of a game.

The second model we consider is the "Quantal Response Equilibria for extensive form games" (henceforth AQRE) by McKelvey and Palfrey (1998). This model posits that each player in the game makes

[21]The only notable difference is in results 6(a) and 6(b). In particular, for treatment 2, in the whole combined sub-session (rounds 9-14), when the opponent lost C13, the average earnings of the Exp player was weakly more when the Exp player was the selected first mover in H13 compared to when he was the selected second mover. In contrast, the ex-post optimality of an Exp player being the selected first mover against an Inexp opponent in the Avoid 9 game was only true for the first two rounds of treatment 1.

errors in assessing his expected payoff from the actions following any given decision node. This implies that players best respond to a given strategy profile with errors and play totally mixed strategies. The AQRE model typically assumes a logit error structure which implies that the higher the expected payoff of an action relative to other actions, the higher the probability of that action being chosen.

The third model we consider is "A Dynamic Level-k Model for Sequential Games" by Ho and Su (2013). This model posits that players are rational, however, they have subjective beliefs about the "level" of rationality of the opponent. Thus, players choose strategies that are optimal given their beliefs about the opponent's "level", but these strategies may not be perfect strategies.

The key difference among the Dynamic Level-k and LFLE models with respect to the data is the following. The LFLE explains the pattern of imperfect play observed (results 4(a) and 4(b): the rate of imperfect play declined as the number of items left reduced) using limited foresight; in contrast, the Dynamic Level-k model, which assumes rationality and relies solely on subjective beliefs to explain the data, does not allow for imperfect play in $A9_{sub}$, $H13_{sub}$ or C13. This is because playing imperfectly after the first/second movers have been decided is a weakly dominated strategy, which implies that no subjective belief about the opponent's "level" can justify such a strategy being chosen. The AQRE model is also not able to simultaneously match the pattern of F/S decisions and the proportion of imperfect play we observe in $A9_{sub}$, $H13_{sub}$ and C13. This is because the expected payoff from perfect actions, which lead to a win, is higher relative to imperfect actions in $A9_{sub}$, $H13_{sub}$ and C13. Therefore, the predicted probability with which imperfect actions ought to be chosen according to the AQRE model is much lower compared to the data. Thus, we find that the LFLE model performs better than the Dynamic Level-k model and the AQRE model in the MLE comparison.

7.1 The Limited Foresight and Learning Equilibrium

The LFLE is a limited foresight equilibrium concept (see *Limited Foresight and Learning Equilibrium* (Rampal (2016)) for a detailed theoretical discussion). The LFLE is applied to finite dynamic games with perfect information.[22] Fix an arbitrary perfect information game G. Given G, and the parameters of the LFLE model, the LFLE generates outcome predictions for G which are testable against the data on G.

[22] Avoid 9 and Human 13 games have one stage of imperfect information, i.e., the simultaneous moves for the F/S decision. Extending the LFLE model to account for this is straightforward.

The LFLE model incorporates two key features: limited foresight and uncertainty about the opponent's foresight. To model the uncertainty, the LFLE model converts G into a standard Bayesian game of imperfect information, say Γ, where each player in G can be one of a set of possible types. The limited foresight feature is captured by the definition of a type. The type of a player denotes his foresight-level. In a multi-stage game, a player's foresight level (or type) is equal to the number of *subsequent* stages that that player can observe from any given move. For example, in the Avoid 9 game, a player with foresight level 1 (or a player with type-1), can observe the decision nodes and action sets in the current stage and the next stage, but cannot observe what the subsequent stages are going to be.

For simplicity, the prior joint distribution on types is assumed to be common knowledge. We allow for one exception to this assumption, which is discussed below. Each player's own type is private information. Given the common prior, and the game G, the LFLE provides a strategy and belief profile for the resulting Γ. The LFLE strategy profile and the common prior imply a distribution over the outcomes observed when testing G. Note that Γ nests G as a special case where all players have "full foresight" with probability 1. To solve for the LFLE where players have limited foresight, we have to use curtailed versions of Γ. This is because if a limited foresight type does in fact have limited foresight, then he must be choosing optimal decisions on the basis of a curtailed version of Γ. We use the Appendix to explain the procedure to calculate LFLE in the games used in our experiment. We now describe the LFLE of the Avoid 9, Computer 13, and Human 13 games described above.

7.1.1 LFLE of the Avoid 9 Game

The free parameter of the LFLE model is the distribution over foresight levels. For simplicity, we model four possible levels of foresight (four types) for each player. Foresight levels 0, 1, 2, and the ex-ante full foresight type, denoted by f. Let (p_0, p_1, p_2, p_f) be the common knowledge prior distribution over the foresight levels (or types). We describe the solution of the LFLE below. As a tie-breaker, we state the LFLE where every player-type plays actions that give equal payoff with equal probability. At any given decision node, let a *Position*, denoted by P, be a particular sum of items removed in all preceding nodes. Let a *perfect* strategy imply the actions $a_4^* = (4 - P)$ for $0 < P < 4$; $a_8^* = (8 - P)$ for $4 < P < 8$, and a uniform distribution over available actions at $P \in \{0, 4, 8\}$. Recall that the perfect strategy is the weakly dominant SPNE strategy in $A9_{sub}$.

LFLE of the Avoid 9 Game:

- The foresight level 0 type randomizes uniformly among available actions at the F/S decision stage and for $P < 5$; at $P \geq 5$, he plays perfectly. His beliefs don't affect strategy.
- The foresight level 1 type chooses F at the F/S decision stage. He randomizes uniformly among available actions for $P < 2$; at $P = 2$, he randomizes uniformly among removing 1 or 2 items. For $P > 2$, he plays perfectly. His beliefs don't affect strategy.
- The foresight level 2 type chooses F at the F/S decision stage. He plays perfectly in $A9_{sub}$, except that he removes 3 items at $P = 0$ if he has a positive prior belief on a type-0 opponent, i.e., if $p_0 > 0$.
- The full foresight type chooses F at the F/S decision stage iff $(p_0 + p_1) \geq 0.5$, else, he chooses S. He removes 1 item at $P = 0$ if he has a positive prior belief that the opponent has foresight level 1, i.e., if $p_1 > 0$. For $P > 0$, he plays perfectly in $A9_{sub}$.

We now explain the LFLE given above. At any move of the **type-0** (foresight level of 0) player in the Avoid 9 game, he cannot observe the stages of the game after his action at that move. But from a given move, say at the F/S decision stage, to provide a basis to choose among the available actions, F or S in this case, the type-0 player must observe payoffs after each of those actions. But the Avoid 9 game doesn't have a terminal node after the F/S decision stage. Therefore, if an action taken at a limited foresight player's foresight horizon does not lead to a terminal node of the underlying game G, our limited foresight model needs to create a synthetic payoff profile for that action. We call this synthetic payoff profile as the *curtailed payoff profile*, as it is constructed after curtailing the underlying game after the foresight horizon of a limited foresight player. We use the following $\frac{(min+max)}{2}$ *curtailment rule* used in Rampal (2016) and Mantovani (2014): from each action at the foresight horizon of each limited foresight type, the *curtailed payoff profile* observed by him provides each player with $\frac{(min+max)}{2}$ of the set of payoffs possible for that player from that action. Applying this rule at all possible moves of the type-0 player, one can understand the LFLE strategy of the type-0 player. At the first stage, both choices, F and S, can lead to him being the first or the second mover, based on the opponent and Nature's decision; further, he can win or lose in either case in the subsequent stages. Thus his *curtailed payoff* after both F and S is $\frac{500+50}{2}$, which, given our tie-breaking assumption, means that the type-0 player chooses F or S with probability $\frac{1}{2}$ each. Next, note that for any $P < 5$, for any action from that P, the set of possible payoffs is the same, and therefore the curtailed payoff profile is the same: (275,125), where we write the payoff of the selected first mover (second mover) first (second). This implies that the type-0 player randomizes among available actions if $P < 5$. But from $P \in \{5, 6, 7\}$, a_8^* gives a curtailed payoff equal to the winner's payoff. Thus, from any $P \geq 5$, the type-0 player plays perfectly. Notice that the opponent's type does not matter for the type-0's behavior, each of his moves is the last move of the "curtailed game" he observes from that move.

Proceeding similarly we can solve for the LFLE actions of each foresight level at each position. This gives us a LFLE strategy for each of the four foresight levels we model here. The details of how we solve for the LFLE are provided in the Appendix. It is worth noting here that the F/S decision of only the full foresight type (type-f) player depends on his prior beliefs. That is, if the type-f player believes with high enough probability that his opponent has a foresight level of 0 or 1, then he chooses F to maximize expected payoff (assuming risk neutrality). Otherwise he chooses S. Prior beliefs of each player-type are assumed to be the same as the common knowledge prior distribution over the foresight levels in all but the following case. When applying the LFLE model over the data from (Exp, Inexp) pairs, we use MLE to estimate a separate distribution $\mathbf{p_{Exp}}$ for the Exp players, and a separate distribution for the Inexp players $\mathbf{p_{Inexp}}$. This is because either type was playing an opponent of a different type, and the players were informed about it. In particular, in $\mathbf{p_{Exp}}$, we allow the Exp players to be only foresight level-0 or full foresight type. Further, we add a parameter $\delta \in [0, 1]$ which captures the proportion of type-f Exp subjects who have a subjective belief that his Inexp opponent is foresight level-0 or 1 with probability at least 0.5, and therefore choose F in the F/S decision stage. Therefore the LFLE strategy of the type-f Exp player in an (Exp, Inexp) pair becomes: choose F with probability δ, and S with probability $(1 - \delta)$; subsequently, play as per the LFLE strategy given above. Although LFLE uses beliefs based on the common-knowledge prior distribution, we have to use subjective beliefs here because the Exp subjects observe that the opponent is from a different Inexp population.[23]

[23] Our conclusions do not change even if we estimate only a single common knowledge prior distribution over the foresight levels for the (Exp, Inexp) groups and impose subjective prior beliefs to be the same as the prior distribution over foresight levels that we estimate using maximum likelihood.

7.1.2 LFLE of Human 13 and Computer 13 Games.

For the game in treatment 2, C13-H13, for simplicity, we model only four foresight levels: foresight levels 0, 1, 4 and the ex-ante full foresight player, denoted by f. Let $\mathbf{p} = (p_0, p_{1,} p_4, p_f)$ be the common knowledge prior distribution on these foresight levels. We solve for the LFLE of H13 and C13 separately. This is because after C13 ends, at the beginning of H13, except the type-f player, the beliefs of the other types do not affect the LFLE in H13. Therefore the history of moves in C13 does not matter for the types other than f. For the type-f player, let the updated belief on the four possible types of the opponent be $\mathbf{u} = (u_0, u_1, u_4, u_f)$. Note that for the Inexp subjects, who get not information about the opponent's play in C13, $\mathbf{p} = \mathbf{u}$. The LFLE strategy profile for the **H13** game and the **C13** game is given below. The details are given in the Appendix. Let a *perfect* strategy imply the actions $a_4^* = (4 - P)$ for $0 < P < 4$; $a_8^* = (8 - P)$ for $4 < P < 8$; $a_{12}^* = (12 - P)$ for $8 < P < 12$, and a uniform distribution over available actions at $P \in \{0, 4, 8, 12\}$.

LFLE of the Human 13 Game:

- The foresight level 0 type randomizes uniformly among available actions at the F/S decision stage and for $P < 9$; at $P \geq 9$, he plays perfectly. His beliefs don't affect strategy.
- The foresight level 1 type chooses F at the F/S decision stage. He randomizes uniformly among available actions for $P < 6$; at $P = 6$, he randomizes uniformly among removing 1 or 2 items. For $P > 6$, he plays perfectly. His beliefs don't affect strategy.
- The foresight level 4 type chooses F at the F/S decision stage. He plays perfectly in $H13_{sub}$. He removes 1 item from $P = 4$ if he has a positive belief on the opponent being type-1, i.e., if $u_1 > 0$.
- The full foresight type chooses F at the F/S decision stage iff $(u_0 + u_1) \geq 0.5$, else, he chooses S. He plays perfectly in $H13_{sub}$. He removes 1 item at $P = 4$ if he has a positive prior belief on the opponent being type-1, i.e., if $u_1 > 0$.

LFLE of C13: The LFLE of C13 is identical to the LFLE of H13 except that: (a) The types with foresight levels 1 and 4 randomize uniformly among $\{F, S\}$ at the F/S stage, while the full foresight type chooses S at this stage; (b) the full foresight type and the foresight level 4 type randomize uniformly among removing 1, 2 or 3 items even when $P = 4$ in the C13 game.

Maximum Likelihood Estimation using LFLE. The data analysis using MLE is done separately for the two treatments. For each treatment, an observation is the observed choices of a pair of subjects in a round. Consider an observation from treatment 1. Denote the observation i by o_i. Suppose o_i can be observed if the strategy profiles in $S^2(o_i) \ni (s_1, s_2)$ are played, where the selected first mover chooses the strategy s_1, and the selected second mover chooses the strategy s_2.[24] Then the likelihood of o_i can be calculated using the LFLE of the Avoid 9 game and the prior distribution over the foresight levels: (p_0, p_1, p_2, p_f).

$$Prob(o_i) = \sum_{(s_1,s_2)\in S^2(o_i)} [\sum_{j,k\in\{0,1,2,f\}} Prob(s_1|\text{foresight level}\, j,\, \mathbf{p})p_j Prob(s_2|\text{foresight level}\, k,\, \mathbf{p})p_k] \qquad (3)$$

Where $Prob(s_1|\text{foresight level}\, j)$ can be calculated using the LFLE of the Avoid 9 game. In some (1.4%) of the observations in treatment 1's combined sub-session we find that subjects play imperfectly at position 5 or more (4 or lesser items left). Such outcomes have 0 probability according to LFLE as even the foresight level 0 plays perfectly at position 5 or more. In treatment 2's combined sub-session, there is no such observation. To make the likelihood function finite, we deal with the 1.4% outcomes where $Prob(o_i)$ equals to 0 by using the "uniform error rate" ϵ used by Ho and Su (2013) (the Dynamic Level-k model discussed below) and Costa-Gomes et al (2001), among others. Let the number of possible outcomes be N. Then $\epsilon \in (0, \frac{1}{N})$ denotes the error probability that each of the possible N outcomes will occur. With remaining probability $[1 - \epsilon N]$, the model's prediction holds. Then $Prob^{LFLE}(o_i)$ in application to the data becomes:

$$Prob^{LFLE}(o_i) = \epsilon + (1 - \epsilon N)Prob(o_i) \qquad (4)$$

Suppose each observation o_i in treatment 1 has frequency f_i in the data. In treatment 1, N (the total

[24] A strategy maps each history of play to an action. However, in the Avoid 9 game, beliefs play a role in determining the LFLE actions at only three information sets: (i) the F/S decision of the type-f; (ii) & (iii): the choice of types 2 and f at $P = 0$. We take this into account in the LFLE stated above for Avoid 9, H13 and C13 games.

number of different outcomes possible) is 447. Then the log likelihood function for the LFLE model in treatment 1 is given by:

$$LogL(LFLE|\mathbf{p}) = \sum_{i=1}^{447} f_i log(Prob^{LFLE}(o_i)) \tag{5}$$

We can similarly calculate the log likelihood for treatment 2 data using the LFLE of the H13 and C13 games, and $\mathbf{p_{T2}} = (p_0, p_1, p_4, p_f)$. In general, maximizing the likelihood using LFLE implies searching over $(\mathbf{p}, \epsilon)$ to maximize (5). In treatment 1, we estimate one $\mathbf{p}$ for observations with (Exp, Exp) groups and another $\mathbf{p}'$ for observations with (Inexp, Inexp). However, for the (Exp, Inexp) pairs, as discussed above, we estimate a separate distribution $\mathbf{p_{Exp}}$ for the Exp players, a separate distribution $\mathbf{p_{Inexp}}$ for the Inexp players, and the parameter $\delta \in [0,1]$ which is the probability that a type-f Exp player chooses F in the F/S stage (and plays perfectly in the subsequent stages).[25] We estimate only a single error term ϵ for all treatment 1 data.

In treatment 2's combined sub-session, we don't include any error term as no 0-probability outcome was observed. We estimate only one $\mathbf{p}_{T2}$ for the treatment 2 data as no information about the opponent's Exp/Inexp status was provided to any subject. According to the LFLE model, at the beginning of the H13 part, the distribution on the foresight levels changes based on the outcome in C13. Table 16 in the Appendix describes the change in distribution on the foresight levels due to the outcome of win (W) or loss (L) in the C13 part of a round. According to the LFLE model, the Exp players observe these outcomes and update their belief. However, beliefs change the subsequent strategy in H13 only for the full foresight type (type-f) Exp player who wins in his own C13 part. Intuitively, when an Exp subject observes that the opponent lost (won) C13, he updates and puts more (less) weight on the opponent being foresight level-0 or 1. Therefore, given our estimate of the prior distribution over foresight levels, if the opponent loses (wins), then the full foresight type player chooses F (S) in H13.

7.2 Ho and Su (2013): The Dynamic Level-k model in Sequential Move Games

The Ho and Su (2013) (henceforth HS) model applies the Level-k model to sequential move games. The version of level-k they apply has the following properties. The level-0 strategy is to randomize uniformly

[25]Recall that the rationale behind this is that we allow δ proportion of type-f Exp players in (Exp, Inexp) pairs to have the subjective belief that the probability that the Inexp opponent's foresight level is 0 or 1 is more than $\frac{1}{2}$.

among the available actions at each move. The Level-1 strategy is a sequentially rational best response to the Level-0 strategy at every move. The Level-2 strategy is a sequentially rational best response to the Level-1 strategy at every move, and so on. Given these Level strategies, *rational agents* pick what they deem is the optimal strategy given their subjective belief about their opponent's Level.

In their words: "A player i chooses the optimal rule L_{k^*} in round $(t+1)$ from the rule hierarchy $\{L_0, L_1, ..., L_S\}$ based on belief $B^i(t)$ to maximize expected payoffs." This modeling implies that the probability of level-0 being chosen in a game is 0. This assumption of theirs agrees with the findings of Costa-Gomes and Crawford (2006) and Crawford and Iriberri (2007a 2007b). In particular, they state that "Because all players best respond given their beliefs, L_0 will not be chosen by any player and only occurs in the minds of the higher-level players."

In the Avoid 9 and H13 (Human 13) games: (i) the Level-0 strategy is to randomize uniformly among all available actions at all moves; (ii) therefore the Level-1 strategy is to choose F and subsequently play perfectly; (iii) the Level-≥ 2 strategies are all the same: choose S and play perfectly. In the C13 (Computer 13) game: (i) the Level-0 strategy leads to a loss with probability 53/54, due to uniform randomization among available actions at each move; (ii) the Level-1 strategy is to choose S and then play perfectly, as there is no extra incentive to win as the first mover, and even the level-0 player plays perfectly as the second mover with probability 1/27. (iii) The Level-2 strategy is to choose S and then play perfectly, as the Level-1 strategy implies perfect play. We argue that in C13, no subject could have had subjective beliefs that the computer is following the level-0 strategy because the subjects were told that: *"the computer plays perfectly to win. That is, if there is a choice or a sequence of choices that the computer can make to win against you, the computer will make that choice or sequence of choices without making any mistake."* Given that all players, including those who choose Level-1, are assumed to be rational, and therefore know how to play the weakly dominant perfect strategy, i.e., they know how to "play perfectly to win," it is implausible that some of these players would form subjective beliefs that the computer randomizes among available actions at any given decision node.[26]

Thus, a distribution over Level-1 and Level-2 strategies generates a distribution over the outcomes of the Avoid 9 game in treatment 1, and the C13-H13 round in treatment 2. We search for the distributions

[26]Our treatment 2 MLE results for the comparison with the Ho and Su (2013) model of dynamic Level-k are robust to the scenario where subjects still chose the Level-1 strategy, and believed that the computer plays the level-0 strategy.

that maximize the likelihood of the treatment 1 and treatment 2 data separately. Note that we need only the Level-1 and Level-2 strategies because all strategies above Level-2 are the same: choose S and play perfectly. The fact that the level-k model assumes that the Level-0 does not exist, implies that the Level-k model puts 0 probability on several observations of the data.

Recall that an observation is the observed choices of a pair of subjects in a round. Because of the level strategies described above, in treatment 1 or treatment 2, any observation where the selected second mover fails to play perfectly in $A9_{sub}$ or $H13_{sub}$ gets 0 probability according to the HS model. Further, any observation in treatment 2 where one of the subjects loses in C13 also gets 0 probability. They tackle this 0 probability problem by introducing an error probability, ϵ, which is the *minimum probability* of each possible outcome (a sequence of choices of a pair of subjects) that can occur given the game being tested. Consider again the outcome o_i which can be observed only if the selected first mover and the selected second mover choose strategies s_1, s_2 respectively such that $(s_1, s_2) \in S^2(o_i)$. Let the probability that players choose the L_1 strategy be $P(L_1)$, and $P(L_2) = 1 - P(L_1)$. If there are a total of N (for example $N = 447$ in the Avoid 9 game) possible choice pairs, then the Level-k model's probability that outcome o_i occurs is:

$$Prob^{L_k}(o_i) = \epsilon + (1 - \epsilon N) \sum_{(s_1,s_2)\in S^2(o_i)} [\sum_{j,k\in\{1,2\}} Prob(s_1|L_j)P(L_j)Prob(s_2|L_k)P(L_k)] \quad (6)$$

Then the likelihood function of the dynamic Level-k model is given by:

$$LogL(Level_k|P(L_1)) = \sum_{i=1}^{N} f_i log(Prob^{L_k}(o_i)) \quad (7)$$

Thus maximizing the likelihood implies searching over $P(L_1)$ and the error term ϵ. The HS (2013) Level-k model allows for belief $B^i(t)$ to change as a function of rounds (captured by t), based on either past experiences of a player (their "closed-loop" approach) or based on mental simulation of all possible experiences (their "open-loop" approach) by a player. They do not allow beliefs to change within a round. As beliefs change so does the distribution over the chosen levels. But it is the proportions of the chosen levels, $P(L_1)$, $P(L_{\geq 2})$, that generates the distribution over observed outcomes (as $P(L_{\geq 2}) = 1 - P(L_1)$, the $P(L_1)$ estimate captures the level distribution completely). We do not use their learning model in our

MLE exercise, but account for this learning using other methods.

In the MLE for treatment 1, we estimate one $P(L_1)$ for observations with (Exp, Exp) groups and another $P'(L_1)$ distribution for observations with (Inexp, Inexp). However, for the (Exp, Inexp) pairs, we estimate a separate distribution $P(L_1)_{Exp}$ for the Exp players, and a separate $P(L_1)_{Inexp}$ for the Inexp players because either type was informed that they are playing an opponent of a different type. Thus, their subjective beliefs may have been affected by this information.[27] We estimate only a single error term ϵ for all treatment 1 data.

In the MLE for treatment 2, we allow for a different distribution over levels in *each* round, which is the strongest possible effect learning can have, further, we also allow the level distribution to change between the C13 and H13 parts of a round. Our conclusions do not change upon doing these robustness checks. The HS Level-k model assumes that no one chooses the level-0 strategy. Therefore, without the error term, the Level-k model puts 0 probability on the outcome in which one of the group members lost in C13, which is the case for 57 percent of the combined sub-session data. However their error term covers for the latter eventuality and allows us to perform an MLE using the HS model.

7.3 The Agent Quantal Response Equilibrium

The Agent Quantal Response Equilibrium (AQRE) model of McKelvey and Palfrey (1998) extends the Quantal Response Equilibrium to extensive form games. The AQRE model introduces a separate additive *payoff disturbance* error term to the expected payoff associated with each action for each player, from each possible move of that player. In particular, let $\overline{u}_{ija}(b)$ is the expected payoff of player i from playing action a at the information set h^i_j, given that the behavioral strategy profile for all players is b. Then the player's "actual" payoff in their model is $\hat{u}_{ija} = \overline{u}_{ija}(b) + \epsilon_{ija}$. They assume that ϵ_{ija} is i.i.d. according to type I extreme value distribution with c.d.f. $F(\epsilon_{ija}) = e^{-e^{-\lambda\epsilon_{ija}}}$. Let $A(h^i_j)$ be the action set at the information set h^i_j. Then in an AQRE b, the probability of i choosing action a at h^i_j is $b^i_j(a) = [e^{\lambda\overline{u}_{ija}(b)}]/[\sum_{a'\in A(h^i_j)} e^{\lambda\overline{u}_{ija'}(b)}]$. The parameter λ generates a certain AQRE b. The behavioral strategy profile b in turn implies a probability distribution over the observed outcomes.

[27] In treatment 1, we do not model learning across rounds as Ho and Su (2013) do because: first, we focus on just the first three rounds. This means that there is very little learning opportunity (we focus on just three rounds because that is when there is an actual difference between Exp and Inexp subjects, as discussed below). Second, unlike their data, in treatment 1, we don't have exact repetitions of the same game in the same environment. Any player can be matched to a different type of player in each new round. Last, we estimate four independent level distributions, which, we argue, captures learning well.

In the MLE for treatment 1, we estimate one λ for observations with (Exp, Exp) groups and another λ' for observations with (Inexp, Inexp). For the (Exp, Inexp) pairs, we estimate a separate λ_{Exp} for the Exp players, and a separate λ_{Inexp} for the Inexp players. In the MLE for treatment 2, we estimate a single λ parameter. The AQRE model does not model updating/learning, however, the probability of the outcome observed in a round of C13-H13 is a product of the probabilities of the observed outcomes in the C13 part and the H13 part of that round. Each outcome in the C13 part has a probability calculated using the win (in C13) and loss (in C13) probability, determined by λ. The effect of win or loss in C13, by model is summarized in Table 16 in the Appendix.

In terms of our earlier discussion, let o_i be an observed outcome in the data, and let f_i be its observed frequency. Then:

$$Prob^{AQRE}(o_i) = \sum_{(s_1,s_2)\in S^2(o_i)} [Prob(s_1|b(\lambda))Prob(s_2|b(\lambda))Prob(b(\lambda))] \tag{8}$$

Then the likelihood function of the AQRE model is given by:

$$LogL(AQRE|\lambda) = \sum_{i=1}^{N} f_i log(Prob^{AQRE}(o_i)) \tag{9}$$

8 Comparative Data Analysis of Behavioral Models

We report MLE results of the three models discussed in section 7 with respect to data. We focus only on the data from the combined sub-sessions of the two treatments. In total 56 different types of outcomes were observed in treatment 1, and 245 types of outcomes were observed in treatment 2, not counting outcomes as different even if the sequence of moves in C13 was different.

8.1 Treatment 1: The Avoid 9 Game

In Table 6, we report the MLE results from the first 3 rounds of the combined sub-session in treatment 1.[28]

[28]This is because those are the rounds where there is a clear difference between the level of understanding of Inexp and Exp subjects, which is the focus of this paper. The order of the three models, and the significance of the differences remains unchanged if we take the whole combined sub-session data into account.

Treatment 1 MLE Result- *The likelihood rankings of the three models using pairwise Vuong (1989) test is: LFLE* $>^{**}$ *AQRE; LFLE* $>^{***}$ *HS; AQRE>HS.*

Table 6, reports the likelihood and parameters of the three different models discussed in section 7: (i) Limited Foresight and Learning Equilibrium (LFLE), (ii) The Dynamic Level-k Model in Sequential Games of HS (2013), and (iii) The AQRE model. We separately estimate the parameters for each model for the three types of pairs that were possible in the combined sub-session: (Exp, Exp), (Exp, Inexp) and (Inexp, Inexp). The LFLE model has the highest likelihood for each type of pair, and therefore also overall. AQRE comes in second, and the Dynamic Level-k model is third. To investigate the statistical significance of these results, we conduct pairwise Vuong (1989) tests. The difference between the likelihood of the LFLE model and the AQRE model is significant with a p-value of 0.03 and a z-statistic of 2.16. The difference between the likelihood of the LFLE model compared to the Dynamic Level-k model is significant with a z-statistic of 3.15, and a p-value < 0.01. The difference between the likelihood of the AQRE model compared to the Dynamic Level-k model is statistically insignificant.[29] The Kawagoe and Takizawa (2012), Level-k model uses a logit error structure similar to the AQRE model. The LFLE and the HS model both use error-less strategies and the same error structure. Comparing the logit Level-k model to LFLE is comparing the LFLE model to a level-k model with different error structures. However, if one insists on this comparison, we find that the likelihood of the logit Level-k model is -899.45, which is insignificantly lower than the LFLE. The Akaike Information Criterion comparison also favors the LFLE model, but the Bayesian Information Criterion favors the logit level-k model.[30]

Thus, we conclude that the results from treatment 1 points towards LFLE as the weakly more likely explanation of the observed outcomes. The key driver of the difference among the likelihoods of the LFLE model and the Level-k model is that without the error term, the LFLE model puts zero probability on 2.2 percent of the data, whereas the Level-k model (both logit and HS) puts zero probability on 13.4 percent of the data. This is because the latter model can't explain imperfect play after the first/second mover has been decided, without an error term.

[29] For the whole combined sub-session the likelihoods are -2245, -2290, and -2294 for the LFLE, HS and AQRE models. The LFLE model has a significantly higher likelihood (p-value < 0.01)

[30] We don't use the logit error term for the LFLE model because for the logit model it doesn't just matter which action gives more expected payoff. The amount of difference also matters. Thus the logit prediction for the LFLE model is not robust to the "curtailment rule" used.

Table 6

Relative MLE performance of the theoretical models.

Model	Inexp Inexp	Inexp Exp	Exp Exp	Ln(L)
LFLE	-260.85	-487.84	-141.86	-890.55
$Prob(0,1,2,f)$	(0.25, 0.23, 0.31, 0.2)	Exp: (0.05, 0, 0, 0.95); $\delta = 0.31$ Inexp: (0.48, 0, 0.32, 0.2)	(0, 0, 0.05, 0.95)	
ϵ	$8{\times}10^{-5}$			
Level-k	-285.34	-503.4	-148.33	-937.07
$Prob(L_1)$	0.65	Exp: 0.31; Inexp: 0.60	0.05	
ϵ	$3{\times}10^{-4}$			
AQRE	-272.82	-505.27	-143.84	-921.93
λ	0.015	Exp: 0.017; Inexp: 0.019	0.028	

Notes. The table reports the log likelihood and parameters of the LFLE, HS Level-k, and AQRE models for the first three rounds of the combined sub-session. The sample size is 231.

8.2 Treatment 2

There are 1591 different sequences of moves possible in the H13 game alone. Therefore, the total number of different sequences of moves possible in a round of treatment 2 is more than 10 million. Thus, to give the Ho and Su model a fair chance (which has an error term where each sequence must have the same minimum error probability of occurring), and for analytical tractability, we divide the observations from treatment 2 into 81 broad categories.[31] The details of this categorization are in the Appendix. We categorize the observations on the following basis: (i) Exp/Inexp combination of a pair; (ii) the outcomes of a pair in their respective C13 game against the computer (iii) perfect/imperfect play by the selected first and second movers in H13.

Treatment 2 MLE Result- *The likelihood rankings of the three models using pairwise Vuong (1989) test is: LFLE* $>^{***}$ *AQRE; LFLE* $>^{***}$ *HS; HS* $>^{***}$ *AQRE.*

Table 7 states the MLE results and parameters for each of the three models from section 7 with respect

[31]No such division was done for treatment 1 as the number of possible outcomes was only 447.

to the data from the combined sub-session of treatment 2 (rounds 9-14). The LFLE model has the highest likelihood, followed by the Dynamic Level-k model, which in turn has a significantly higher likelihood than the AQRE model.

Table 7

MLE results for Treatment 2

Model	Log Likelihood	Parameters
LFLE	-820.23	$Prob(0, 1, 4, f) =$ (0.12, 0, 0.38, 0.5)
Level-k	-1034.09	$\epsilon = 0.0075$. $Prob(L_1|$ Exp, Opponent Won) = 0.13
		$Prob(L_1|$ Inexp, No Info) = 0.11
AQRE	-1352.6	$\lambda = 0.0056$

Notes. The table reports the log likelihood and parameters of the LFLE, HS Level-k, and AQRE models for the combined sub-session. The sample size is 295.

Vuong's test for comparing the likelihoods of non-nested models shows that the likelihood of the LFLE is significantly greater than the likelihood of the Level-k model (test statistic 7.95, $p-value \approx 0$) or the AQRE model (test statistic 12.06, $p-value \approx 0$). With a test-statistic of 7.95, the likelihood of the Dynamic Level-k model is also significantly greater than the likelihood of the AQRE model with a p-value of approximately 0. Note that the results in Table 7 are for the LFLE model without an error term. While for the Dynamic Level-k model of Ho and Su (2013), 57% of the data would have 0 probability sans the error term.

The Ho and Su (2013) model allows for changes in level distribution across repetitions of a game. This level distribution changes because of changes in beliefs. One can argue that the two parts of the C13-H13 round are two repetitions of a very similar game and that the Exp subjects could have formed different subjective belief based on observing the opponent's play in C13. Therefore in the results given in Table 7, we have allowed for players to choose different levels in H13 based on the information they get from their opponent's behavior in C13. In C13, all players choose Level-2 or above because they are rational and they are explained that the computer plays perfectly (the MLE results on Ho and Su (2013) and resulting comparisons do not change if we allow subjects to choose the Level-1 strategy in C13). After

C13 ends, one of three cases happen (i) an Exp subject observes his opponent lost in C13 (ii) an Exp subject observes his opponent won in C13 (iii) a subject observes nothing. Case (i) implies that the HS model puts 0 probability on that observation. We allow players to mix differently between Level-1 and Level-2 (the two possible levels) in H13 for cases (ii) and (iii).

As a robustness check, in the HS model, we allow for learning across rounds, i.e., we allow for a separate distribution on levels (and separate additional parameters), in each of rounds 9 through 14. The log likelihood of the data as per the sequential level-k model with such a specification is -993.96, which doesn't change our conclusions.[32] As a last robustness check we try the KT model of Level-k which uses a logit error structure. The likelihood of the KT model is -904.68, which is also significantly lesser (z-statistic of 3.55, and a p-value of 0.0004) than the likelihood of the LFLE.[33]

8.3 Discussion

The data suggests that in the first treatment, 13.4 percent of the times, subjects played imperfectly after the first and the second mover had been decided. Thus, most deviations from the perfect strategy occur at the F/S decision stage, which can also be accounted for by the Level-k model, which uses the lack of iterative reasoning among subjects as the reason for this finding. The MLE results favor the LFLE model, which puts 0 probability on only 2.2 percent of the data without the error term, as opposed to 13.4 for the Level-k models. However, the logit Level-k model is a very close competitor, even though the comparison among different error structures is arguably inappropriate. Thus, the treatment 1 data does not show a high enough incidence of a failure to play the weakly dominated strategy in $A9_{sub}$ to conclusively distinguish among models.

The data from the second treatment shows that there was a high incidence of limited foresight, as 41.4 percent of subjects lost their interaction with the computer in C13 when there was no incentive to do so,

[32] We also test a Cognitive Hierarchy (Camerer, Ho and Chong (2004)) model adapted to sequential move games. Their model includes the empirical existence of a level-0 player who randomizes uniformly. Further, they constrain beliefs to be "partial rational expectations" using prior beliefs. As a robustness check we relax this condition and search over the distribution of possible Levels that maximizes likelihood. The resulting maximum likelihood is -873-56. The LFLE model has a significantly higher likelihood (z-statistic of 2.91 and p-value of 0.004) than even this "free" distribution cognitive hierarchy model. The difference becomes higher if we impose the "partial rational expectations" condition.

[33] This result is for the case where subjects cannot choose Level-1 in C13 (because of them being rational and it being explained to them that the computer plays perfectly), however all subjects are allowed to switch to any level once C13 ends and before H13 begins. Even if 13% of the Inexp subjects are allowed to choose Level-1 in C13 (that is, allowed to hold beliefs that the computer randomizes uniformly) in the KT model, its log-likelihood is -863.9, which is significantly lower, with a z-statistic of 2.01 (two tailed p-value of 0.045) than the log likelihood of the LFLE model.

and it was a weakly dominated strategy (given that it was announced and explained that the computer plays perfectly, playing imperfectly in C13 was a strictly dominated strategy). Thus, limited foresight and uncertainty about the opponent's foresight played a primary role in generating the data in treatment 2. The MLE results indicate that the LFLE model fits the treatment 2 data the best.

Another indicator of limited foresight is the comparison between the percentage of cases in which imperfect play was observed in $A9_{sub}$ with the percentage of such cases in the similar, but longer, $H13_{sub}$. For the $A9_{sub}$ this percentage was 10.57 percent, for the latter it increased to 24.23 percent, which is a significant increase with a p-value of approximately 0. Further, there also seems to be evidence of uncertainty about the opponent's foresight. Notice that the LFLE implies that at $P = 0$ in the Avoid 9 game, the type-f player removes 1 item if he believes that an opponent with foresight level of 1 exists. And similarly at $P = 4$ in the H13 game, the type-f player removes 1 item for the same belief. Now notice that according to the MLE estimates of LFLE, the type-f is the most likely type in both treatments 1 and 2. Thus, the LFLE can partially explain results 5(a) and 5(b) which show that players choose to remove 1 item from losing positions. According to LFLE, the uncertainty about the opponent's type is what makes the type-f player remove 1 item, as this maximizes the probability of the opponent making a mistake.

9 Conclusion

We report results from an experiment using a sequential move "winner take all" game which we constructed by adding a First/Second mover decision stage to a "race game". In the game we constructed, one can attain a prize of 200 ECUs by choosing a sure-shot winning strategy. One can attain a higher prize in this game by winning from a losing position. Winning from a losing position is possible only if the opponent makes a mistake and doesn't choose his weakly dominant strategy. The results in treatment 1 indicate that experienced subjects, who understand how to win the "winner take all" game, are more likely to risk losing to try to attain the higher prize when they are told that their opponent is inexperienced. In treatment 2 we replace the exogenous information about the opponent's experience level with information about the play of the opponent in a closely related zero-sum race game against the computer. We found that if the experienced subjects were shown that their opponent lost the race game versus the computer, then the experienced subjects were more likely to risk losing to attain the higher prize in the "winner take all" game against that opponent. The results in treatment 2 indicate that experienced subjects become better at guessing their opponent's level of experience after observing more moves of the opponent.

These findings point to a systematic failure of SPNE within the paradigm of *selfish and rational* behavior. We can be sure that the behavior paradigm is indeed selfish because the game is "winner-take-all." We know that experienced players are acting rationally in choosing to deviate from the SPNE strategy, and risking a loss, because we can observe that they converge to SPNE behavior and play the "sure-win" strategy when playing another experienced opponent. Further, even if one argues that subjects being told their opponent's experience-level is not very relevant to dynamic games in the real world, the finding that subjects observe past behavior and infer the experience-level of the opponent and then act on this inference suggests a need for theoretical investigation of the data reported here.

We test the relative performance of three models in explaining this data: (a) the Limited Foresight and Learning Equilibrium (LFLE) (b) the Agent Quantal Response Equilibrium for extensive form games (AQRE), and (c) the Dynamic Level-k Model for Sequential Games (Ho and Su (2013)). Comparing the likelihoods of these models with respect to the data, we find that the data from both treatments favors the LFLE explanation. The reason for this finding is that the Level-k model, without errors, puts zero probability on weakly dominated play. But weakly dominated play is observed in 13.4 percent of the treatment 1 data and 57 percent of treatment 2 data. In contrast, the LFLE model, which explains weakly dominated play in dynamic games using a model of limited foresight and uncertainty about the opponent's foresight, explains the observed patterns of play well.

Appendix

Treatment 1 Data Analysis

Difference in Proportion of "First Mover" Choice by Round and Opponent's Type. We analyze the round-by-round behavior of Exp subjects between rounds 13-20 in Table 8. For example, in round thirteen, 42 percent of the Exp subjects who faced an Inexp subjects chose F, while 8 percent of the Exp subjects who faced an Inexp opponent chose F. This difference is significant at 1 percent, with a p-value of 0.001. The difference in proportion of Exp subjects choosing F based on opponent being Exp or Inexp remains significant from round 13 through round 17. The disappearance of this difference after round 17 suggests that by round 18, almost all the Inexp subjects had acquired experience and learnt the "sure-win" perfect strategy of the second mover in $A9_{sub}$, and that the Exp subjects were aware of this.

Table 8

Percentage of first mover choices of Exp subjects by round and opponent's experience

Round	13	14	15	16	17	18	19	20
Opponent Inexp	42%	24%	31%	17%	10%	0	2%	0
Opponent Exp	8%	2.5%	6.3%	0	0	0	3%	3%
p-value: two tailed	0.001	0.006	0.009	0.015	0.058	-	0.782	0.327

Learning Speed Difference. Table 9, second row, reports the percentage of F choices of Exp players playing other Exp players in their first 8 rounds of playing the Avoid 9 game. The fourth row of Table 9 reports the percentage of F choices made by Inexp players when facing another Inexp player in *their* first 8 rounds, which are rounds 13-20 of the session, of playing the Avoid 9 game. The bottom row reports the p-value comparing the second row to the fourth row, for each column. In their first three rounds, both Exp and Inexp subjects behave similarly, but for their rounds 5, 6, and 7 (session round numbers 17,18 and 19), Inexp subjects choose F in a significantly smaller percentage of cases against other Inexp players, as compared to the Exp players in their rounds 5, 6 and 7 of playing the Avoid 9 game.

Table 9

Percentage of F Choices of Exp against Exp, Inexp against Inexp in their respective first 8 rounds.

Round (Exp)	1	2	3	4	5	6	7	8
Exp against Exp	78%	64%	55%	47%	37%	36%	18%	10%
Round (Inexp)	13	14	15	16	17	18	19	20
Inexp against Inexp	82%	70%	53%	34%	10%	14%	0%	4.5%
Two tailed p-value of percentage diff	0.65	0.5	0.39	0.19	0.002	0.011	0.007	0.237

The findings in Table 9 are confirmed by a probit estimation in Table 10. We set the F/S decisions of Exp against Exp and Inexp against Inexp in their respective first 8 rounds as the dependent variable. We use session dummies as controls, and one other independent variable: a dummy for "the player is Exp."

The probit result is the following. We find that the dummy for "the player is Exp" significantly increases the probability (p-value 0.037) of choosing F compared to the baseline of "the player is Inexp."

Table 10: Treatment 1 learning speed difference

	Dependent variable	
	(1): Perfect play	(2): choosing$F$
Subject is Exp	-0.27*(0.14)	0.24**(0.12)
Constant	0.84***(0.25)	0.36(0.25)
Controls	Session Dummies	Session Dummies

Notes. Model (1) dependent variable takes value 1 when the selected second mover plays perfectly. Model (2) dependent variable takes value 1 when the player chooses F. Figures in parenthesis are robust standard errors. Column 2 has 460 observations with 147 clusters, and column 3 has 920 observations with 154 clusters.

Treatment 2 Data Analysis

Difference in Proportion of "First Mover" Choice by Round and Opponent's Performance in C13. We report the round-by-round behavior of Exp subjects based on opponent's performance for rounds 9-14 in Table 11. For example, in round 10, Exp subjects chose F 62 percent of the times against an opponent who lost C13 versus the computer, and 26 percent of the times against an opponent who won C13 versus the computer. This difference is significant at LOS 5 percent with a two tailed p-value of 0.015.

Table 11

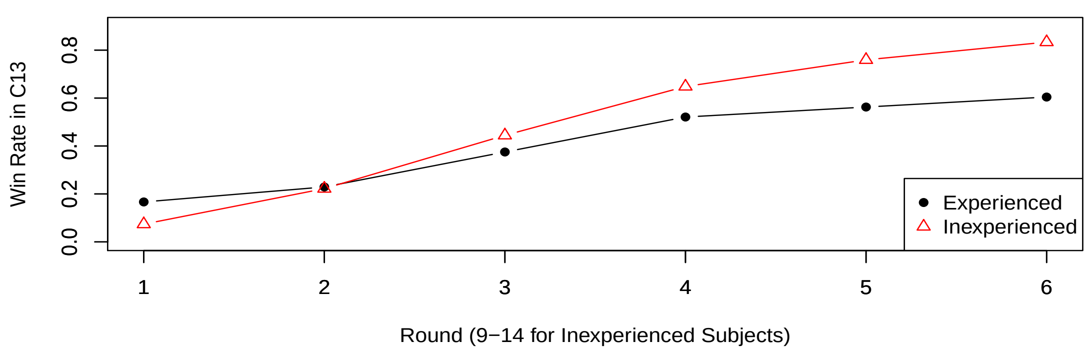
Percentage of times Exp subjects chose F in H13 by Round and Opponent Performance

Round	9	10	11	12	13	14
Opponent lost C13	40%	62%	37%	41%	36%	28%
Opponent won C13	39%	26%	21%	20%	11%	12%
p-value of two-tail test	0.94	0.015**	0.25	0.124	0.046**	0.305

Result 3(b): *Faster learning speed of Inexp subjects*: During their 6 rounds in the combined sub-session, the Inexp subjects learned the perfect strategy in the Computer 13 game faster than Exp subjects in their first 6 rounds in the training sub-session.[34]

Figure 4 depicts the win rate of Inexp and Exp subjects in the C13 part of their respective first 6 rounds. The win rate is defined as the number of wins of type i subjects in the C13 part divided by total number of type i subjects, where $i \in \{Inexp, Exp\}$. Table 12 (below), second row, reports the percentage of wins of Inexp players playing the computer in the C13 part of their first 6 rounds (rounds 9-14 of the session), by round. The third row reports the same for Exp subjects in their first six rounds (rounds 1-6 of the session). In their first 4 rounds, both Exp and Inexp subjects perform statistically equivalently (see Table 12). But in their rounds 5 and 6 (rounds 13 and 14 of the session), Inexp subjects win a significantly higher proportion of their C13 games, as compared to the Exp players in the rounds 5 and 6 of the training sub-session. The p-values of these differences are 0.035 and 0.017, respectively. Recall that Inexp subjects, unlike the Exp subjects, do not get information about how their opponent played the C13 part of a round. So Inexp subjects have lesser information in their rounds 2 (round 10 of the session) through 6 (round 14 of the session), compared to Exp subjects in their rounds 2 through 6. However, like for result 3(a), the data does not provide a clear answer. That is, separating Inexp subjects by how often they played Exp subjects produces no systematic pattern of learning difference.

Figure 4: Learning Speed Comparison



Notes. The figure depicts the behavior of experienced subjects in rounds 1-6 of the session and the inexperienced subjects' behavior in rounds 9-14 of the session (which are the inexperienced subjects' first 6 rounds). The figure depicts the round-wise proportion of experienced subjects who won C13 (black dots) and the proportion of inexperienced subjects who lost C13 (red triangles). It is a dominant strategy to play perfectly in C13 and win. The latter proportion is significantly higher, with a p-value < 0.05, for rounds 5 and 6. This shows the faster learning speed of the inexperienced subjects.

[34] We focus on the C13 game because it provides a clean measure of the proportion of subjects who understood the perfect strategy. It is a weakly dominant strategy to play perfectly (strictly dominant given that subjects were told that the computer plays perfectly) in this game. The F/S decision in C13 was not dependent on the belief about the ability of the human opponent to play perfectly. Further, the only way to win against the computer was to play perfectly at each decision node, including the F/S decision stage.

Table 12

Percentage of Wins in the C13 part of respective first 6 rounds, by types.

Round (Inexp)	9	10	11	12	13	14
Inexp vs Computer	7.4%	22.2%	44.4%	65%	76%	83.3%
Round (Exp)	1	2	3	4	5	6
Exp vs Computer	16.7%	23%	37.5%	52%	56%	60.4%
p-value of two-tail test	0.147	0.933	0.477	0.192	0.035**	0.017**

Result 4(b): *The rate of imperfect play declined significantly as the players got closer to the end of the H13 game.* That is, as the position increased, or as the number of items left reduced, the rate of imperfect actions reduced. In particular, no subject lost from a winning position with 4 or lesser items left in the H13 games played in the combined sub-session.

Table 13 shows the observed rates of imperfect play by position in the H13 games played in treatment 2. We show the rates separately for three parts of the data: (i) session (training and combined sub-sessions together); (ii) the first three rounds of Exp subjects in their training sub-session; and (iii) the first three rounds of Inexp subjects in the combined sub-session. Recall that a position is a particular sum of items removed. For each of (i), (ii), and (iii), a test of difference in proportions reveals that the rate of imperfect play is significantly more (p-values < 0.01) from a position in 1-3 (10-12 items remaining) as compared to 5-7 (6-8 items remaining), and from the latter compared to a position in 9-11 (2-4 items remaining).

Table 13

Rate of imperfect play by position

	Position		
	1-3	5-7	9-11
Session	21	5	0.8
Exp: first 3 rounds	45.5	21.7	3.5
Inexp: first 3 rounds	34.8	12.2	0

Notes. The figures are in percentage. The figures are reported from positions from which a perfect action is distinguishable from an imperfect action. A perfect action makes the position 4, 8, and 12 from a position in $\{1,2,3\}$, $\{5,6,7\}$, and $\{9,10,11\}$ respectively.

Result 5(b): *Choice proportions at losing positions in H13.* At the losing positions 0, 4, and 8 in H13, subjects were significantly more likely (with p-values of approximately 0 for each) to choose to remove 1 item than the next most likely alternative. This results also holds true for Exp subjects in the combined sub-session. These results can be seen in Table 14 below.

Table 14

Proportion of choices from losing positions

Positions	Session				Exp in Combined sub-session			
	Remove 1	Remove 2	Remove 3	p-value	Remove 1	Remove 2	Remove 3	p-value
0	40.7	29.8	29.5	0.0004	41.1	34.9	24	0.3048
4	48	24.5	27.5	0	43.4	29.2	27.4	0.0269
8	53.3	23.9	22.8	0	53.3	23.9	22.8	0.0003

Notes. The figures are in percentages. The p-values are two tailed p-values comparing the proportion of times 1 was chosen, as compared to the next most chose alternative.

Result 6(b): *Ex-post optimality of being "First Mover" in H13 against an opponent who lost C13.*

In the combined sub-session (rounds 9-14), when the opponent lost C13, the average earnings of the Exp player was weakly more when the Exp player was the selected first mover in H13 rather than when he was the selected second mover. Thus, in the combined sub-session, it was ex-post optimal for an Exp player to choose F in the H13 game when facing an opponent who lost his/her C13 game with the computer. This can be seen in the Table 15 below.

Table 15

Ex-post average earnings of the Exp subject in H13 against opponent who lost C13

Round	Earning as Second Mover	Earning as First Mover	p-value
9-14	187.7 (18)	193.2 (27.8)	0.8
9-10	181.8 (23.2)	163.9 (35)	0.427
11-12	200 (32.2)	320 (53.9)	0.035**
13-14	185 (45)	162.5 (67.5)	0.743

LFLE Details

Avoid 9 Game

We first complete our description of the LFLE for the Avoid 9 game.

The **type-1** player has foresight level of 1. Le h be an arbitrary move of the type-1 player. Let a^0 be an action available at h, and a^1 be an action available in the next stage following $h \to a^0$. Then from h, the type-1 player observes all sequences $\{h \to a^0 \to a^1 \to$*curtailed payoff profile*$\}$. At the F/S decision stage, the type 1 player observes all actions of the selected first mover, each action with an associated curtailed payoff profile. The curtailed payoff profile for all observable sequences is $(\frac{500+50}{2}, \frac{200+50}{2})$. Thus, the type-1 player chooses F. Similarly, from $P \leq 1$, the type-1 player observes that for all observable sequences, the curtailed payoff profile is $(\frac{500+50}{2}, \frac{200+50}{2})$. Thus he randomizes among available actions at $P \leq 1$. Both these decisions are regardless of the opponent's type or the succeeding action.

From $P = 2$, the opponent's move in the next stage matters for the type-1 player: if he thinks that his opponent will play perfectly if the sum of items removed is 5 or more, then he should choose an action such that $a < (5 - P)$. How does the LFLE deal with the limited foresight type's perception about other players' types' actions? To answer this, we first define *total foresight.* Let a limited foresight type's *total foresight* be the sum of (i) the stage number that the limited foresight type is moving at, and (ii) the level of foresight of that limited foresight type. Consider a player-type, X, at a move h. X has a certain total foresight. X is observing a curtailed version of the underlying game of perfect information. For our purposes, the LFLE definition boils down to three **rules of thumb**: **(a)** X knows the LFLE actions of the player types with lesser total foresight than him; **(b)** X assumes that equal or higher total foresight types play a sequentially rational strategy based on the curtailed version of the underlying game he observes at h, and X chooses an action at h that is a sequentially rational best response to such a strategy profile; and **(c)** X's perceived beliefs of equal or higher total foresight types, used to determine their sequentially rational strategy in (b), are Bayes' rule consistent with the strategy profile that obeys (a) and (b).

By rule-of-thumb (b), the type-1 player assumes that any type of his opponent will play perfectly to attain the winner's payoff given a move at $P \geq 5$. Thus, for $1 < P < 4$, the type-1 player randomizes over the action set $\{a : P + a < 5\}$. From $P \geq 4$, the type-1 player has enough total foresight to play perfectly.

From any move h, the **type-2** player observes all sequences $\{h \to a^0 \to a^1 \to a^2 \to$*curtailed payoff* *profile*}, where a^1 is an action available at the next stage after $h \to a^0$, and a^2 is an action available at $h \to a^0 \to a^1$. At the F/S decision stage, the type 2 player chooses F regardless of beliefs because after every observable sequence, the curtailed payoff profile is $(\frac{500+50}{2}, \frac{200+50}{2})$. As the selected first mover with $P = 0$, by rule (b), the type-2 player chooses[35] $a^0 = 3$. From $P \geq 1$, the type-2 player plays perfectly observing that choosing a_4^* and a_8^* guarantees a winner's payoff.

From $P \geq 1$, for identical reasons as the type-2 player, the type-f player also plays perfectly. As the first mover at stage two (i.e., at $P = 0$), the type-f player removes 1 item to maximize the probability that the opponent (type-0 or type-1) will *not* play a_4^*. The maximized probability is $(p_0 + p_1) \times \frac{2}{3}$. Thus, the type-$f$ player chooses F iff his belief on the opponent being type-0 or type-1 is high enough.

Human 13 and Computer 13 Games

At the F/S decision stage, the **type-0** player chooses F/S with probability $\frac{1}{2}$ each, as he observes that the curtailed payoff profile is (275,275) after both F and S. From any move with $P < 9$, the type-0 player observes the curtailed payoff profile (275,125) after every available action. Therefore from any such move, the type-0 player randomizes uniformly among $1, 2, 3$. The threshold position after which all foresight types, including the type-0 player, play perfectly is 9, as opposed to 5 for the Avoid 9 game. This is because in H13, if $P \geq 9$, then the curtailed payoff after $a_{12}^* = (12 - P)$ is the winner's payoff.

The **type-1** player, from any move h, observes all sequences $\{h \to a^0 \to a^1 \to$*curtailed payoff profile*$\}$. The type-1 player chooses F at the F/S decision stage, as he observes only (275,125) as the possible curtailed payoff profile for all observable sequences from the first stage. For any move where $P \leq 5$, the type-1 player observes the curtailed payoff profile (275,125) after any observable sequence of actions. Therefore if $P \leq 5$, the type-1 player randomizes uniformly among $\{1, 2, 3\}$. From $P \in \{6, 7\}$, the type-1 player puts equal weight on a^0 such that $P + a^0 < 9$, because he assumes (by rule (b)) that if $P + a^0 \geq 9$, any type of his opponent will choose $a^1 = a_{12}^*$, giving him the loser's payoff. At $P = 8$, the type-1 player

[35] This is because, given type-0's LFLE action at $P \leq 3$, removing 3 items maximizes the probability that his opponent will choose a^1 such that $a^0 + a^1 > 4$. And if $a^0 + a^1 > 4$, then the type-2 player can choose $a^2 = a_8^*$ to obtain the winner's payoff as the curtailed payoff. Due to his limited foresight, this is the only sequence observable to him in which he wins as the first mover. Notice that the only position at which the belief about the opponent's type matters for type-2 is at $P = 0$. If the type-2 player puts positive probability on the opponent being type-0, he removes 3 items as the first mover at $P = 0$. This belief is consistent with the type-0 player randomizing at the F/S decision stage.

is assured a loss, and thus he randomizes uniformly among $\{1,2,3\}$. From $P \geq 9$, the type-1 player plays perfectly.

The **type-4** player has a high enough foresight level such that, except at $P = 4$, and $P = 0$, he plays perfectly to attain the winner's payoff. At $P = 0$, i.e., as the first mover beginning the game, the type-4 player removes 3 items,[36] and at $P = 4$, he removes 1 item.[37] At the F/S decision stage, the type-4 player observes that if the first mover chooses 1 item at each move, then the curtailed payoff profile after the fifth stage action can only be (275,125). As $275 > 200$, the maximum possible for the second mover, the type-4 player chooses F.

The type-f player plays exactly as the type-4 from $P \geq 1$ for identical reasons. As the first mover, the type-f player randomizes because he knows (by rule (a)) that the probability of an opponent's mistake is the same from $P = 1, 2$ or 3. At the F/S decision stage, the type-f player chooses F iff his updated belief (updating happens only if he observes opponent's actions in C13; the updating uses Bayes' rule and the LFLE of C13) on the opponent being type-0 or type-1 (belief on type-0 and type-1 opponent is denoted as u_0 and u_1, respectively) is high enough. The probability that the opponent makes a mistake as the second mover is $(u_0 + u_1)(\frac{2}{3} + \frac{1}{3}.\frac{2}{3})$. Thus the threshold for $u_0 + u_1$ is 0.375, beyond which the type-f player chooses F.

In the LFLE for the **C13** (Computer 13) game, there are only one two changes compared to the LFLE for the H13 game: (i) in the first stage, types 1 and 4 play F/S with 50% chance each. This is because in C13, winning as the first or the second mover gives the same payoff: 500. (ii) Types 4 and f do not favor any particular action from $P = 4$ or $P = 1$, as they are playing against the computer and they are informed that the computer plays perfectly.

Data Categorization in Treatment 2

Recall that stage 1 of the H13 game is the F/S decision stage. Let I denote the selected first mover in H13, and II denote the selected second mover in H13. We first categorize the data from $H13_{sub}$ into three categories: (a) I played an arbitrary action at $P = 0$, II played imperfectly at stage three, I made

[36] The choice at $P = 0$ is sequentially rational because the type-4 player's limited foresight at stage two allows him to observe only one sequence where he wins as the first mover: if the second mover makes the sum of items removed *more than* 4 after the third stage move.

[37] At $P = 4$, the type-4 player attains full total foresight in the H13 game. Because of rule (a), the choice at $P = 4$ is optimal with respect to the LFLE, as it maximizes the probability that the type-0 or type-1 opponent will not play a_8^*.

the sum 4 or 8 (whichever was possible) in stage four; (b) I played an arbitrary action in stage two, and II played perfectly (made the sum of items removed 4) in stage three; (c) I played an arbitrary action in stage two, II played imperfectly in stage three, I also failed to make the sum 4 or 8 in stage four. These categories can be safely interpreted as (a) arbitrary-imperfect-perfect; (b) arbitrary-perfect; and (c) arbitrary-imperfect-imperfect. This is because in every case in which a player made the sum 8, two stages later he made the sum 12 and won the game. Further, in 97 percent of the data, if a player made the sum 4, two stages later, he made the sum 8.

Each of (a), (b), and (c) has three distinct cases possible for the choices in the F/S stage of H13. These three cases are: (1) in H13's F/S decision stage, I chose F and II chose F; (2) I chose F and II chose S; (3) I chose S and II chose S. For each of (1)-(3) in the F/S decision stage of H13, we have (a)-(c) in $H13_{sub}$, which makes 9 categories.

Treatment 2's combined sub-session also had two types of players playing together: Inexp and Exp. While the Exp subjects were informed about their opponent's play in C13 before the beginning of H13, the Inexp subjects were uninformed about their opponent's play in C13. Further, there were four possible outcomes for a pair of subjects from the respective play of the two members of the pair against their respective computer in the C13 part of the round: Win-Win, Win-Loss, Loss-Win, and Loss-Loss, where the first term is I's outcome in C13 and the second term is II's outcome in C13. Thus, we need to further broaden the categories of outcomes for treatment 2.

We make a total of 81 categories for the data from treatment 2. The 9 categories above are repeated in 9 cases. Case (i): (Exp_W, $Inexp_W$) , which denotes that I is Exp and he won his C13 game, while II is Inexp and he also won his C13 game. Case (ii): ($Inexp_W$, Exp_W), with the order from (i) swapped. Case (iii): (Exp_W, $Inexp_L/Exp_L$), which denotes that I is Exp and he won his C13 game, while II lost his C13 game, and he can be either Inexp or Exp. [38] Case (iv): ($Inexp_L/Exp_L$, Exp_W) is the same as case (iii) with the order swapped. Proceeding similarly we make case (v): (Exp_W, Exp_W); (vi): ($Inexp_W$, $Inexp_W$); (vii): ($Inexp_W$, $Inexp_L/Exp_L$); (viii) ($Inexp_L/Exp_L$, $Inexp_W$); (ix) ($Inexp_L$, $Inexp_L$). For

[38] We club the Exp_L and $Inexp_L$ I or II in one category, as the Exp player who loses in C13 cannot be type-f according to LFLE. Thus, he has the same strategy in LFLE, regardless of his beliefs. That is, in LFLE, learning about the opponent's type within the round does not make a difference to $Exp'_L s$ strategy, while $Inexp_L$ cannot learn by design. On the other hand, the sequential level-k model implies that an observation where one of the group members loses C13 has 0 probability (not accounting for the error term: ϵ). Thus all such observations are treated homogeneously in their model. The AQRE model also does not account for any learning, and therefore we reduce the number of categories without affecting the result.

each of the cases (i)-(ix), the categories (1)-(9) are repeated, making 81 total categories.

Treatment 2: The Effect of the Outcome in C13

According to the LFLE model, at the beginning of the H13 part, the distribution on the foresight levels changes based on the outcome in C13. Table 16 describes the change in distribution on the foresight levels due to the outcome of win (W) or loss (L) in the C13 part of a round. According to the LFLE model, the Exp players observe these outcomes and update their belief. However, beliefs change the subsequent strategy in H13 only for the full foresight type (type-f) Exp player who wins in his own C13 part. Let p_0, p_1, p_4, and p_f be the initial probabilities of levels 0, 1, 4 and the full foresight player in the LFLE model, then the updated distribution after a win in C13 is $(\frac{p_0}{18}, \frac{11p_1}{108}, \frac{p_4}{2}, p_f)$, and after a loss, this distribution is $(\frac{17.p_0}{18}, \frac{97.p_1}{108}, \frac{p_4}{2}, 0)$. The type-$f$ Exp player will thus choose to be the first mover in H13 iff $(\frac{p_0}{18}+\frac{11p_1}{108})/(\frac{p_0}{18}+\frac{11p_1}{108}+\frac{p_4}{2}+p_f) > 0.375$ after his opponent wins in C13, and $(\frac{17p_0}{18}+\frac{97p_1}{108})/(\frac{17p_0}{18}+\frac{97p_1}{108}+\frac{p_4}{2}) > 0.375$ after his opponent loses in C13. The type-f Inexp player makes his F/S decision just on the basis of the prior distribution on foresights given by p_0, p_1, p_4, and p_f. The $P(W)$ for the AQRE model is calculated backwards.[39] In order to win C13, the human player has to choose second mover, and then play three actions as per the perfect strategy.

Table 16: Updated Distributions Due to C13 Win or Loss

	LFLE	Level-k	AQRE
W	$(\frac{p_0}{18}, \frac{11p_1}{108}, \frac{p_4}{2}, p_f)$	Initial Distribution Maintained	$P(W)$
L	$(\frac{17.p_0}{18}, \frac{97.p_1}{108}, \frac{p_4}{2}, 0)$	$P(observation) = 0$	$1 - P(W)$

References

[1] Agranov, M., E. Potamites, A. Schotter, and C. Tergiman (2012). Beliefs and endogenous cognitive levels: An experimental study. *Games and Economic Behavior 75*(2), 449–463.

[39] The probability of a player's action from a position in C13 being perfect is given by $P_3 = [\frac{2}{3}(exp(500\lambda)/(exp(500\lambda) + 2exp(50\lambda)) + \frac{1}{3}(exp(500\lambda)/(exp(500\lambda) + exp(50\lambda))]$. The probability of the fifth stage action being perfect is given by $P_2 = exp(\lambda(P_3 500 + (1 - P_3)50)/(exp(\lambda(P_3 500 + (1 - P_3)50) + 2exp(50\lambda))$. Similarly, the probability of the third stage action being perfect is given by $P_1 = exp(\lambda(P_2P_3 500 + (1 - P_2P_3)50)/(exp(\lambda(P_2P_3 500 + (1 - P_2P_3)50) + 2exp(50\lambda))$, and P(S) $=exp(\lambda(P_1P_2P_3 500 + (1 - P_1P_2P_3)50)/(exp(\lambda(P_1P_2P_3 500 + (1 - P_1P_2P_3)50) + 2exp(50\lambda))$. Finally $P(W) = P(2^{nd}$ mover$)P_1P_2P_3$.

[2] Alaoui, L. and A. Penta (2015). Endogenous depth of reasoning. *The Review of Economic Studies*, rdv052.

[3] Binmore, K., J. McCarthy, G. Ponti, L. Samuelson, and A. Shaked (2002a). A backward induction experiment. *Journal of Economic theory 104*(1), 48–88.

[4] Binmore, K., J. McCarthy, G. Ponti, L. Samuelson, and A. Shaked (2002b). A backward induction experiment. *Journal of Economic theory 104*(1), 48–88.

[5] Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 861–898.

[6] Costa-Gomes, M. A. and V. P. Crawford (2006). Cognition and behavior in two-person guessing games: An experimental study. *The American economic review 96*(5), 1737–1768.

[7] Crawford, V. P., M. A. Costa-Gomes, and N. Iriberri (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature 51*(1), 5–62.

[8] Crawford, V. P. and N. Iriberri (2007a). Fatal attraction: Salience, naivete, and sophistication in experimental hide-and-seek games. *The American Economic Review 97*(5), 1731–1750.

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

[10] Dufwenberg, M., T. Lindqvist, and E. Moore (2005). Bubbles and experience: An experiment. *The American Economic Review 95*(5), 1731–1737.

[11] Dufwenberg, M., R. Sundaram, and D. J. Butler (2010). Epiphany in the game of 21. *Journal of Economic Behavior & Organization 75*(2), 132–143.

[12] Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics 10*(2), 171–178.

[13] Gill, D. and V. L. Prowse (2014). Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. *Journal of Political Economy, forthcoming*.

[14] Gneezy, U., A. Rustichini, and A. Vostroknutov (2010). Experience and insight in the race game. *Journal of economic behavior & organization 75*(2), 144–155.

[15] Ho, T.-H. and X. Su (2013). A dynamic level-k model in sequential games. *Management Science 59*(2), 452–469.

[16] Ivanov, A., D. Levin, and J. Peck (2009). Hindsight, foresight, and insight: an experimental study of a small-market investment game with common and private values. *The American Economic Review 99*(4), 1484–1507.

[17] Johnson, E. J., C. Camerer, S. Sen, and T. Rymon (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory 104*(1), 16–47.

[18] Kawagoe, T. and H. Takizawa (2012). Level-k analysis of experimental centipede games. *Journal Of Economic Behavior & Organization 82*(2), 548–566.

[19] Kőszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 1133–1165.

[20] Kreps, D. M. and R. Wilson (1982). Sequential equilibria. *Econometrica: Journal of the Econometric Society*, 863–894.

[21] Le Coq, C. and J. T. Sturluson (2012). Does opponents experience matter? experimental evidence from a quantity precommitment game. *Journal of Economic Behavior & Organization 84*(1), 265–277.

[22] Levin, D. and L. Zhang (2016). Bridging level-k to nash equilibrium. *Working Paper*.

[23] Levitt, S. D., J. A. List, and E. Sally (2011). Checkmate: Exploring backward induction among chess players. *The American Economic Review 101*(2), 975–990.

[24] Mantovani, M. (2014). Dems working paper series.

[25] McKelvey, R. D. and T. R. Palfrey (1998). Quantal response equilibria for extensive form games. *Experimental economics 1*(1), 9–41.

[26] Palacios-Huerta, I. and O. Volij (2009). Field centipedes. *The American Economic Review 99*(4), 1619–1635.

[27] Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, 97–109.

[28] Slonim, R. L. (2005). Competing against experienced and inexperienced players. *Experimental Economics 8*(1), 55–75.

[29] Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, 307–333.

[30] Wang, S. W., M. Filiba, and C. F. Camerer (2010). Dynamically optimized sequential experimentation (dose) for estimating economic preference parameters. *Manuscript submitted for publication*.

[31] Rampal, J. (2016). Limited Foresight and Learning Equilibrium. *Working Paper.* Available at: https://sites.google.com/site/jeevantrampalecon/Research.

Subject Instructions

Treatment 1

INSTRUCTIONS for Type D (Inexp) Subjects

This is an experiment in the economics of decision making. The Ohio State University has provided funds for conducting this research. The instructions are simple, and if you follow them carefully and make good decisions, you may earn a CONSIDERABLE AMOUNT OF MONEY which will be PAID TO YOU IN CASH at the end of the experiment.

Please pay careful attention to the instructions. If you have any questions at any time, raise your hand.

Do not communicate with other participants during the experiment. Put your phone away and make sure it is in silent mode or turned off.

Upon entry you were each assigned a unique subjectID, which you will be asked to enter several times. Please keep this carefully.

This is the second (first part was bargaining) of the two parts of the experiment. You will play 8 rounds of the AVOID Removing the 9th Item game described below. Your total payment will be a sum of your payment from the two parts of the experiment, and your show up fee ($5).

In each round of the AVOID Removing the 9th Item Game:

1. Each of you will be asked to fill in your subjectID. Please do so accurately. Then each of you will be randomly assigned a new unique participant whom you will interact with. You will not know the identity of your assigned participant. We shall refer to him/her as "the other." Decisions made by you and the other in a particular round will affect how many experimental currency units you earn in that round. You will be told if the other is of type S or type D. Type S participants have already played 12 rounds of the AVOID Removing the 9th item game. Type D participants (like you) participated in a different, completely unrelated first part of the experiment, and are now playing the second part's AVOID removing the 9th item game with other type D and type S participants.

2. Once the game begins, you will be asked to decide between being the first mover (you get to make the first choice) or the second mover (the other makes the first choice) in the subsequent task. The computer will choose one of you or the other with 50% chance each, and implement their first/second mover choice.

3. You and the other will take subsequent decisions alternately.

4. You will see all your past choices and all the past choices of the other at all points during the round.

5. You and the other will alternately choose the number of items to remove from a box containing 9 items. You can only choose to remove 1, 2, or 3 items with any given choice. You can't choose to remove 0 items. Of course, you can't choose to remove more items than there are left in the box. For example: if 7 items have been removed before your move, you can only choose to remove either 1 or 2 items at your move. The round ends when the 9th item is removed:

   (a) The one who removes the 9th item from the box receives 50 experimental currency units (ECUs) for the round.

   (b) If the other removes the 9th item, and you are the first mover, you receive 500 ECUs.

(c) If the other removes the 9th item, and you are the second mover, you receive 200 ECUs.

You will play 8 rounds of this game in the second part. In each round, the other may be of type S or type D. The type of the other will be communicated to you at all points within a round.

Your earnings for the experiment: part one earning will be equal to your earning in a randomly selected round from part one. Part two earning will be equal to your earning in a randomly selected round from part two. Your total earnings will be a sum of your show-up fee and your earnings from each of the two parts of the experiment.

The conversion rate for ECU in the experiment is: 1USD=80ECUs for the first part, and 1USD=60ECUs for the second part. Your earnings will be rounded up to the nearest dollar.

INSTRUCTIONS for Type S (Exp) Subjects

This is an experiment in the economics of decision making. The Ohio State University has provided funds for conducting this research. The instructions are simple, and if you follow them carefully and make good decisions, you may earn a CONSIDERABLE AMOUNT OF MONEY which will be PAID TO YOU IN CASH at the end of the experiment.

Please pay careful attention to the instructions. If you have any questions at any time, raise your hand.

Do not communicate with other participants during the experiment. Put your phone away and make sure it is in silent mode or turned off.

Upon entry you were each assigned a unique subjectID, which you will be asked to enter several times. Please keep this carefully.

The experiment has two parts. In both parts, you will play several rounds of the AVOID Removing the 9th Item game described below. Your total payment for the experiment will be a sum of your payment from the two parts plus your show up fee ($5).

In each round of the AVOID Removing the 9th Item Game:

1. Each of you will be asked to fill in your subjectID. Please do so accurately. Then each of you will be randomly assigned a new unique participant whom you will interact with. You will not know the identity of your assigned participant. We shall refer to him/her as “the other.” Decisions made by

you and the other in a particular round will affect how many experimental currency units you earn in that round. You will be told if they are of type S, or if they are type D. Type S will participate in several rounds of the same AVOID removing the 9th item game (described here) in both parts of the experiment. Type D will participate in a different, completely unrelated first part. After they finish the first part of the experiment, you will be informed. They will then join you in the second part of the experiment, i.e., they will also participate in several rounds of the AVOID removing the 9th item game in the second part, the same as you. You will be told when the type D subjects join and the second part begins.

2. Once the "Avoid removing the 9th item" game begins, you will be asked to decide between being the first mover (you get to make the first choice) or second mover (the other makes the first choice) in the subsequent task. The computer will choose one of you or the other with 50% chance each, and implement their first/second mover choice.

3. You and the other will take subsequent decisions alternately.

4. You will see all your past choices and all the past choices of the other at all points during the round.

5. You and the other will alternately choose the number of items to remove from a box containing 9 items. You can only choose to remove 1, 2, or 3 items with any given choice. You can't choose to remove 0 items. Of course, you can't choose to remove more items than there are left in the box. For example: if 7 items have been removed before your move, you can only choose to remove either 1 or 2 items at your move. The round ends when the 9th item is removed:

    (a) The one who removes the 9th item from the box receives 50 experimental currency units for the round.

    (b) If the other removes the 9th item, and you are the first mover, you receive 500 ECUs.

    (c) If the other removes the 9th item, and you are the second mover, you receive 200 ECUs.

You will play 12 rounds of this game in the first part and 8 rounds of this game in the second part. In the first part the other can only be of type S. In the second part, the other may type S or type D. The type of the other will be communicated to you at all points within a round. Part one earning will be equal to your earning in a randomly selected round from part one. Part two earning will be equal to your earning in a randomly selected round part two. Your total earnings will be a sum of your show-up fee and your earnings from each of the two parts of the experiment.

The conversion rate for ECU in the experiment is: 1USD=60ECUs. Your earnings will be rounded up to the nearest dollar.

Treatment 2

INSTRUCTIONS for Type D (Inexp) Subjects

This is an experiment in the economics of decision making. The Ohio State University has provided funds for conducting this research. The instructions are simple, and if you follow them carefully and make good decisions, you may earn a CONSIDERABLE AMOUNT OF MONEY which will be PAID TO YOU IN CASH at the end of the experiment. If you have any questions at any time, raise your hand. Do not communicate with other participants during the experiment. Put your phone away and make sure it is in silent mode or turned off.

The experiment has Two Components. You already completed the first (bargaining) component. In this second component, we will be using a game called "Avoid Removing the 13th Item Game." This game is played by two players. There is a box containing 13 items. The two players make choices alternately. Each player can choose to remove 1, 2 or 3 items from the box at their move. The players remove 1, 2 or 3 items alternately until 0 items are left in the box. One cannot remove 0 items or more than 3 items. Also, one cannot remove more items than are left in the box. The goal is to avoid removing the last/13th item. The player who makes his opponent remove the 13th item achieves the goal.

You are a type D subject. Type D subjects play a different bargaining game for the first component of the experiment. Type S subjects play the same round described below throughout the experiment including the first component. In this component you will play 6 rounds of the round described below.

In every round of this experiment:

Each of you will be asked to fill in your subjectID. Please do so accurately. You will be assigned a new random subject every round. You will interact with this subject for the whole round, and then randomly re-matched. You will not know the identity of your assigned subject. We shall refer to him/her as "the other" or "the human opponent". Every round will have 2 parts.

Part 1 of a Round: Play the computer

You and your human opponent will separately play the AVOID Removing the 13th Item Game (Avoid 13th for short) with the computer. *The computer plays perfectly to win. That is, if there is a choice or a sequence of choices that the computer can make to win against you, the computer will make that choice or sequence of choices without making any mistake.*

When you play the computer:

1. You will decide if you are the first mover or the second mover in the Avoid 13th game vs the computer.

2. Your choice for being the first or second mover vs the computer will be implemented. You will then commence playing the Avoid 13th game versus your computer as the first/second mover, as per your decision.

3. If you avoid removing the 13th item against the computer, you earn 500ECUs from the first part of the round, regardless of your first mover/second mover decision. If you have to remove the 13th item, you earn 50ECUs from the first part. 50ECUs is the minimum payment for participating in each part of the round.

Part 2 of a Round: Play the human opponent

1. You and the human opponent will then play the AVOID 13th Game with each other. Before the Avoid 13th game begins, both you and the human opponent will make the choice between being the first mover or second mover in the Avoid 13th game. One of your or the human opponent's first/second mover choice will be implemented with 50% chance each.

2. For the second part, if you AVOID Removing the 13th Item as the First mover, you earn 500ECUs, while if you AVOID Removing the 13th Item as the Second Mover, you earn 200ECUs. If you have to remove the 13th item, you earn 50ECUs.

After the round ends, you will be randomly re-matched with another subject.

Your earnings for a round will be the sum of your earnings from the first and the second part of the round. To calculate your earnings from this second component, one round will be randomly selected from the 6 rounds you play.

Your earnings from the experiment will be the sum of your earnings from the two components (bargaining and Avoid 13th game) plus the show-up fee ($5). The conversion rate for ECU in this second component is: 1USD=120ECUs. The conversion rate for ECU in the first component is: 1USD=120ECUs. Your total earnings will be rounded up to the nearest dollar.

INSTRUCTIONS for Type S (Exp) Subjects

This is an experiment in the economics of decision making. The Ohio State University has provided funds for conducting this research. The instructions are simple, and if you follow them carefully and make good decisions, you may earn a CONSIDERABLE AMOUNT OF MONEY which will be PAID TO YOU IN CASH at the end of the experiment. If you have any questions at any time, raise your hand. Do not communicate with other participants during the experiment. Put your phone away and make sure it is in silent mode or turned off.

The experiment has Three Components. We explain the first two components here. The third component is unrelated and doesn't affect your earning from the first two components.

In the first two components, we will be using a game called "Avoid Removing the 13th Item Game." This game is played by two players. There is a box containing 13 items. The two players make choices alternately. Each player can choose to remove 1, 2 or 3 items from the box at their move. The players remove 1, 2 or 3 items alternately until 0 items are left in the box. One cannot remove 0 items or more than 3 items. Also, one cannot remove more items than are left in the box. The goal is to avoid removing the last/13th item. The player who makes his opponent remove the 13th item achieves the goal.

Type D subjects play a different bargaining game for the first component of the experiment. You are a type S subject. That means you will play the same round described below throughout the experiment.

In every round of this experiment:

Each of you will be asked to fill in your subjectID. Please do so accurately. You will be assigned a new random subject every round. You will interact with this subject for the whole round, and then randomly re-matched. We shall refer to him/her as "the other" or "the human opponent". Every round will have 2 parts.

Part 1 of a Round: Play the computer

You and your human opponent will separately play the AVOID Removing the 13th Item Game (Avoid 13th for short) with the computer. *The computer plays perfectly to win. That is, if there is a choice or a sequence of choices that the computer can make to win against you, the computer will make that choice or sequence of choices without making any mistake.*

When you play the computer:

1. You will decide if you are the first mover or the second mover in the Avoid 13th game vs the computer.

2. You will be shown your human opponent's first mover/second mover decision in his/her interaction with his/her computer. You will then be asked about the type of your human opponent, i.e., your human opponent's type is ____(S or D). If you answer correctly, 100ECUs (Experimental Currency Units) will be added to your earning from the round. Please note that this step is only for type S subjects. Type D subjects are never shown their opponent's moves vs the computer and never asked questions about their opponent's type.

3. Your choice for being the first or second mover vs the computer will be implemented. You will then commence playing the Avoid 13th game versus your computer as the first/second mover, as per your decision.

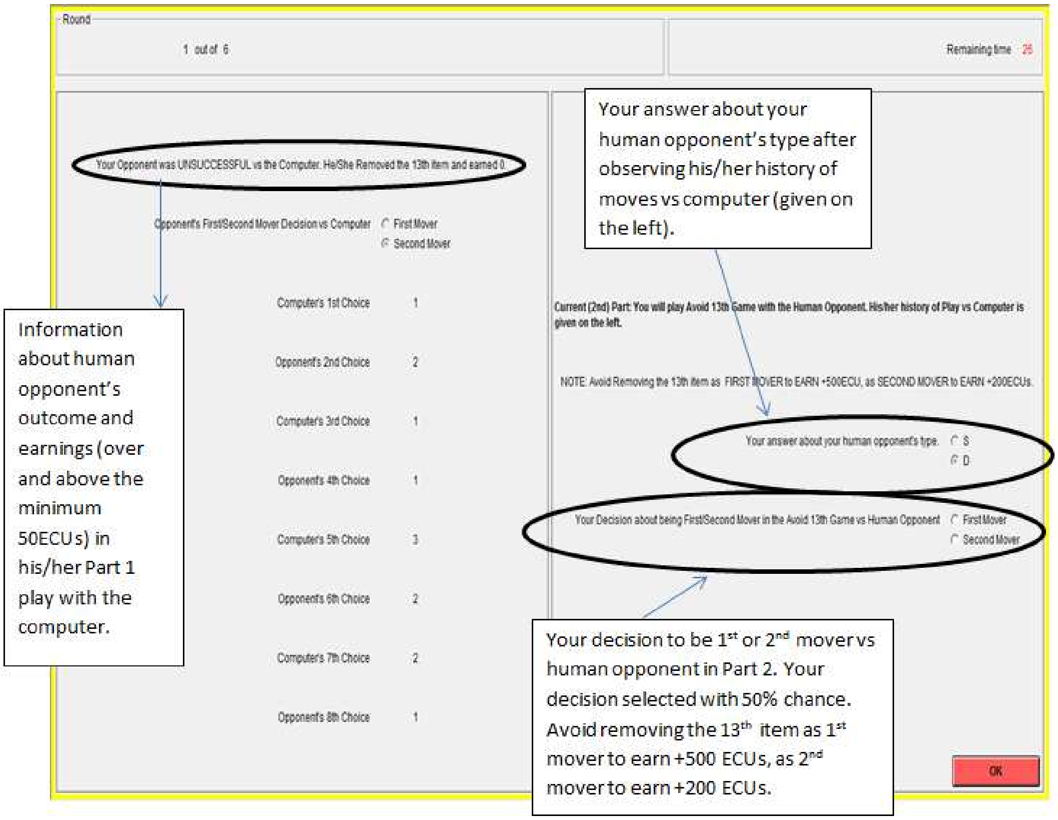
4. If you avoid removing the 13th item against the computer, you earn 500ECUs from the first part of the round, regardless of your first mover/second mover decision. If you have to remove the 13th item, you earn 50ECUs from the first part. 50ECUs is the minimum payment for participating in each part of the round.

Part 2 of a Round: Play the human opponent

1. You will be shown the complete history of your human opponent's moves vs the computer and his/her resulting outcome in the first part of that round. You will again be asked about the type of your human opponent, i.e,. your human opponent's type is ____(S or D). A correct answer will add a further 100ECUs to your earning from the round. (See the screenshot below). Please note that this step is only for type S subjects. Type D subjects are never shown their opponent's moves vs the computer and never asked questions about their opponent's type.

2. You and the human opponent will then play the AVOID 13th Game with each other. Before the Avoid 13th game begins, both you and the human opponent will make the choice between being the first mover or second mover in the Avoid 13th game. One of your or the human opponent's first/second mover choice will be implemented with 50% chance each.

3. For the second part, if you AVOID Removing the 13th Item as the First mover, you earn 500ECUs, while if you AVOID Removing the 13th Item as the Second Mover, you earn 200ECUs. If you have to remove the 13th item, you earn 50ECUs.



After the round ends, you will be randomly re-matched with another subject.

Your earnings for a round will be the sum of your earnings from the first and the second part of the round, including your earnings from the questions about your human opponent's type.

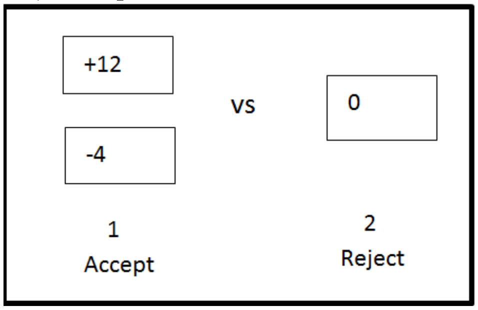
As stated above, this experiment will use three components:

First Component: Type S vs Type S. You will play 8 rounds (every round has two parts: play the computer, then play the human opponent) as described above. Every round your human opponent will be randomly redrawn from among the Type S subjects. That is, in the first component your opponent is always of type S. One of these 8 rounds will be randomly drawn to calculate your earning from this component.

Second Component: (Type S or D) vs (Type S or D). You will play 6 rounds as described above. Every round your opponent will be randomly redrawn from all the subjects in the experimental session. That is, in the second component your opponent may be of type S or type D. One of these 6 rounds will be randomly drawn to calculate your earning from this component.

Third Component: Risky choice study. This component is a short and unrelated study with 20 rounds.

In each round you will be asked to make a choice among two options. According to your choice, you may end up losing some money earned in the previous two components. Press the 1 and 2 keys to make your choice, as explained below:



Option one, called ACCEPT (on the left of screen) consists of a possible reward and a loss. If you pick this option, the computer flips a fair digital coin (chances are 50-50). In case of heads, you earn additional money. In case of tails, you lose an amount that is subtracted from your previous earnings. To choose this option, press 1. Option two, called REJECT (on the right) is one reward. If you pick this option, you will get this reward for sure. To choose this option, press 2. At the end of this third component, one of these 20 rounds will be randomly selected by the computer. Your choice in that round will determine your earning in this component. If you chose Accept in the selected round, your earning will be 20% of the amount drawn (reward or loss) from the computer's coin toss. If you chose Reject in the selected round, your earning will be 20% of the sure amount. Note: This last component will begin with a short practice with no payment. Your earnings will be the sum of your earnings from the three components plus the show-up fee ($5). The conversion rate for ECUs in the first two components is: 1USD=120ECUs. Your total earnings will be rounded up to the nearest dollar.

Bargaining Instructions: Same across treatments

This is an experiment in the economics of decision making. The Ohio State University has provided funds for conducting this research. The instructions are simple, and if you follow them carefully and make good decisions, you may earn a CONSIDERABLE AMOUNT OF MONEY which will be PAID TO YOU IN CASH at the end of the experiment.

Please pay careful attention to the instructions. If you have any questions at any time, raise your hand.

Do not communicate with other participants during the experiment. Put your phone away and make sure it is in silent mode or turned off.

Upon entry you were each assigned a unique subjectID, which you will be asked to enter several times. Please keep this carefully.

There are two parts of this experiment. You will be given instructions for part two later. Part two is unrelated to this part. During this first part of the experiment you will participate in several bargaining rounds. At the end of this part of the experiment, one of the bargaining rounds you participated in will be chosen at random for calculating your final payment.

Your total payments in this experiment will be a show up fee ($5), plus the sum of your earnings from the two parts of the experiment.

In each round, each of you will be randomly assigned a new unique participant whom you will interact with. You will not know the identity of your assigned participant. We shall refer to him/her as "the other." Decisions made by you and the other in a particular round will affect how many experimental currency units you earn in that round. There shall be a total of 8 rounds of bargaining in part one of the experiment.

In each round of bargaining:

1. The first mover of the round shall be decided randomly by the computer. Suppose the two players in a particular group are Ms. X and Mr. Y. (This is just an example, we don't match based on gender).

2. The first mover, say, Ms. X, will decide how to split a "pie" of 1000 Experimental Currency Units (ECUs) among herself and the other, say, Mr. Y. In particular, she will specify how much of the pie she wants to give to the other as her "offer" and how much her "demand" is. Her "offer" and "demand" can total up to no more that the size of the pie, i.e., 1000 ECUs. If her "offer" plus "demand" is more than 1000ECUs, the computer will guide her so that she doesn't make that mistake.

3. The other, Mr. Y, will then view this "offer" to him and "demand" for Ms. X herself and decide whether to Accept this offer or to Reject it with an offer and demand of his own. If Mr. Y accepts the first offer, (then Mr. Y's offer and demand entry are meaningless as the round ends) the round

ends and the first offer becomes Y's earnings for the round, while Ms. X gets her first demand as the earnings for the round. If Mr. Y doesn't accept Ms. X's first offer then he rejects the first offer with an offer for Ms. X and a demand for himself out of a reduced pie of 600ECUs. Again, this offer and demand can't add up to be more than 600ECUs. If they do, the computer will guide Mr. Y so that he can't make that mistake.

4. Ms. X will then decide whether to Accept this second offer or to Reject it with a last (third) offer for Mr. Y and a demand for herself out of a pie of 360ECUs. If Ms. X accepts the second offer, the round ends. Ms. X's last (third) demand and offer can't add up to be more than 360ECUs. If they do, the computer will guide Ms. X so that she can't make that mistake. If Ms. X accepts Mr. Y's offer then her last offer and demand are meaningless, as she will get Mr. Y's offer and Mr. Y will get his own demand.

5. If Ms. X rejects the second offer and makes an offer and demand of her own, then Mr. Y will then decide whether to Accept Ms. X's third and last offer or to Reject it.

6. If Mr. Y accepts this last offer, then the round ends and the last offer becomes his earnings for the round, while Ms. X gets her last demand as the earnings for the round.

7. If Mr. Y rejects the last offer, then both X and Y get 0ECUs for the round and the round ends.

The experiment will use ECUs that will be converted to USD at the rate of 1USD = 80ECUs for this part of the experiment. In addition, you will receive USD5 as your show-up fee for the whole experiment, and your earning from second part (whatever the earning may be). You are not allowed to talk to another participant. Please feel free to ask questions during the experiment by raising your hand. I will come to you and answer your questions. Good luck!