DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

ITERATED DOMINANCE AND ITERATED BEST-RESPONSE IN EXPERIMENTAL 'P-BEAUTY CONTESTS'

Teck Hua Ho
The Anderson School
UCLA

Keith Weigelt
The Wharton School
University of Pennsylvania

Colin Camerer
Division of Social Sciences
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 974

July 1996

Iterated Dominance and Iterated Best-Response in Experimental 'P-Beauty Contests'

Teck Hua Ho Keith Weigelt Colin Camerer

Abstract

We study a dominance-solvable 'p-beauty contest' game in which a group of players simultaneously choose numbers from a closed interval. The winner is the player whose number is the closest to p times the average, where $p \neq 1$. The numbers players choose can be taken as an indication of the number of steps of iterated reasoning about others they do. Choices in the first period show that the median number of steps of iterated reasoning is either one or two. Repeating the game produces reliable convergence to the unique Nash equilibrium. Choices in later periods are consistent with subjects' best-responding to previous choices, or iterating one step and best-responding to best responses. (Choices are not as consistent with 'learning direction theory' which embodies elements of belief-free reinforcement models). Variation in the values of p, the number of players, and whether subjects played a similar game before, all affect choices and learning.

Iterated Dominance and Iterated Best-Response in Experimental 'P-Beauty Contests'*

Teck Hua Ho Keith Weigelt Colin Camerer

1 Introduction

Picture a thin country 1000 miles long, running north and south, like Chile. Several natural attractions are located at the northern tip of the country. Suppose each of n resort developers plan to locate a resort somewhere on the country's coast (and all spots are equally attractive). After all the resort locations are chosen, an airport will be built to serve tourists, at the average of all the locations and the natural attractions. Suppose most tourists visit all the resorts equally often, except for lazy tourists who visit only the resort closest to the airport; so the developer who locates closest to the airport gets a fixed bonus of extra visitors.

Where should the developer locate to be nearest to the airport?

The surprising game-theoretic answer is that all the developers should locate exactly where the natural attractions are. This answer requires at least one natural attraction at the northern tip, but does not depend on the fraction of lazy tourists or the number of developers (as long as there is more than one).[1]

To see how this result comes about, denote developers' choices by mileage numbers on the coastline (from 0 to 1000) as $x_1, x_2, \ldots, x_n$. Locate all m of the natural attractions at 0. Then, the average location is $A = \frac{x_1+x_2+\ldots+x_n}{n+m} = \frac{n}{n+m} \cdot \bar{x}$. If we define the fraction $\frac{n}{n+m}$ as p (and note that $p < 1$ as long as $m \geq 1$), then the developer who is closest to A, or $p \cdot \bar{x}$, wins a fixed amount of extra business (from the lazy tourists).

*Teck Ho and Colin Camerer were sponsored in part by National Science Foundation Grant SBR 95-11137 and Keith Weigelt by Wharton's Reginald Jones Center for Policy, Strategy, and Organization. We thank two anonymous referees, Rosemarie Nagel, Lisa Rutstrom, Dale Stahl, seminar participants at the University of Pennsylvania, UCLA, and the MacArthur Foundation Preferences Group, for helpful comments, and Hongjai Rhee for tireless research assistance.

[1] Douglas Gale (1996, section 4) describes a related class of "dynamic coordination" games in which the returns to investing at time t can depend on how much others invest after t. For example, a firm pioneering a new product standard benefits by subsequent entrants using the same standard. In equilibrium, all firms invest immediately.

This game was first discussed by Herve Moulin (1986, p. 72) and studied experimentally by Rosemarie Nagel (1995). The game can be solved by iterated application of dominance. The largest possible value of A is $1000 \cdot p$ so any choice of x above $1000 \cdot p$ is dominated by choosing $1000 \cdot p$. If developers believe others obey dominance, and therefore choose $x_i < 1000 \cdot p$, then the maximum A is $1000 \cdot p^2$ so any choice larger than that is dominated. Iterated application of dominance yields the unique Nash equilibrium, which is for everyone to locate at zero. No matter where the average of the other developers' locations is, a developer wants to locate between that average and the natural attractions (which is where the airport will be built); this desire draws all the developers inexorably toward exactly where the attractions are.

We call these 'p-beauty contest' games[2] because they capture the importance of iterated reasoning John Maynard Keynes described in his famous analogy for stock market investment (as Nagel, 1995, pointed out). Keynes (1936, pp.155-156) said

> ...professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole.... It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practise the fourth, fifth and higher degrees.

In Keynes's 'newspaper competition' people want to choose exactly the same faces others choose. Professional investment is not quite like this. Think of the time at which investors choose to sell a rising stock as picking a number. When many investors choose to sell, the stock crashes; the time of the crash is around the average number (selling time) chosen. Then professional investment is a p-beauty contest (with $p < 1$) in which investors want to sell a few days ahead of the crash—picking a number equal to p times the average number—but not too far ahead.

Our paper reports experiments on p-beauty contest games. These games are ideal for studying an important question in game theory—how many iterations of dominance to players apply? The games are also useful for studying learning. Both iterated dominance and learning have broad implications for economics.

Our central contribution is application to p-beauty contests of a pair of structural models—a model of first-period choices, and a separate model of learning—similar to those used recently by Dale Stahl and Paul Wilson (1994, 1995), Debra Holt (1993), and

[2]The term 'guessing game' has some precedence (it is used in two published papers) but does not distinguish these games from most others. We introduce the name 'p-beauty contest' because the name is apt and the Keynesian passage is well-known.

others. In the model of first-period choices, players are assumed to obey different levels of iterated dominance. We estimate from the data what the distribution of the different levels is most likely to be. In the learning model, players are assumed to use different levels of iterated best response. These models give a parsimonious way to empirically characterize the levels of iterated dominance and iterated best-response.

1.1 Iterated dominance

Iterated dominance is perhaps the most basic principle in game theory. Games in which iterated application of dominance determine a unique equilibrium are called 'dominance-solvable'. Some games in this class include finitely-repeated prisoners' dilemma, 'centipede' (e.g., Richard McKelvey and Thomas Palfrey, 1992), the 'electronic mail game' (Ariel Rubinstein, 1988), Cournot duopoly, and others.

In a dominance-solvable game, reaching the equilibrium requires some minimal number of steps of iterated dominance (which we call the rationality-threshold). There are good reasons to doubt that players behave as if they have more than a couple of steps of iterated rationality. For example, iterated reasoning is cognitively difficult. And high levels of iterated rationality are not easily justified by natural selection arguments (see Stahl, 1993). [3]

Furthermore, most experimental studies indicated limited application of iterated dominance, perhaps 1–3 steps. The p-beauty contest is a sharper instrument for measuring iterated dominance than other games, however, because it is a constant-sum game so apparent violations of dominance due to altruism, cooperativeness, etc., are less likely.

The level at which iterated dominance is violated is central for two questions in economics. First, many economists (e.g., Charles Plott, 1992) argue that phenomena which appear irrational could be due to rational players expecting some others to behave irrationally. Potential examples include cooperation in finitely-repeated prisoners' dilemma games, the winner's curse, escalating bids in the dollar auction, and price bubbles in experimental markets (e.g., David Porter and Vernon Smith, 1995). The p-beauty contest game tests this conjecture directly by measuring the fractions of players who violate dominance, violate one step of iterated dominance, etc. (Some players do violate dominance, so the strong form of the conjecture is false.)

Second, Ho and Weigelt (1996) show that subjects playing coordination games with multiple equilibria appear to use rationality-threshold as a selection principle (they coor-

[3] Natural selection works if those who are rational do better, and those who are not get eliminated by natural selection. However, in games with unique dominance-solvable equilibria the equilibrium strategy is only a best response to all other strategies if the rationality-threshold is one (as in one-shot prisoners' dilemma.) Therefore, players who choose nonequilibrium strategies, but stay one step ahead of the equilibration process, will be selected for and players who immediately choose equilibria will be selected against. If there is a steady flow of naive new players, those players who choose equilibria 'too quickly' will gradually die out.

dinate on lower-threshold equilibria). The levels of iterated dominance needed to reach equilibria could therefore help predict which equilibria will be selected. Equilibrium selection is a long-standing problem in economics because of multiple equilibria in models of search and trading, the macroeconomy, intrafirm investment, etc.

1.2 Learning and iterated best-response

The predominant view in modern game theory is that equilibria in all but the simplest games are reached by a learning or evolutionary process rather than by reasoning. Many processes have been studied (e.g., Paul Milgrom and John Roberts, 1991; Alvin Roth and Ido Erev, 1995; Kenneth Binmore, John Gale and Larry Samuelson, 1994) but more careful empirical observations are needed to judge which rules describe learning best.

While P-beauty contests are uniquely suited to studying iterated dominance, they are also a useful tool for studying learning empirically for three reasons. First, even if players do not choose the Nash equilibrium initially, choices are likely to move toward that point. Precisely because convergence is not immediate, there is healthy variation in the data which we can use to estimate which adaptive dynamics fit the convergence process best. Second, because of the structure of the game, adaptive learners, who simply learn from past observations, choose different numbers than sophisticated learners who realize others are adapting and then best-respond to them. The difference in choices by adaptive and sophisticated types can be used to estimate the proportions of those types in the subject population. Third, since choices are steadily converging toward a prediction at the boundary of the strategy space (either 0 or 200), approaches which assume players just repeat previously-successful strategies are likely to lag behind the subjects' actual learning. So these games are good for pitting reinforcement approaches against more complex learning rules. (That competition is not conducted here, but see Stahl, in press).

Our experiments use several variants of the p-beauty contest game. First, we compare 'finite-threshold' games (with $p > 1$) in which the equilibrium can be reached in a finite number of steps of iterated dominance, with 'infinite-threshold' games with $p < 1$, in which the equilibrium can not quite be reached in finitely-many iterations of dominance. (The developer-location game mentioned in the introduction is an infinite-threshold game.) The finite-threshold case is useful because it helps bound the number of steps of iterated dominance people use naturally.

Two other comparisons, between different group sizes n and values of p, are used to study whether the number of levels of iterated dominance (and levels of best-responding) are robust across parameter variations. Group size is particularly interesting because players in smaller groups exert more influence on the mean number; if they recognize this, they should choose lower numbers and converge more quickly.

To study learning, we estimate a structural model in which some fraction of level-0' players choose a weighted average of previous winning numbers, other players best-

respond to the winning numbers in the previous periods, other players best-respond to anticipated best responses, and so on. Versions of some familiar learning rules (e.g., fictitious play and Cournot best-response dynamics) are nested within this class. We also estimate a structural form of the learning direction' theory tested by Nagel (1995), and earlier by Reinhard Selten and Rolf Stoecker (1986).

1.3 Key Results

Our experiments and analysis yield several key results:

- First-period choices are widely distributed and far from equilibrium, but subsequent choices converge to equilibrium (particularly in the finite-threshold game).
- First-period choices are consistent with a median of 2 steps of iterated dominance in the infinite-threshold game, and 1 step in the finite-threshold game. The estimated proportions are spread across levels 0–3 (at least 10% in each).
- Choices after the first period are consistent with learning rules in which a large fraction of level 0' subjects choose a weighted average of previous winning numbers (weighting the previous winner most strongly). Half the subjects (level 1s) best-respond to the level 0's.
- Subjects with experience in a previous game generally have higher best-response levels. The parameter estimates are also sensitive to both p and the group size.

2 The p-beauty Contest Game

In our experiments, a group of n subjects simultaneously choose a number from a closed interval $[L, H]$. The subject whose number is closest to p times the group mean wins pi. Denote the winning number by $w = p \cdot \bar{x} = p \cdot \frac{x_1+x_2+...+x_n}{n}$. Subjects' payoffs are determined as follows. Denote the set of winners I^* to be $argmin_i\{|x_i - w|\}$ (the set of players whose choices are closest to w). Each winner $i \in I^*$ obtains a monetary prize of $\frac{n \cdot \pi}{|I^*|}$ and the remaining group members receive nothing. Variants of this game are denoted by $G([L, H], p, n)$.

Consider two variants of this game with n players: 1) finite threshold $FT(p, n) = G([100, 200], p, n)$ and 2) infinite threshold, $IT(p, n) = G([0, 100], p, n)$. While both games have an unique dominance solvable equilibrium, $IT(p, n)$ requires an infinite level of iterated reasoning to solve the game whereas $FT(p, n)$ requires only a finite level. Figures 1a-b illustrate this with $FT(1.3, n)$ and $IT(0.7, n)$. In these figures, the level of iterated rationality is indicated by $R(i)$. For example, $R(2)$ means that subjects are rational and know that others are rational. Figure 1a shows that the threshold level needed to solve $FT(1.3, n)$ is 3. Subjects with zero levels of iterated rationality may

choose numbers from $[100, 130)$ (i.e., $R(0)$[4]). Rational players will choose a number from [130, 200] because 130 dominates any number in [100, 130)(i.e., $R(1)$). (This illustration assumes a very large number of players, so that players can ignore their own effect on the mean and winning number). Mutually rational players deduce it is in their interests to choose a number from [169, 200] (i.e., $R(2)$). To guarantee all subjects choose the unique equilibrium of 200 requires a threshold level of 3 (since a subject at $R(2)$ could choose less than 200). When $p = 1.1$, the threshold level is 8.

In Figure 1b, the threshold level for the $IT(0.7, n)$ game is infinite. Rational players will only choose a number from [0, 70] because any number in (70, 100] is dominated by 70 (i.e., $R(1)$). Applying the same reasoning, mutually rational R(2) players will only choose a number from [0, 49] (i.e., $0.7 \cdot 70$). With $k-$levels of iterated rationality, players will pick numbers from $[0, 0.7^{k+1} \cdot 100]$. All players will choose the unique equilibrium of 0 only if iterated rationality is infinite (i.e., $0.7^k \to 0$ as $k \to \infty$).

3 Experimental Design

To investigate the degree of iterated rationality required to each equilibrium we studied games with infinite (IT) and finite (FT) thresholds, and different values of p. The design also varied the group size n, to test whether smaller or larger groups behave differently. In addition, to study whether learning transfers across different games, each subject played one IT game and one FT game (counterbalanced for order). Experiments were conducted using the following game pairs.

$$IT(0.7, 3), FT(1.3, 3)$$
$$IT(0.9, 3), FT(1.1, 3)$$
$$IT(0.7, 7), FT(1.3, 7)$$
$$IT(0.9, 7), FT(1.1, 7)$$

Each game pair consists of one FT and one IT game. Groups of 3 were randomly assigned to one of the first two pairs and groups of 7 to one of the last two pairs. To maintain a constant expected payoff per subject across groups of different sizes, the payoff for winning was directly proportional to the group size. The expected payoff per subject per round was \$0.50.[5]

[4]Irrational players may also choose a number outside of [100, 130) by chance, thus the number of players choosing between [100,130) is a lower bound on the number of $R(0)$ players.

[5]The standard deviations of payoffs are 3.22 for IT(p,7), 2.48 for FT(p,7), 2.12 for IT(p,3), and 1.95 for FT(p,3). If subjects are equally-skilled, the theoretical standard deviations are 3.89 ($n = 7$) and 2.24 ($n = 3$). Note that there is less variation in actual payoffs than predicted by the equal-skill benchmark. Most of the difference is due to the fact that subjects shared the prize in the event of a tie. Simulating the standard deviation of actual payoffs that would result if the whole prize was given to a randomly-chosen subject in the event of a tie yields standard deviations are extremely close to equal-skill benchmark. Note that this finding casts some doubt on models in which players have persistent differences in skill, effort, or reasoning ability, etc. that create payoff differences. Or the standard deviation of payoffs may not be sensitive enough to detect individual differences.

We ran 55 experimental groups with a total of 277 subjects. There were 27 groups of size 3, and 28 groups of size 7. Table 1 summarizes our experimental design.[6]

Subjects were recruited from a business quantitative methods class at a major undergraduate university in Southeast Asia. They were assigned to experimental sessions randomly and each participated in one session.

A typical session was conducted as follows. Subjects reported to a room with chairs placed around its perimeter, facing the wall, so subjects could not see the work of others. Subjects were randomly assigned seats, subject numbers, and given written instructions (see Appendix). After all subjects were seated, an administrator read the instructions aloud, and subjects were given the opportunity to publicly ask questions. During the experiment, subjects were not allowed to communicate with each other. Before round 1 began, all subjects were publicly informed of the relevant number range $[L, H]$, and the value of p. Then round 1 began. Subjects were asked to choose a number from the relevant range $[L, H]$, and record this number on a slip of paper. An administrator then collected the responses of all subjects, calculated the average number of the group and publicly announced this number. Payoffs were then privately announced to each subject for that round (the winner(s) received a positive payoff, all others received \$0).[7] Then the next round began. All rounds were identical, and the game lasted for 10 rounds[8]. After the 10th round was completed, subjects participated in a second 10 round game. The order in which the games were played was determined randomly. Subjects remained in the same group, but the number range [L, H], and the parameter p were both changed. The experimental procedure remained the same. After this second 10 rounds, subjects summed their earnings over all 20 rounds, and were paid their earnings in cash. Experiments lasted approximately 40 minutes, and subjects earned on average \$10.00.

4 Basic Results

This section summarizes basic results. Later sections report more refined estimates of the number of levels of iterated dominance subject use, and estimates of learning models.

Result 1: First-period choices are far from equilibrium, and roughly normally distributed

[6]Raw data are available from the authors.

[7]A referee wondered whether announcing who won each round might be a poor design choice in this game(though announcing payoffs regularly is the standard protocol in experimental economics). The concern is that wealth effects alter incentives and create surprising outlying 'spoiler' responses of 100. In section 4 we point out that there is no systematic evidence that these spoiler choices came from subjects who were satiated in money from winning repeatedly. Furthermore, a good reason for announcing payoffs is that payoff history is necessary to provide reinforcement and allow tests of reinforcement learning models.

[8]Each subject played the game with the same group members for all 10 rounds. Since the game was constant-sum, there was no reason for subjects to develop reputation or tacitly collude to increase overall payoffs.

around the interval midpoint. Choices converge toward equilibrium over time.

Figures 2a-h (IT) and 3a-h(FT) show histograms of the frequencies of choices by subjects in each condition. Only 2.2% of the subjects chose the equilibrium in the first period. Most first-period distributions are sprinkled around the interval midpoint. Choices converge toward the equilibrium in later periods.

Table 2 summarizes the degrees of iterated dominance suggested by number choices. The table has four panels, one for each value of p, adding 3- and 7-person groups and experience levels together for each p. Each row reports the number of choices violating each level of iterated dominance (in conjunction with lower levels). The last row of each panel reports the number of players who chose the unique equilibrium prediction.[9] For example, in condition $FT(1.3, n)$ in round 1, 35 subjects out of 140 (25%) exhibited zero levels of iterated dominance because they chose numbers in the interval $[100, 130)$, 68 subjects (48.6%) exhibited only one level of iterated dominance, and so forth.

Table 2 shows that substantial numbers of subjects violate each of the lowest levels of iterated dominance (and consequently, very few choose the equilibrium), particularly in earlier rounds. (Later rounds are consistent with higher levels of iterated dominance, but that is probably learning rather than more sophisticated iterated reasoning per se.) Section 5 below gives more precise estimates.

Result 2: Choices are closer to equilibrium for games with finite-thresholds, and for games with p further from 1.

Comparing the Figure 2 and 3 histograms shows that choices are closer to equilibrium across most rounds in the finite-threshold games, compared to infinite-threshold games. For example, look at Figures 2a and 3a: Starting from the second round the proportion of equilibrium choices in the FT game (Figure 3a) begins to grow toward 100%, while convergence is much slower in the IT game (Figure 2a).[10] Pooling rounds, the frequencies of equilibrium play are highly significantly different in FT and IT games (51.6% vs. 4.9%, $(\chi^2 = 1493, p < 0.001)$.

In addition, more choices are at equilibrium in games with p further from zero ($p = 1.3$ vs. $p = 1.1$, $\chi^2 = 171.6, p < 0.001$); and $p = 0.7$ vs. $p = 0.9 (\chi^2 = 5.9, p < 0.05)$. [11] Analyses-of-variance (ANOVA) using each group's mean choice for the first or last five rounds also shows a strong effect of p, at significance levels from .000 to .08, for each of the two groups of rounds (1–5 or 6–10) and threshold levels. For example, means for $p = .7$ are lower than means for $p = .9$ for the first five rounds (31.57 vs. 44.66) and the last five rounds (17.76 vs. 27.83).

[9] Levels of rationality higher than the threshold are indistinguishable, and pooled in 'Equilibrium Play'.

[10] Note that the tall bars in the back corner of Figures 2 represent frequent choices of 1–10, not zero.

[11] While the threshold levels for both games are infinite, the level of iterated rationality needed for any number in the closed interval is higher when $p = 0.9$. For instance, to reach 10, $IT(n)$ requires a level of iterated rationality of 6 where $IT'(n)$ requires a level of 21. This implies that threshold convergence is faster for IT relative to IT'.

Result 3: Choices are closer to equilibrium for large (7-person) groups than for small (3-person) groups.

Figures 2a and 2e illustrate the typical effect of group size: Larger groups (Figure 2a) choose higher numbers at the start, and converge to equilibrium much more quickly than the small groups (Figure 3a).

Across rounds, the proportions of equilibrium play by subjects in 3- and 7- person groups are significantly different in finite-threshold games (39.6% vs. 56.6%, $\chi^2 = 67.3, p < 0.001$) and marginally significant in infinite-threshold games (3.7% vs. 5.4%, $\chi^2 = 3.4, p < 0.10$). Means of larger groups also start closer to the equilibrium for the FT games and, for both kinds of games, lie closer to the equilibrium in every round. ANOVAs on group means aggregated over rounds 1–5 or 6–10 show highly significant differences across group sizes in comparisons for all values of p and experience levels (p-values range from .003 to .014).

The group size effect goes in a surprising direction because each member of a small group has a larger influence on the mean and should choose closer to equilibrium if they take account of this. For example, if $p = .7$ and you think others will choose an average of 50, you should choose the solution to $C = .7 \cdot \frac{C+(n-1)\cdot 50}{n}$, which is 30.4 if $n = 3$ and 33.3 if $n = 7$. But large groups choose lower numbers. Perhaps, as a referee suggested, adjusting for n takes extra thought which limits the number of steps of iterated reasoning subjects do. This represents an interesting puzzle for future research.

Result 4: Choices by experienced subjects are no different than choices by inexperienced subjects in the first round, but converge faster to equilibrium.

Figures 1a-b illustrate the effects of experience in IT games. Experienced subjects (who previously played an FT game with an equilibrium of 200) choose similar numbers to inexperienced subjects in the first round, but converge much faster.

Two types of learning transfer by experienced subjects can be distinguished: 'Immediate transfer' (if choices are closer to equilibrium in the first round), and 'structural transfer' if convergence is faster across the ten rounds.[12]

In general, there is little immediate transfer because experienced subjects' choices in the first round are not much different than the choices of inexperienced subjects.[13] But there is some evidence of structural transfer because ANOVAs show that group means of experienced subjects are closer to equilibrium in the first 5 rounds ($F = 4.60, p < .04$)

[12]Without detailed cognitive theory, there are no firm ground on which to make sharper predictions. We note that psychological literature on transfer is generally pessimistic about transfer of 'deep structure of similar problems (e.g., Mark Singley and John Anderson, 1989). However, this result and further analyses reported in section 6 show that subjects do converge more quickly with experience, hinting at some deep structure learning.

[13]The only exception is the $p = .7$ case where experienced subjects are further from equilibrium, exhibiting 'negative transfer'. For $p = .7$, $t = -2.14, p < 0.01$ for $n = 7$ and $t = -2.15, p < 0.01$ for $n = 3$.

and about the same in the last 5 rounds ($F = .07, p < .80$) for FT games, and closer to equilibrium in IT games ($F = 2.62, p < .11$ and $F = 15.67, p < .001$). And overall, the proportion of equilibrium play is significantly higher for experienced subjects ($\chi^2 = 28.2$ (FT) and $\chi^2 = 125.6$ (IT), both $p < 0.001$)

Result 5: Extreme choices of 100 ('spoilers') occur occasionally, often after previous losses, and are seldom repeated.

A small, eye-catching feature of the histograms in Figures 2-3 are choices of extreme numbers close to 100. We focus on choices of exactly 100. There are 87 such spoiler choices (3.1%) in FT games and 52 (1.9%) in IT games. (Nagel, 1995, reports .9%.) Most were chosen by subjects who spoiled only once (88.6%). While spoilers are rare, it is natural to wonder what causes them because a choice of 100 is as far from the equilibrium as possible.

There are several plausible explanations. Incentives may change if repeat winners become satiated in money (or altruistically want to let others win), or if previous losers become frustrated and express their annoyance by making the wrongest possible choice (hence, the term 'spoiler'). Or subjects may be choosing 100 hoping to win, naively thinking they can alter the mean enough to win if they choose the most extreme number. Alternatively, they may be trying to strategically manipulate the learning process—knowing others respond to previous winners, they try to spoil the learning process by changing the mean temporarily, so they can win when others overreact to the unusual mean.[14]

Facts about spoilers help distinguish these explanations. Spoiling choices are evenly distributed across rounds and are equally common in the last round (casting doubt on the learning-manipulation theory).[15] Spoilers tended to follow low-payoff rounds, and did not raise average payoffs.[16] These facts reject the altruism explanation, which incorrectly predicts spoiling in later rounds and following high payoffs. The facts are consistent with frustration, since spoiling tends to follow low payoffs, and naive attempts to win by raising the mean, which generally fail and are quickly abandoned.

In the analyses below, we both include and exclude spoilers. (An alternative is to estimate the proportion of these 'level -1' types as Stahl, in press, did.) Some parameter estimates are affected, but the basic conclusions are unchanged when spoilers are excluded.

[14]In John Duffy and Nagel, 1996, and in further experiments of ours, we set the winning number to p times the median, which is less easily influenced by spoilers than the mean. There are fewer spoilers in these experiments, consistent with the naive and strategic-manipulation explanations.

[15]66 spoilers occur in rounds 1-5, 73 in round 6-10, and 16 in the last round. The average round numbers in which spoiling first occurs are 6.11 and 4.78 for IT and FT games, which are only slightly different from the average round number of 5.5 (t=1.31 and -2.23).

[16]The average per-round earnings in pre-spoiling rounds were \$.37 and \$.47 in IT and FT games, marginally below the expected earnings of \$.50 (t=-2.43, -1.31). Post-spoiling earnings averaged \$.29 and \$.45, substantially less than expected profits (t=-4.20, -1.51) but not significantly different than pre-spoiling profits (t=-1.05,.36).

5 Further Results: Levels of Iterated Dominance

The simplest method for approximating the level of iterated dominance revealed by choices is to count the number of choices in each of the intervals $[0, p^{k+1} \cdot 100]$ (or the corresponding intervals when $p > 1$). (These figures were reported above in Table 2.)

For example, since 6.6% of the subjects chose numbers in $(90, 100]$ in $IT(0.9, n)$ game, then we can conclude that <u>at least</u> 6.6% violated dominance. Since 5.1% of the subjects chose numbers in $(81, 90]$, we can conclude that at least 5.1% of the subjects violate the conjunction of dominance, and one level of iterated dominance. (Or put differently, we can be sure these subjects are not performing two or more levels of iterated dominance.)

But these numbers are simply a lower bound on the rates at which various levels of rationality are violated. The bounds cannot be tightened without using some method for distinguishing how many of the 5.1% subjects in the interval $(81, 90]$, for example, are violating dominance and how many are obeying dominance but violating one step iterated dominance. If these subjects could be assigned to one level or the other, the estimates of levels of iterated dominance revealed by choices can be sharpened considerably.

The method we use posits a simple structural model of how dominance-violating 'level 0' players choose, assume that 'level 1' players obey dominance but believe that others are level 0 players, etc. (as in Stahl and Wilson, 1994, 1995).

We begin with the assumption that level 0 players choose numbers randomly from a truncated normal density with mean μ and variance σ^2.[17] Level L players are assumed to believe that all other players (besides themselves) choose from the level $L-1$ distribution $B_{L-1}(x)$. Believing this, they mentally simulate $n-1$ draws from the level $L-1$ distribution and compute the average of those draws. (For reasons we explain below, assume they allow these draws to be correlated with correlation ρ.) Then they choose p times the average (including their own choice), giving a distribution that satisfies

$$B_L = \frac{p}{n} \cdot (B_L + \sum_{k=2}^{n} B_{L-1}^k) \tag{1}$$

where B_{L-1}^k is the k-th draw from random variable B_{L-1}. This gives a random variable for level L players' choice

$$B_L = \frac{p}{n-p} \cdot \sum_{k=2}^{n} B_{L-1}^k \tag{2}$$

[17] We also tried a uniform distribution over all possible number choices, but the uniform almost always fit worse than the normal. Also, the normal distributions always have some mass m at numbers outside the permissible interval. We truncate the distribution at the endpoints and assume that all mass outside the endpoints is piled up at the nearest endpoint. An alternative technique is to normalize the distribution within the endpoints, dividing the distribution by 1-m. The latter technique did not produce reliable maximum-likelihood estimates in many cases, especially for $p > 1$.

Notice that since the level 0 distribution is truncated at 100 (for the $p < 1$ case), the level 1 distribution is automatically truncated at $[\frac{p\cdot(n-1)}{n-p} \cdot 100]$; so level 1 players never violate dominance. Similarly, since the level 2 distribution is truncated at $[\frac{p\cdot(n-1)}{n-p}]^2 \cdot 100$, level 2 types never violate dominance and never violate one step of iterated dominance.[18]

It is easy to show that the mean and variance of B_L obey the following recursive relationships:

$$E(B_L) = \frac{p \cdot (n-1)}{n-p} \cdot E(B_{L-1}), \tag{3}$$

$$Var(B_L) = \frac{p^2}{(n-p)^2} \cdot [(n-1) + 2\rho \cdot \frac{(n-1)\cdot(n-2)}{2}] \cdot Var(B_{L-1}). \tag{4}$$

An important feature of this model is that if $p < 1$, as the level L rises, the variance in the distribution of choices B_L falls[19] (because the term $\frac{p^2}{(n-p)^2} \cdot [(n-1) + 2\rho \cdot \frac{(n-1)\cdot(n-2)}{2}]$ is less than one). The variance falls because the players are assumed to take an average of $n-1$ other players' choices, which will have less variance than an individual choice. This implies that the level 1 players' distribution will be rather narrow, the level 2 players' distribution narrower still, and so forth. Allowing higher-level players to perceive a nonzero correlation among (simulated) choices by lower-level players ρ slows down the rate of reduction in $Var(B_L)$ with L, and turns out to fit the data much better than the restriction $\rho = 0$.

The assumptions above give a distribution of first-period choices by each of the level types. The crucial problem is how to 'assign' a level type to players who choose numbers x_i that different types might choose. Take the $p = .7$ case as a clarifying example. Suppose a player chooses 63. This choice could come from a level 0 player or from a level 1 player. We assign this choice to a level 0 type iff a level 0 type is more likely to have made that choice than a level 1 type (i.e, iff $B_0(63) > B_1(63)$). (Note that a level 2 type would never pick 63, i.e., $B_2(63) = 0$, and similarly for higher-level types.)

Put more formally, assume that a fraction ω_L of the players are of level L, and ω_L^b is the fraction of choices assigned to L in each level-of-dominance interval or 'bin' b. The

[18]One criticism of this method is that it assumes all players think they are 'smarter (or reason more deeply) than others. While this is logically impossible, it is consistent with a large body of psychological evidence showing widespread overconfidence about relative ability (see, e.g., Dan Lovallo and Camerer, 1996). An alternative approach includes some degree of 'self-consciousness': Level L types to believe that a fraction ω_L of others are level L types like themselves, then perhaps impose (or test) the rational expectations assumption that the perceived ω_L is close to the econometrician's best estimate, given the data (as in McKelvey and Palfrey, 1992; McKelvey and Palfrey, 1995, 1996). The empirical problem with self-consciousness in p-beauty contests is that players who think enough others are like themselves are inexorably led toward the Nash equilibrium; but choices are very far from the equilibrium.

[19]In general, the same thing can be said for the case if $p > 1$ as long as ρ is small. However, if ρ is close to 1, then we can have $Var(B_L) > Var(B_{L-1})$.

total proportion of level L types is $\omega_L = \sum_{b=0}^{L_m} \frac{N_b \cdot \omega_L^b}{N}$ (where L_m is the maximum level estimated (3 in our analyses)and N_b is the number of observations in bin b). Define the observations x in bin b by those x which satisfy $[\frac{p(n-1)}{n-p}]^{b+1} \cdot 100 < x \leq [\frac{p(n-1)}{n-p}]^b \cdot 100$. (There are a total of $L_m + 1$ bins). For observations in bin b, the distribution function is

$$B(x) = \sum_{L=0}^{b} \omega_L^b \cdot B_L(x) \tag{5}$$

Of course, $\sum_{L=0}^{L_m} \omega_L = 1$. Then the log-likelihood of observing a sample $(x_i, i = 1, \ldots, N)$ is:

$$LL_1(\mu_1, \sigma_1, \rho, \omega_L^b; b = 0, L_m; L = 0, \ldots, L_m) = \sum_{i=1}^{N} Log(B(x_i)). \tag{6}$$

The objective is to maximize the log-likelihood LL_1 by choosing $\mu_1, \sigma_1, \rho, \omega_L^b; b = 0, L_m, L = 0, \ldots, L_m$.

The left columns of Table 3 report maximum-likelihood parameter estimates for IT and FT games, using only first round data. Separate analyses include and exclude 'spoiler' choices of 100.

In both games the estimates of ω_i show substantial proportions, at least 12%, in all level categories from 0 to 3. (Higher-level types are included in level 3). The median level is two for IT games and one for FT games. The IT games also have a larger fraction of high-level types than FT games.

The estimated means of the normal distribution from which level 0s choose—70 for IT and 116 for FT games—are far from the interval midpoint and from equilibrium, but this is plausible since only level 0 types choose numbers far from equilibrium so the estimate of μ must be far from equilibrium to explain those observed choices. The estimated correlation ρ is 1.00 in both cases. This implies that higher-level subjects are choosing much more variable numbers than would be predicted if they were simply best-responding to an average of independent choices by others. Their behavior is consistent with players choosing against a 'representative-agent player' or composite, neglecting variation in the sample mean.[20]

[20]The correlation ρ might be a game-theoretic incarnation of the 'representativeness' heuristic in statistical judgment (see, for example, Daniel Kahneman and Amos Tversky, 1982). People using representativeness judge likelihoods of samples by how well they represent a population or process. Representativeness inadvertently neglects other statistical properties like variation—in this case, higher-level players neglect the fact that independent draws tighten the variance of the average they best-respond to. A related phenomenon has been observed in experiments on 'weak link' coordination games, in which a player's payoff depends upon his action and the minimum action chosen by others. The distributions of first-round choices in these games is strikingly similar across groups of different sizes (even though the chance of getting a low minimum rises sharply as the group grows), as if players represent all other players as a single composite.

The rightmost columns of Table 3 show parameter estimates using Nagel's (1995) data with $p = 1/2$ and $p = 2/3$. Estimates of normalized level percentages (from her Figure 2) are shown in parentheses. Our estimates using her $p = 2/3$ data suggest more level 0s and fewer level 3s than in our $p < 1$ data.

Our method for estimating the proportions of level types and Nagel's method generally give similar results. Her method posits an n-step reasoning process which begins from a reference point, 50 for the numbers reported in Table 3.[21] Our method estimates this starting point, instead, giving a level 0 distribution mean of $\hat{\mu} = 52.2$ for $p = 2/3$ and $\hat{\mu} = 36.0$, for $p = 1/2$. She then tests the theory by counting the frequencies of choices in number intervals corresponding to various reasoning levels. Our structural method, in contrast, uses all the data and assigns each observation to some level of reasoning (based on relative likelihood), giving a more complete picture. For example, her method does not classify the 20% of subjects who choose greater than 50 in the first round in the $p = 2/3$ game (normalizing the percentages she reported spreads the 20% evenly over level categories). Our method mostly classifies these 20% as level 0s and consequently, we estimate a much higher level 0 proportion (28% versus 13%).

6 Further Results: Learning Model Estimation

In this section we estimate two classes of learning models to understand the dynamic process by which choices change over rounds.

6.1 Iterated Best-response Models

The first class of learning models posits various levels of 'iterated best-response'. This model parallels the analysis in the last section of levels of iterated dominance underlying first-round choices, but applies the same basic ideas to learning over rounds.

In the model, level-0 learners simply choose a weighted sum of winning numbers in previous rounds. Level-1 learners assume all others are level-0 learners and best-respond to anticipated choices by level-0 learners. Level 2 learners best-respond to level 1 learners, and so forth.

These levels capture the distinction between adaptive learning (responding only to previous observations) and sophisticated learning (best-responding to anticipated play by others) which is discussed, among others, by Milgrom and Roberts (1991). Level 0 learners are adaptive; higher levels are sophisticated.

[21] Nagel (1994, pp. 23-24) also reports estimates using a reference point of 100. With that reference point, she estimates fewer low-level types and more high-level types, which pushes her estimates even further from ours.

To express the model formally, denote subjects' choices at round $t \in \{1, 2, \ldots, 10\}$ by $x_1(t), x_2(t), \ldots, x_n(t)$. The winning number at round t, $w(t) = p \cdot \bar{x}(t) = p \cdot \frac{x_1(t)+x_2(t)+\ldots+x_n(t)}{n}$. Suppose a subject of level L forms a guess $G_L^j(t)$ about what another subject j will choose. Given $G_L^j(t)$, the subject chooses a best response, to maximize his or her expected payoff. That is, the subject will choose $B_L(t)$ such that

$$B_L(t) = p \cdot \frac{B_L(t) + \sum_{j=2}^{n} G_L^j(t)}{n}. \tag{7}$$

Or

$$B_L(t) = \frac{p}{n-p} \cdot \sum_{j=2}^{n} G_L^j(t). \tag{8}$$

The guess of level L subjects at time t is assumed to be the best response of level $L-1$s (hence ther term 'iterated best-response), i.e.,

$$G_L^j(t) = B_{L-1}(t) \tag{9}$$

The level 0 subject j is assumed[22] to choose randomly from a normal density with mean μ equal to a weighted sum of the R previous winning numbers (where R corresponds to level of recall), and variance σ^2. That is,

$$\mu = \sum_{s=1}^{R} \beta_s \cdot w(t-s). \tag{10}$$

The parameters β_s capture the influence of past winning numbers on the current choice. In addition, the correlation between subject choices in any level is ρ (for the same reasons given in the previous section).

Assume that a fraction α_L of the players are level L best-responders, and compose a mixture of the underlying distributions to form an overall distribution of probability of number choices, B. Then $B(x)$ is given by:

$$B(x) = \sum_{L=0}^{L_m} \alpha_L \cdot B_L(x). \tag{11}$$

where L_m is the highest level allowed (restricted to three in our estimates).

[22]In an earlier draft we assumed level 0 types chose a weighted sum of previous winning numbers and their own previous choices. However, the coefficients on previous choices were rarely significant so at the suggestion of referees those terms were dropped from specification (6.4). Our earlier analysis also allowed only level 1 types and referees wisely coaxed us to do this more general analysis.

When $L_m = 1$, level 0 learners choose from a normal distribution with mean given by (6.4) above and level 1 learners choose from a normal distribution with mean given by $\frac{p\cdot(n-1)}{n-p} \cdot [\sum_{s=1}^{R} \beta_s \cdot w(t-s)]$. This implies that when p is further from 1, and n is small, choices will be closer to equilibrium (because the fraction $\frac{p\cdot(n-1)}{n-p}$ is less (greater) than one for IT (FT) games). Our experiment was designed to test these predictions.

In addition, variants of three familiar special cases are nested in the general model. These are Cournot dynamics (Augustin Cournot, 1838), a variant of fictitious play (George Brown, 1958), and a hybrid case in which previous observations are given geometrically declining weight.

1. **Modified Fictitious Play**: Fictitious play learning rules assume that the probability of another player's future choice is best predicted by the empirical frequency of that choice in previous plays. A plausible variant of this applied to the p-beauty contest game is that subjects are all level 1 learners who expect all others to choose an equally-weighted average of the numbers they chose in the past. Modified fictitious play can thus be tested by restricting all types to be level 1 ($\alpha_1 = 1$), and $\beta_1 = \ldots = \beta_R = \beta$. (A further restriction is $\beta = 1/R$ but, as we shall see, that is strongly rejected.)

2. **Geometric Weighted Average**: Fictitious play weights all previous observations equally. A more plausible model assigns geometrically decreasing weights to older observations, then averages them. The declining weight model will fit learning better if subjects realize that choices come from a nonstationary distribution (or others are learning too), and therefore give more recent observations more weight. In this model, all types are level 1 ($\alpha_1 = 1$) and $\beta_s = \beta^s$.

3. **Cournot Dynamics**: Cournot best-response dynamics assumes that players guess others will repeat their most previous choices—i.e., all types are level 1 ($\alpha_1 = 1$) and $\beta_1 = 1, \beta_2, \ldots, \beta_R = 0$.

The log likelihood of observing a sample of N subjects over a total of 10 periods is given by:

$$LL_2 = \sum_{i=1}^{N} \cdot \sum_{s=1}^{10} Log(B((x_i(s))) \tag{12}$$

Some subtle issues arise in implementing the estimation. The standard method in estimating models with R lags is to exclude the first R rounds of data. Since we estimate models with R up to 3, this means discarding 30% of the data. Inspired by a referee's call to fix initial conditions of the model somehow, we do so by estimating a hypothetical 'initial winning number' W_0. Level 0 learners are assumed to act as if they had observed the winning number W_0, before making their first round choice. Continuing along the

same lines, to estimate the model with $R = 3$ we estimate hypothetical winning numbers W_{-1} and W_{-2}. This method fixes the initial conditions, uses all the data, and uses the same data for different R values so they can be fairly compared.

Table 4 reports parameter estimates for recall lengths R=1,2 and 3, with spoilers included and excluded. Generally R=3 fits best, as indicated by χ^2 statistics in the bottom rows. But the estimates of weights on the winning numbers two and three periods back, $\hat{\beta}_2$ and $\hat{\beta}_3$, are low, so a model which assumes only one period of recall would be an adequate approximation for some purposes.

The estimates of learner-level proportions $\hat{\alpha}_i$ show that in IT games, nearly half the players are level 1 learners and the rest are roughly evenly divided between levels 0 and 2. In FT games types are evenly divided between levels 0 and 1. All models restricted to only one type are strongly rejected in favor of the many-type model.[23]

Estimates of the hypothetical initial winning numbers $\hat{W}_0$ are quite plausible, around 50 for IT games and 150 for FT games. The estimated standard deviations of level 0 choices, $\hat{\sigma}$, are reasonable too.

The estimates of $\hat{\beta}_1$ are above one for IT games and below one for FT games[24], which may seem odd since it implies that level 0 learners are picking numbers that are further from equilibrium than the previous winning number. But keep in mind that level 1 learners choose a fraction $\frac{p\cdot(n-1)}{n-p}$, of their guess about the average level 0 choice. When this fraction is multiplied by the typical $\hat{\beta}_1$, the product is usually around one or lower (for IT games), which captures the idea that level 1 players choose numbers below previous winning numbers. If the estimates of $\hat{\beta}_1$ were much lower, that would force the level 1 choices to be 'too small' to fit the data well.

We also estimated the learning-model parameters separately for each of the eight treatment combinations for both IT and FT games.[25]

Parameter estimates are significantly different for different values of p and n, but not in an interesting way. Parameter estimates are different for inexperienced and experienced subjects in IT games in a way that is interesting: Experience shifts the level with the highest estimated proportion up about one level (e.g., $\hat{\alpha}_1 = .572$ and $\hat{\alpha}_2 = .159$ for inexperienced, versus $\hat{\alpha}_1 = .290$ and $\hat{\alpha}_2 = .417$ for experienced). At the same time, the estimated initial winning numbers are generally <u>further</u> from equilibrium in the experienced groups ($\hat{W}_0 = 48.23$, versus 40.74 for inexperienced). This interesting combination

[23]The χ^2 statistics are 1916 and 200 ($\alpha_0 = 1$ only), 2464 and 720 ($\alpha_1 = 1$), 3283 and 1320 ($\alpha_2 = 1$), and 2430 and 1844 ($\alpha_3 = 1$) for IT and FT games respectively.

[24]The discrepancy in estimates $\hat{\beta}_i$ shows that parameters for IT and FT learning differ (the differing α_i estimates show this as well), so it can be rejected as a general theory of learning with invariant parameters.

[25]The eight analyses are not reported for the sake of brevity. In addition, the estimation is not very reliable for $FT(p, 3)$ games because the sample size is smaller, and convergence to the equilibrium of 200 is often so rapid and uniform across subjects that there is little variation to fit (as is evident in Figures 3e-f).

suggests that previous experience in FT games creates a 'negative transfer' to initial choices in IT games (shown by higher values of $\hat{W}_0$) but also creates positive structural transfer, evidenced by more high-level learners. These findings also sharpen the result 4 mentioned in section 4 above, showing how the more powerful learning model analysis yields new insight.

The Cournot, fictitious play, and geometrically declining weight restrictions on β_i are all strongly rejected.[26] This is not surprising given the estimates in Table 4, because $\hat{\beta}_1$ is usually far from one (rejecting Cournot) and $\hat{\beta}_2$ and $\hat{\beta}_3$ are close to zero (rejecting the other two theories).

6.2 Learning Direction Theory

Unlike the models above, which assume some players formulate a belief and choose a best response given their beliefs, 'belief-free' models posit direct relations between choices and observable variables. For example, Nagel (1995) describes a 'learning direction' theory and shows how it accounts for many features of her data. We describe this theory and test it in some detail because it represents a potentially important part of Nagel's contribution, and its performance may give clues about the likely performance of other belief-free 'reinforcement' theories (which conceptualize learning as a direct behavioristic relationship between previous outcomes and future choices).

The idea in learning direction theory is that players move in the direction of choices which are optimal ex post. Nagel adds the plausible complication that players change by not simply raising and lowering their number choices, but by adjusting the ratio of the number they chose to the previous winning number (the 'adjustment factor'), lowering the ratio if their number was too high and raising the ratio if their number was too low. This boils down to the following prediction:

$$\begin{aligned} \hat{x}_i(t) &< x_i(t-1) \cdot \frac{w(t-1)}{w(t-2)} \quad \text{if } x_i(t-1) > w(t-1) \\ \hat{x}_i(t) &> x_i(t-1) \cdot \frac{w(t-1)}{w(t-2)} \quad \text{if } x_i(t-1) < w(t-1) \end{aligned} \tag{13}$$

where, as before, $\hat{x}_i(t)$ denotes predicted choice in period t. Nagel tests learning direction theory by frequency counts of how often the bounds in (6.7) are violated. We do the same, but also reformulate the theory to make it more comparable to the best-response models. In the form stated by Nagel, learning direction theory makes a set- theoretic prediction. Figure 4 illustrates. Values of $w(t-1) - x_i(t-1)$ are shown on the x-axis

[26]For all data including spoilers, using $R = 3$, the χ^2 statistics for Cournot, fictitious play, and geometric weights are 2678, 2572, and 2512 for IT games, and 1246, 644, and 654 for FT games.

and $\frac{x_i(t)w(t-2)}{x_i(t-1)w(t-1)}$ on the y-axis. Now divide the space into quadrants around the point x=0, y=1 (i.e., where $w(t-1) = x_i(t-1)$ and $x_i(t) = x_i(t-1)w(t-1)/w(t-2)$).

Learning direction theory predicts that if actual values were plotted on Figure 4, most of the data would lie in the two shaded quadrants to the right and above of (0,1) (numbers are too low in t-1, adjust upward in t), and to the left and below of (0,1) (numbers too high, adjust downward).

Purely to compare this kind of prediction parametrically with the other learning models, we express the set-theoretic prediction stated by Nagel as a very general response function that is continuous and locates data only in the upper right and lower left quadrants. Consider the following response function:

$$\hat{x}_i(t) = [\gamma_1 - \frac{\gamma_2}{1 + e^{-\gamma_3 \cdot (x_i(t-1) - w(t-1))}}] \cdot x_i(t-1) \cdot \frac{w(t-1)}{w(t-2)} \tag{14}$$

As $\gamma_3 \to \infty$,

$$\begin{aligned} \hat{x}_i(t) &= (\gamma_1 - \gamma_2) \cdot x_i(t-1) \cdot \frac{w(t-1)}{w(t-2)} \quad \text{if } x_i(t-1) > w(t-1) \\ \hat{x}_i(t) &= \gamma_1 \cdot x_i(t-1) \cdot \frac{w(t-1)}{w(t-2)} \quad \text{if } x_i(t-1) < w(t-1) \end{aligned} \tag{15}$$

This response function is a variant of a stochastic choice model in which the likelihood of choosing A over B, say, varies continuously and nonlinearly with the difference between their utilities. In (6.8), the ratio $\frac{\hat{x}_i(t)}{x_i(t-1)} \cdot \frac{w(t-2)}{w(t-1)}$ varies nonlinearly with the difference $x_i(t-1) - w(t-1)$. The free parameters γ_i specify the location and curvature of the function. Keep in mind that this response function is a very specific form of learning direction theory (cf. Stahl, in press), different than the form studied by Nagel and designed only to make it more comparable to the parameterized iterated best-response models. Rejecting it does not imply rejection of the set-theoretic version of the theory.

The response-function form of learning direction theory can be tested by assuming normal errors in responses and choosing parameters to maximize the log likelihood of the data $x_i(t)$. We then test various restrictions on the parameters γ_i which restrict the shape of the response function. For example, if γ_3 is finite the response function is continuous– Figure 4 shows an example– and has γ_1 and $\gamma_1 - \gamma_2$ as asymptotes in the upper and lower quadrants. Furthermore, for the response function to pass through the origin (allowing only outcomes in the upper right and lower left quadrants) requires that at $x_i(t-1) - w(t-1) = 0$, the coefficient in (6.8) is 1, which implies $\gamma_1 - \frac{\gamma_2}{2} = 1$.

A simpler nonparametric test of learning direction theory is whether players usually change their adjustment factors, raising and lowering number choices, in a way that depends on previous choices and winning numbers. Table 5a shows transition frequencies from three possible conditions on period $t-1$ choices– choices are below, equal to, or above the winning number– to three conditions for period t choices. (A fourth condition for period $t-1$ choices, called 'never fail' means that learning direction theory predicts a bound which can never be violated. [27])

Learning direction theory predicts most transitions will lie on the diagonal categories (printed in bold). Most transitions do lie in these categories but a substantial fraction of observations lie in the wrong cells (37% and 46% in infinite- and finite-threshold games).

A more complicated test uses maximum-likelihood estimation to estimate the parameters γ_i of the response function (6.8), and tests whether the restrictions that learning direction theory imposes are satisfied. Table 5b shows estimates and log likelihoods (using data from the second round on).

Stricter forms of learning direction theory can be characterized as making any of three predictions: (a) $\gamma_1 > 1, \gamma_1 - \gamma_2 < 1$ (i.e., the response function asymptotes in the shaded quadrants in Figure 4); (b) $\gamma_1 - \frac{\gamma_2}{2} = 1$ (the response function passes through the point (0,1) in Figure 4); and (c) $\gamma_3 = \infty$ (the response function is a step function, discontinuous at the point (0,1)).

Table 5b gives maximum likelihood estimates for γ_1, γ_2, γ_3, and σ, estimated separately for FT and IT games. In both games, restriction (a) is not rejected but restrictions (b) and (c) are. These rejections indicate that the response-function form organizes the data adequately if the response function simply asymptotes into the predicted quadrants, but not if it satisfies the stronger properties of sharp inflection and passing through the point (0,1). (The latter rejection means the response function 'crosses into' an unpredicted quadrant, perhaps the strongest evidence against the theory from this analysis.) The substantial number of wrong transitions shown in Table 5a also indicate that there is room for improvement in the set-theoretic form of the theory as well.[28]

Finally, while the response-function form of the learning direction approach and the iterated best-response models are not nested, when they are compared using an information criterion which penalizes theories for using extra degrees of freedom, the iterated best-response theory does substantially better.[29]

[27]For example, for infinite-threshold games with $p < 1$, suppose $w(t-1)$ and $w(t-2)$ are 30 and 20 and $x(t-1)$ is 70. Learning direction theory predicts that since $x(t-1)$ was higher than the previous winning number, $x(t)$ must then be less than $x(t-1)w(t-1)/w(t-2)$, or 105, which will always be satisfied. The category 'never fail' includes all these cases.

[28]Nagel and Duffy suggested learning direction theory may work well in the early rounds of an experiment (like the four rounds in their papers) and not as well in the later rounds, which explains the mixed performance in our ten-round analysis.

[29]The Akaike criterion (AIC), for example, subtracts $k(ln(n)+1)$ from 2 times the log likelihood, where n is sample size and k is the number of free parameters (see, e.g., Hamparsum Bozdogan, 1987).

7 Discussion

7.1 Rationality-thresholds and iterated dominance

Standard game theoretic models are indifferent to task complexity since players are assumed to be fully rational. Our results show that choices reveal a limited number of steps of iterated dominance (which could be taken as a sharp measure of the degree of bounded (mutual) rationality). Since iterated dominance is limited, the minimal number of iterated dominance steps needed to solve a dominance-solvable game, which we call the rationality-threshold, is correlated with equilibrium convergence and behavior.

Many previous studies also showed modest levels of iterated dominance. Randolph Beard and Richard Beil (1994) studied mutual rationality and found that about 50% of the subjects violated it, in a game presented in a tree form and played sequentially. Using a similar game, Andrew Schotter, Weigelt, and Charles Wilson (1994) found that half their subjects violated mutual rationality, and about 20% violated weak dominance in a matrix-form game. (These percentages were much lower in an extensive-form game.) John van Huyck, Daniel Wildenthal, and Raymond Battalio (1994) study a five-strategy variant of a prisoners' dilemma in which choices corresponds to levels of iterated dominance, and found about three levels of iterated dominance. Camerer, Barry Blecherman, and David Goldstein (1995) studied the 'electronic mail game' introduced by Ariel Rubinstein (1988) and observed about two levels of iterated dominance. In McKelvey and Palfrey's study (1992) of centipede games, most subjects reveal 2 or 3 levels of iterated dominance.

In a study which essentially created the approach we used, Stahl and Wilson (1995) define 'level-0' players as those who choose strategies randomly and equally-often, level-1 players as those who optimize against lower level (level-0) players, and so forth. In three-strategy matrix games, they estimate that most players are level-1 or level-2 but the games are not designed to discriminate levels higher than that.

As we mentioned in the introduction, <u>all</u> these games are nonconstant-sum. In experimental applications (where dollar payoffs cannot be presumed to map exactly into utilities), a test of dominance is a joint test of utility-maximization and the self-interest assumption that utility depends only on one's own dollar payoffs. For example, in McKelvey & Palfrey's centipede games, 15-20% of the players who arrive at the final node violate dominance by 'passing'. Passing means taking 20% of a \$32 pie (earning \$6.40 while the other player gets \$25.60) instead of 80% of a \$16 pie (earning \$12.80 while the other gets \$3.20).

Players who obey dominance (maximizing utility) but care substantially about others'

We reestimated the iterated best-response model starting with second round data (excluding W_{-2}) to make it comparable to learning direction theory. The results yield AICs of -20724 (IT) and -12208 (FT) for learning direction and -18259 and -11588 for iterated-best response. (More negative numbers are bad.)

payoffs[30] will violate the conjunction of dominance and self-interest by passing at the final node. In addition, a third of the subjects who reach the penultimate, second-to-last node, either violate dominance, or believe others will, by passing at that node. Indeed, the centipede game is the rule, not the exception: In all the games mentioned previously in this section, subjects who violate the self-interest assumption may appear to violate dominance.

This point is also evident in the large literature on bargaining and public goods games. In these games, a large portion of players violate the conjunction of dominance and self-interest by rejecting low offers, giving money to others, contributing to a public good, or cooperating in a prisoner's dilemma (e.g., Camerer and Richard Thaler, 1995; David Sally, 1995; John Ledyard, 1995). But these results overstate the degree of dominance violation if players are simply being fair-minded or altruistic.

As discussed before, the p-beauty contest game allows us to more neatly separate self-interest and iterated dominance.[31] Since our game is constant-sum, actions which violate iterated dominance are not easily reconciled with models of fairness or altruism.[32] Thus, p-beauty contests are a superior game for isolating degrees of iterated dominance from the complications of self-interest violations.

In the original work on these games, Nagel (1995) used large groups various values of p (above and below one), playing for four rounds. For $p = 2/3$, for example, she reports an average initial choice around 36, which correspond to about 2 levels of iterated dominance (see also Table 3), or one step of reasoning from a reference point of 50. She also observes gradual convergence toward equilibrium.

John Duffy & Nagel (1995) studied 12-16 person groups with $p = 1/2$, in which the target number was p times the mean, median, or maximum number chosen in a group. (All these games have the same equilibrium of zero.) They find no substantial difference between mean and median games, and higher choices in the maximum game.

Our results extend these earlier findings in several ways. We draw a novel distinc-

[30] For example, assume a linear additive 'social utility function' in which player i's utility for the allocation (x_i, x_j) is $x_i + \alpha x_j$. Then passing is utility-maximizing iff $\alpha > .29$.

[31] Another way to separate the two is to collect information-processing measures other than choices. For example, Eric Johnson et al (1995) show using measures of attention that players in a three-stage sequential bargaining game often do not look past the first stage, violating the computational underpinning of iterated dominance.

[32] The claim that p-beauty contests are fairness-free depends on the model of fairness being considered. In Matthew Rabin's (1993) approach, for example, 'niceness' and meanness' of actions alter players' social utilities away from pure self-interest, then a standard equilibrium concept is applied to the altered utilities. In p-beauty contests with sufficiently small stakes, this idea could create fairness equilibria in which everybody picks a number, say 50, and shares the prize equally. The equilibrium is supported by the fact that a deviation would 'meanly' enhance the deviator's payoff at the expense of others, leading to punishments which end at the Nash equilibrium in which the prize is still shared equally so deviation doesn't pay. Empirically, these fairness equilibria probably become fragile as the number of subjects increases (which may explain the anomalous group size effect we observe) and the small effect of fairness may also be squashed by the modest stakes we use.

tion between games with finite and infinite rationality-thresholds and show that finite-threshold games converge more quickly and reliably.[33] Our use of a Stahl-Wilson type structural model of levels of reasoning, estimated from first round choices, gives a sharper characterization of levels of iterated dominance. Using different group sizes reveals a puzzling effect– smaller groups learn slower. Playing different games sequentially shows evidence of positive learning transfer.

7.2 Learning

By using ten rounds instead of four (and collecting eight times as much data), we got a fuller picture of learning. The extra data enable reliable estimates of learning models. The tests of various learning models suggest some stylized facts:

- Players appear to be influenced by up to three previous winning numbers.
- The data are consistent with presence of adaptive (level 0) learners who simply respond to experience, and sophisticated (level 1 and higher) learners who best-respond to lower-level learners. Thus, any learning model which hopes to describe well should include both types.
- Familiar learning models which are special cases of our approach, including Cournot best-response dynamics (which looks back only one period) and fictitious play, are clearly rejected.
- Restrictions corresponding to a form of the 'learning direction' model Nagel proposed can be rejected. An information-criterion-based comparison of learning direction shows it fits worse than iterated best-response.

7.2.1 Payoff reinforcement

Best-response models of the sort estimated above do not directly alter the probability of choosing a strategy according to the strategy's previous payoff.

An entirely different class of 'payoff reinforcement' models ignore beliefs, and assume that the propensity to play strategies depends on a specific initial propensity and on how strategies have been 'reinforced' (based on observed success or failure) in the past. Reinforcement models of this kind were widely used to study animal behavior, and human learning in the heyday of 'behaviorism', until the 1960s or so when they were largely abandoned for models with more cognitive detail.

[33]Nagel conducted three sessions with $p = 4/3$ and a number interval [0,100]. First-period data look roughly like a reflection of $p = 2/3$ data around the midpoint of 50, although many choose 33 and 100. However, with choices in [0,100] and $p > 1$, no numbers are ruled out by any level of iterated dominance (in contrast to our design). Also, 0 and 100 are both equilibria in her design, though 0 is not trembling-hand perfect, but 200 is the unique equilibrium in ours.

Reinforcement models have been applied to many games recently.[34] In future research we plan to apply these models to p-beauty contest data. However, there is a reason to think reinforcement models will not fit p-beauty contest data well. Winning numbers converge toward a boundary-prediction equilibrium over time. Reinforcement models assume that a strategy which wins one time gets reinforced, and is likely to be chosen again. Then convergence will occur sluggishly because 'old' winning numbers will continue to get picked (unless previous reinforcements are very quickly forgotten). The basic problem is that it is hard to build endogenously into reinforcement the fact that (as post-experiment comments indicate) players realize winning numbers are converging.[35]

Stahl (1994) proposes a different model in which decision rules are reinforced rather than specific strategies (number choices). Rule learning has some precedence in related literature (e.g., Richard Dawkins on culturally-transmitted 'memes) and squarely addresses the shortcomings of strategy reinforcement just mentioned. In Stahl's rule-learning approach, subjects choose one of K decision rules, where decision rule k chooses a number equal to $p^k \cdot w(t-1)$ after observing a previous winning number $w(t-1)$. Rules are reinforced by their expected payoff. Reinforcing rules, rather than specific numbers, does reinforce reasoning the 'best' number of steps ahead. Fitting this model to Nagel's data, and using several other free parameters, Stahl finds that initial propensities toward k between 0-2 are about equal, and in about half the sessions propensities move toward $k = 2$ over four periods. In addition, he rejects a variety of alternative models (some nested, some not) including behavior reinforcement, forecasting of changes in the ratios of winning numbers, and direction learning.

Another concern is that strategy-reinforcement theories must posit an initial propensity to choose various strategies. (Initial propensities are usually taken to be random or fit to first round play, though in Stahl's rule-learning approach they are estimated, and show evidence of heterogeneity.) Evidence of transfer of experience across similar games can be interpreted as evidence about how initial propensities change. Analysis from the iterated best-response model showed that first-period choices are somewhat affected by previous experience. This implies that initial propensities reflect experience in a related game. In addition, experienced players converge faster than inexperienced ones (and are estimated to use higher levels of iterated best-response). But in reinforcement terms, quicker learning means payoffs have more impact, which implies that initial propensities are lowered by experience. Together these results indicate that initial propensities are

[34] See Roth and Erev (1995) on bargaining games, Erev and Roth (1995) on games with mixed strategies, Amnon Rapoport, Darryl Seale, Erev, and James Sundali (1995) in entry games.

[35] A referee opined that this criticism of reinforcement slays a straw man because, like forecasting price levels during inflationary times, subjects will know enough to focus on the 'deflation rate' of winning numbers rather than their levels. We take the point, but it would be even better for reinforcement models to generate such a realization naturally rather than building it into the space of what is being reinforced. In addition, in p-beauty contests with interior equilibria in $(0, 100)$ it is not clear that it is more sensible to forecast inflation or deflation rather than levels (forecasting changes may create overshooting), so which strategies are reinforced becomes an empirical question. A closely related problem is that it is hard for reinforced learners to learn to play repeated game strategies like alternation or trigger strategies, unless many-trial patterns are reinforced.

altered by experience, but at lower overall levels (or equivalently, experienced subjects are more responsive to early-period reinforcements). Since this is certainly a complicated way to think about the influence of transferred learning, a more cognitively detailed approach is surely worth exploring further.

References

[1] Binmore, K., Gale, J., and Samuelson, L. 'Learning to Be Imperfect: The Ultimatum Game,' Game and Economic Behavior, 8, 56-90, 1995.

[2] Randolph Beard and Richard Beil 'Do People Rely on the Self-interested Maximization of Others - An Experimental Test,' Management Science, pp. 252-262, 1994.

[3] Bozdogan, Hamparsum. 'Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extensions,' Psychometrika, 52, 345-370, 1987.

[4] Brown, G. 'Iterative Solution of Games by Fictitious Play,' In Activity Analysis of Production and Allocation, New York: John Wiley & Sons, 1951.

[5] Cachon, Gerard and Camerer, Colin F. 'Loss-avoidance and forward induction in experimental coordination games,' Quarterly Journal of Economics, February 1996, 111, 165-194.

[6] Camerer, Colin, Blecherman, Barry, and Goldstein, David. 'Iterated Dominance and Learning in 'Electronic Mail Game' Experiments.' California Institute of Technology working paper, 1996.

[7] Camerer, Colin and Richard Thaler. 'Anomalies: Dictators, Ultimatums, and Manners,' Journal of Economic Perspectives, 1995.

[8] Cournot, A. *Recherches sur les principes mathematiques de la theorie des richesses.* Translated into English by N. Bacon as Researches in the Mathematical Principles of the Theory of Wealth. London: Haffner, 1960.

[9] Crawford, V. P. 'Adaptive Dynamics in Coordination Games,' Econometrica, 63, 103-143.

[10] Duffy, John and Nagel, Rosemarie. 'On the Robustness of Behavior in Experimental Guessing Games,' University of Pittsburgh Working Paper, 1995.

[11] Erev, Ido and Roth, Alvin. 'On the Need for Low Rationality, Cognitive Game Theory: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria,' University of Pittsburgh Working Paper 1996.

[12] Friedman, M. Essays of Positive Economics. Chicago: Chicago University Press, 1953.

[13] Frydman, R. 'Towards an Understanding of Market Processes: Individual Expectations, Learning, and Convergence to Rational Expectations Equilibrium,' American Economic Review, 72, 652-668.

[14] Garvin, Susan and Kagel, John. 'Learning in Common Value Auctions: Some Initial Observations,' Journal of Economic Behavior and Organization, Vol. 25, 1994, 351-372.

[15] Gilboa, Itzhak. 'The Complexity of Computing Best-Response Automata in Repeated Games,' Journal of Economic Theory, 45, 342-352.

[16] Holt, Debra. 'An empirical model of strategic choice with an application to coordination games,' Department of Economics, Queen's University, 1993.

[17] Ho, Teck-Hua and Weigelt, Keith. 'Task Complexity, Equilibrium Selection, and Learning: An Experimental Study,' Management Science, 42, 659-679.

[18] Johnson, Eric, Camerer, Colin, Talia Rymon and Sankar Sen. 'Limited Computation and Fairness in Sequential Bargaining Experiments,' University of Pennsylvania Working paper, 1996.

[19] Jordan, James. 'Bayesian Learning in Normal Form Games,' Games and Economic Behavior, 3, 60-81.

[20] Keynes, John Maynard 'The General Theory of Interest, Employment and Money,' London: Macmillan, 1936.

[21] Kreps, David. Game Theory and Economic Modelling, Oxford: Oxford University Press, 1990.

[22] Ledyard, John. 'Public Goods: A Survey of Experimental Research,' in J. Kagel & A. E. Roth (Eds.), Handbook of Experimental Economics. Princeton: Princeton University Press, 1995, pp. 111-194.

[23] LiCalzi, Marco. 'Fictitious Play by Cases,' Games and Economic Behavior, 11, 64-89, Oct 1995

[24] Lovallo, Dan and Camerer, Colin F. 'Overconfidence and excess business entry: An experimental approach, California Institute of Technology working paper, 1996.

[25] McKelvey, Richard D. and Palfrey, Thomas R. 'An Experimental Study of the Centipede Game,' Econometrica, 1992, 60, 803-836.

[26] McKelvey, Richard D. and Palfrey, Thomas R. 'An Experimental Study of the Centipede Game,' Econometrica, 1992, 60, 803-836.

[27] McKelvey, Richard D. and Palfrey, Thomas R. 'Quantal response equilibria for normal form games,' Games and Economic Behavior, 1995, 10, 6-38.

[28] McKelvey, Richard D. and Palfrey, Thomas R. 'Quantal response equilibria for extensive form games,' Caltech working paper 1996.

[29] Myerson, Roger. Game Theory: Analysis of Conflict, Boston: Harvard University Press, 1992.

[30] Milgrom, Paul and Roberts, John. 'Adaptive and Sophisticated Learning in Normal Form Games,' Games and Economic Behavior, 3(1) February 1991.

[31] Milgrom, Paul and Roberts, John. 'Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities,' Econometrica, 58(6): Nov 1990, 1255-1277.

[32] Moulin, Herve. Game Theory for the Social Sciences, New York University Press, New York, 1986.

[33] Nagel, Rosemarie. 'Reasoning and learning in guessing games and ultimatum games with incomplete information: An experimental investigation, Doctoral dissertation, Universitat Bonn, 1994.

[34] Nagel, R. 'Experimental Results on Interactive Competitive Guessing,' American Economic Review, 1995, 85, 1313-1326.

[35] Porter, David and Smith Vernon L. 'Futures Contracting and Dividend Uncertainty in Experimental Asset Markets,' Journal of Business, 68, 509-542, October 1995.

[36] Rabin, Matthew. 'Modelling fairness in game theory and economics,' American Economic Review, 83, December 1993, 1281-1302.

[37] Rapoport, A., Seale, D., Erev, R. and Sundali, J. 'Coordination Success in Market Entry Games: Tests of Equilibrium and Adaptive Learning Models,' Management Science, in press.

[38] Roth, Alvin and Erev, Ido. 'Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term,' Games and Economic Behavior, 8, 164-212 (1995).

[39] Rubinstein, Ariel. 'The Electronic Mail Game: Strategic Behavior under 'Almost Common Knowledge" American Economic Review, 79 (1989), 395-391.

[40] Sally, David. 'Conversation and Cooperation in Social Dilemmas: Experimental Evidence from 1958 to 1992,' Rationality and Society, 1994.

[41] Schotter, Andrew; Weigelt, Keith; and Charles Wilson. 'A laboratory investigation of multi-person rationality and presentation effects', Games and Economic Behavior, 1994, 6, 445-468.

[42] Selten R. and Stoecker R. 'End Behavior in Sequences of Finitely Repeated Prisoner's Dilemma Supergames - A Learning Theory Approach,' Journal of Economic Behavior and Organization, 7, 47-70.

[43] Singley, Mark K. And John R. Anderson. The Transfer of Cognitive Skill. Cambridge: Harvard University Press, 1989.

[44] Stahl, Dale. 'The evolution of $smart_n$ players,' Games and Economic Behavior, 5, 1993, 604-617.

[45] Stahl, Dale. 'Rule Learning in a Guessing Game,' Games and Economic Behavior, forthcoming.

[46] Stahl, Dale and Wilson, Paul. 'Experimental evidence of Players' Models of Other Players,' Journal of Economic Behavior and Organization, 1994.

[47] Stahl, Dale and Wilson, Paul. 'On Players Models of Other Players: Theory and Experimental Evidence,' Games and Economic Behavior, 10, 218-254, July 1995.

[48] van Huyck, J. Wildenthal, J. and Battalio, R. 'Tacit Cooperation, Strategic Uncertainty, and Coordination Failure: Evidence from Repeated Dominance Solvable Games,' 1994.

APPENDIX

INSTRUCTIONS

This is an experiment in decision-making. Several research foundations have provided funds for these experiments. If you follow the instructions, and make good decisions, you may earn a sizable amount of money. The amount of money you earn depends on your choices and the choices of other subjects in the experiment.

YOUR DECISION PROBLEM

The experiment will last for 10 rounds. During each round, you will be presented with an identical choice problem. There are n subjects in a group and each of you will simultaneously choose a number between 100 to 200 inclusive. Since each of you has worked in privacy, after making your choice, an administrator will come around and record your number. Your payoff will be determined as follows. Let x_i be person i's number. First the average number $\bar{x}$ will be computed as follows:

$$\bar{x} = \frac{x_1 + x_2 +, \ldots, x_{n-1} + x_n}{7}.$$

The person whose number is the nearest to $1.1 \cdot \bar{x}$ receives \$3.5 and the rest of you receive nothing. If there are identical numbers which are equally close to $1.1 \cdot \bar{x}$, then the \$3.5 prize will be equally divided among the persons who chose these numbers. The administrator will publicly announce $1.1\bar{x}$ and come around and privately inform you whether you win or not.

PAYOFFS

Your dollar earnings for the experiment are the sum of your dollar earnings in every round of the game.

Round	1	2	3	4	5	6	7	8	9	10
Your Number										
$\bar{x}$										
$1.1 \cdot \bar{x}$										
Payoff										

Transfer of Learning	Group Size: 3	Group Size: 7
Finite --> Infinite	FT(1.3,3) --> IT(0.7,3) (7 groups) FT(1.1,3) --> IT(0.9,3) (7 groups)	FT(1.3,7) --> IT(0.7,7) (7 groups) FT(1.1,7) --> IT(0.9,7) (7 groups)
Infinite --> Finite	IT(0.7,3) --> FT(1.3,3) (7 groups) IT(0.9,3) --> FT(1.1,3) (6 groups)	IT(0.7,7) --> FT(1.3,7) (7 groups) IT(0.9,7) --> FT(1.1,7) (7 groups)

Table 1: The Experimental Design

Table 2: Frequencies of Levels of Iterated Dominance Over Round in FT and IT games With Varying P Values

Games/Round	1	2	3	4	5	6	7	8	9	10	Total
FT(1.3,n)											
R(0)	35	9	12	15	9	5	11	3	4	7	110
R(1)	68	34	12	6	7	5	7	3	4	0	146
R(2)	34	67	44	26	27	22	11	11	3	4	249
Equilibrium Play	3	30	72	93	97	108	111	123	129	129	895
FT(1.1,n)											
R(0)	10	2	4	5	2	8	2	5	8	5	51
R(1)	8	1	2	0	2	2	2	0	2	1	20
R(2)	10	4	2	2	1	1	1	0	1	0	22
R(3)	15	12	4	3	3	2	1	3	1	1	45
R(4)	54	42	18	6	1	0	4	2	1	3	131
R(5)	23	42	41	18	9	4	5	2	6	5	155
R(6)-R(10)	14	28	42	61	64	54	40	36	39	33	411
Equilibrium Play	3	6	24	42	55	66	82	89	79	89	535
IT(0.7,n)											
R(0)	37	5	6	5	7	6	6	10	7	8	97
R(1)	39	26	13	8	2	3	5	2	2	1	101
R(2)	15	38	18	12	11	3	3	5	7	5	117
R(3)	12	23	30	23	17	20	14	7	8	13	167
R(4)	13	26	28	22	22	22	26	21	25	16	221
R(5)	4	9	24	19	16	19	19	17	15	17	159
R(6)-R(10)	15	10	21	50	61	47	45	57	51	40	397
> R(11)	2	0	0	1	2	10	11	7	9	16	58
Equilibrium Play	3	3	0	0	2	10	11	14	16	24	83
IT(0.9,n)											
R(0)	9	3	2	1	1	3	1	1	4	3	28
R(1)	7	0	1	1	0	1	0	0	1	0	11
R(2)	17	6	3	1	2	1	1	1	0	1	33
R(3)	9	8	8	4	1	0	0	0	1	1	32
R(4)	14	19	8	10	6	4	3	2	2	1	69
R(5)	5	9	15	6	6	6	5	1	2	1	56
R(6)-R(10)	46	71	69	73	61	39	44	36	30	30	499
> R(11)	27	20	29	40	59	77	77	85	86	89	589
Equilibrium Play	3	1	2	1	1	6	6	11	11	11	53

Table 3: MLEs and Log-likelihoods for Levels of Iterated Dominance (first round data only)

Game parameter estimates	Our Data (Groups of 3 or 7)		Nagel's Data (Groups of 16-18)	
	Spoilers Included	Spoilers Excluded	p = 1/2	p = 2/3
IT(p,n)				
ω_0	18.05	15.93	15.83 (23.94)	28.36 (13.11)
ω_1	20.22	20.74	37.50 (29.58)	34.33 (44.26)
ω_2	13.00	13.33	46.67 (40.84)	37.31 (39.34)
ω_3	48.74	50.00	0.00 (5.63)	0.00 (3.28)
μ	70.00	68.87	36.97 (50.00)	52.23 (50.00)
σ	27.71	27.04	22.12	14.69
ρ	1.00	1.00	0.25	1.00
–LL	1143.15	1128.98	156.9	244.0
FT(p,n)				
ω_0	24.55	21.72		
ω_1	30.32	31.46		
ω_2	12.27	12.73		
ω_3	32.85	34.08		
μ	116.36	117.72		
σ	13.81	13.52		
ρ	1.00	1.00		
–LL	1139.34	1117.02		

Table 4: MLEs and Log-likelihoods for the Iterated Best Response Learning Models

Game parameter estimates	Spoilers Included (N=2770)			Spoilers Excluded (N=2711 (infinite-threshold); 2666(finite-threshold))		
	Recall Period			Recall Period		
	R=1	R=2	R=3	R=1	R=2	R=3
IT(p,n)						
α_0	0.1905	0.2112	0.2102	0.3204	0.3201	0.3173
α_1	0.4452	0.4398	0.4394	0.5100	0.4977	0.4907
α_2	0.2977	0.2892	0.2872	0.1696	0.1521	0.1604
α_3	0.0666	0.0597	0.0633	0.0000	0.0302	0.0316
β_{-1}	1.25	1.40	1.36	1.17	1.27	1.26
β_{-2}	-	-0.14	0.02	-	-0.09	0.06
β_{-3}	-	-	-0.12	-	-	-0.13
W_0	53.03	41.58	42.12	44.04	40.62	38.83
W_{-1}	-	0.00	35.69	-	0.06	43.89
W_{-2}	-	-	0.00	-	-	0.99
σ	31.84	30.35	31.84	19.66	19.43	19.48
ρ	0.00	0.00	0.00	0.00	0.00	0.00
LL	-10507.9	-10499.8	-10493.3	-10008.9	-9990.1	-9981.1
χ^2		16.2	13.0		37.6	18.0
FT(p,n)						
α_0	0.3324	0.4233	0.4184	0.5649	0.5957	0.5267
α_1	0.2729	0.5767	0.5816	0.4351	0.4043	0.4733
α_2	0.3947	0.0000	0.0000	0.0000	0.0000	0.0000
α_3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_{-1}	0.82	0.63	0.61	0.91	0.71	0.69
β_{-2}	-	0.27	-0.01	-	0.22	0.01
β_{-3}	-	-	-0.32	-	-	0.23
W_0	151.92	137.66	163.09	152.62	165.44	168.78
W_{-1}	-	155.00	135.23	-	107.93	159.67
W_{-2}	-	-	100.00	-	-	100.00
σ	47.01	45.16	31.84	27.58	27.09	26.60
ρ	0.00	0.00	0.00	0.00	0.00	0.00
LL	-7126.2	-7092.5	-7078.9	-6413.62	-6395.4	-6380.8
χ^2		67.4	27.4		35.4	29.2

Table 5a: Frequencies of Number Changes

Game Choice in t-1	Choice in t			Total
	x(t) > x(t-1) w(t-1)/w(t-2)	x(t) = x(t-1) w(t-1)/w(t-2)	x(t) < x(t-1) w(t-1)/w(t-2)	
IT(p, n)				
x(t-1) < w(t-1)	**926**	4	514	1444
x(t-1) = w(t-1)	2	**10**	5	17
x(t-1) > w(t-1)	206	54	**431**	691
Never Fail	64	0	0	64
Total	1198	68	950	2216
FT(p, n)				
x(t-1) < w(t-1)	**27**	0	69	96
x(t-1) = w(t-1)	0	**0**	0	0
x(t-1) > w(t-1)	224	238	**598**	1060
Never Fail	996	0	64	1060
Total	1449	238	731	2216

Note: Bold indicates transitions predicted by learning direction theory

Table 5b: MLEs and Log-likelihoods for the Learning Direction Theory

Game restrictions	γ_1	γ_2	γ_3	W_0	σ	-LL
IT(p, n)						
No restrictions	2.23	2.34	0.03	45.30	19.53	10340.7
$\gamma_1 > 1$; $\gamma_1 - \gamma_2 < 1$	2.23	2.34	0.03	45.30	19.53	10340.7
$\gamma_1 - \gamma_2/2 = 1$	2.13	2.26	0.03	43.22	19.58	10346.3
$\gamma_3 =$ infinity	1.07	1.05	infinity	41.28	32.58	11499.3
All three restrictions	1.74	0.48	infinity	51.23	30.62	12063.7
FT(p, n)						
No restrictions	2.68	3.26	0.003	180.69	54.94	6081.8
$\gamma_1 > 1$; $\gamma_1 - \gamma_2 < 1$	2.68	3.26	0.003	180.69	54.94	6081.8
$\gamma_1 - \gamma_2/2 = 1$	3.34	4.68	0.004	168.76	55.83	6097.7
$\gamma_3 =$ infinity	1.17	0.09	infinity	185.47	56.65	6138.5
All three restrictions	1.15	0.30	infinity	178.63	56.10	6173.9

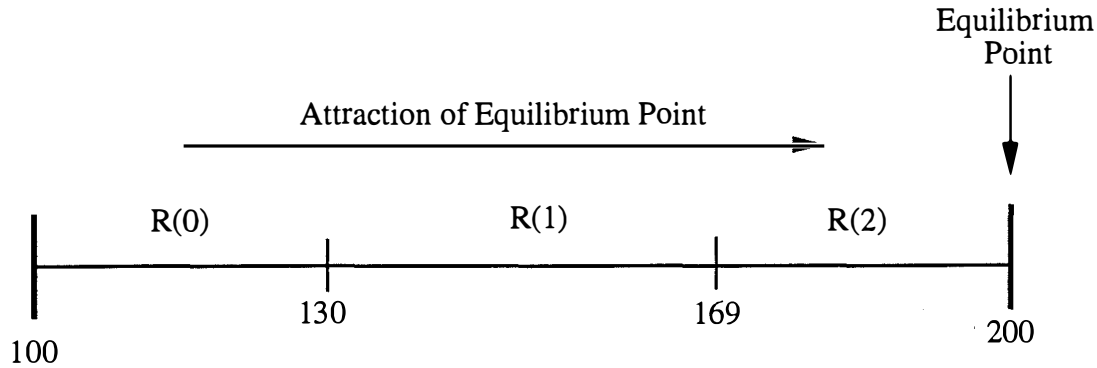


Figure 1a: A Finite-Threshold Game, FT(n) = ([100,200], 1.3, n)

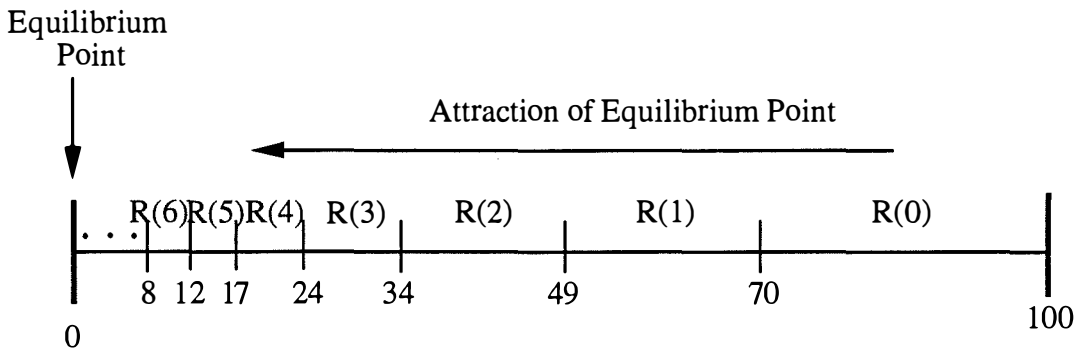


Figure 1b: An Infinite-Level Threshold Game, IT(n) = ([0,100], 0.7, n)

Figure 2a: Inexperienced Subjects' Choices Over Round in IT(0.7, 7)

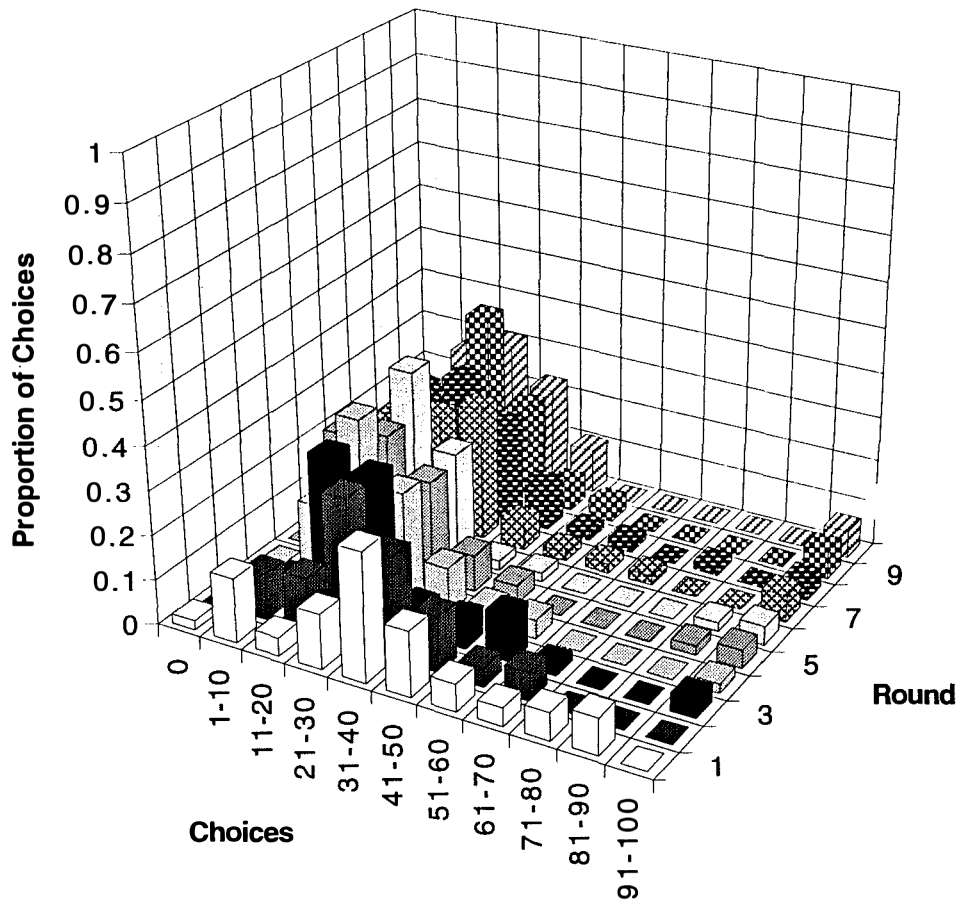


Figure 2b: Experienced Subjects' Choices Over Round in IT(0.7, 7)

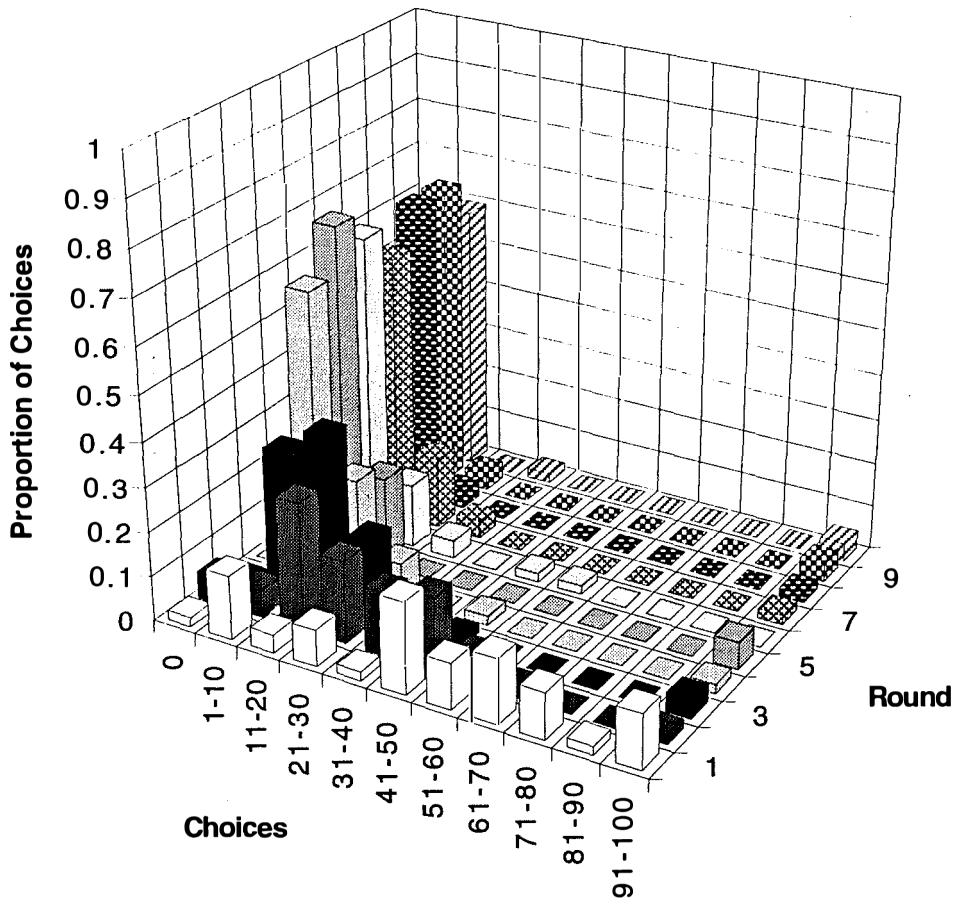


Figure 2c: Inexperienced Subjects' Choices Over Round in IT(0.[illegible], 7)

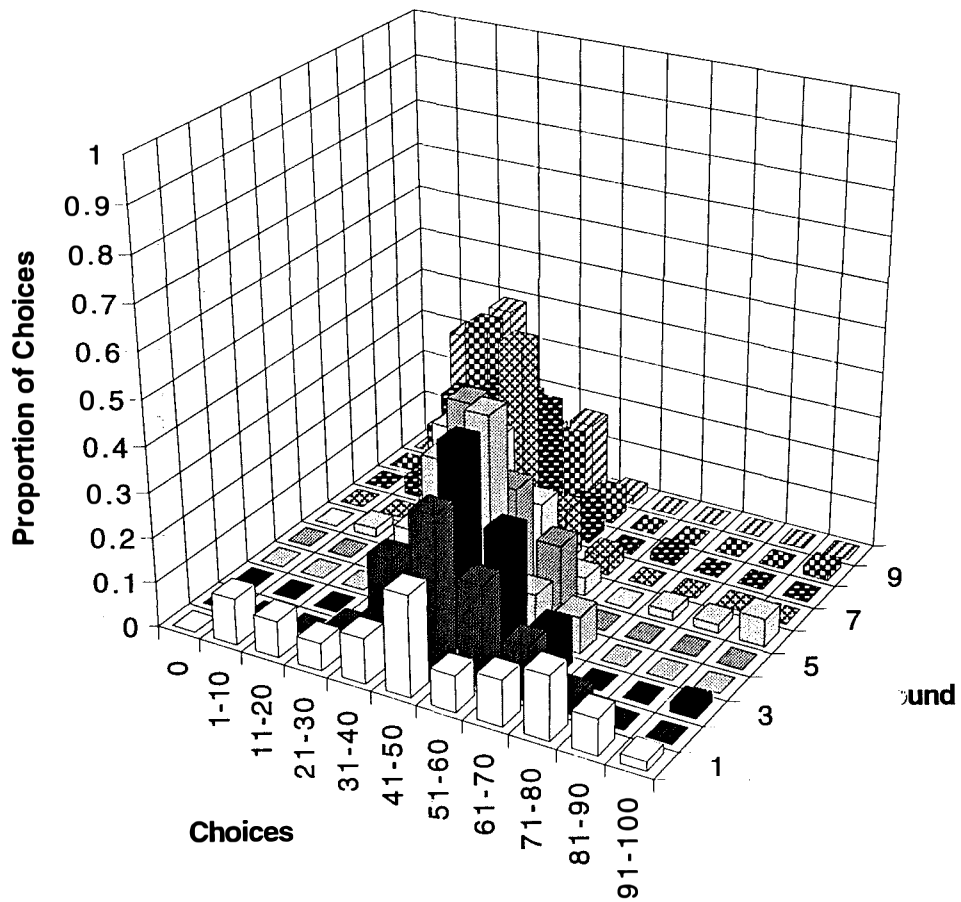


Figure 2d: Experienced Subjects' Choices Over Round in IT(0.9,7)

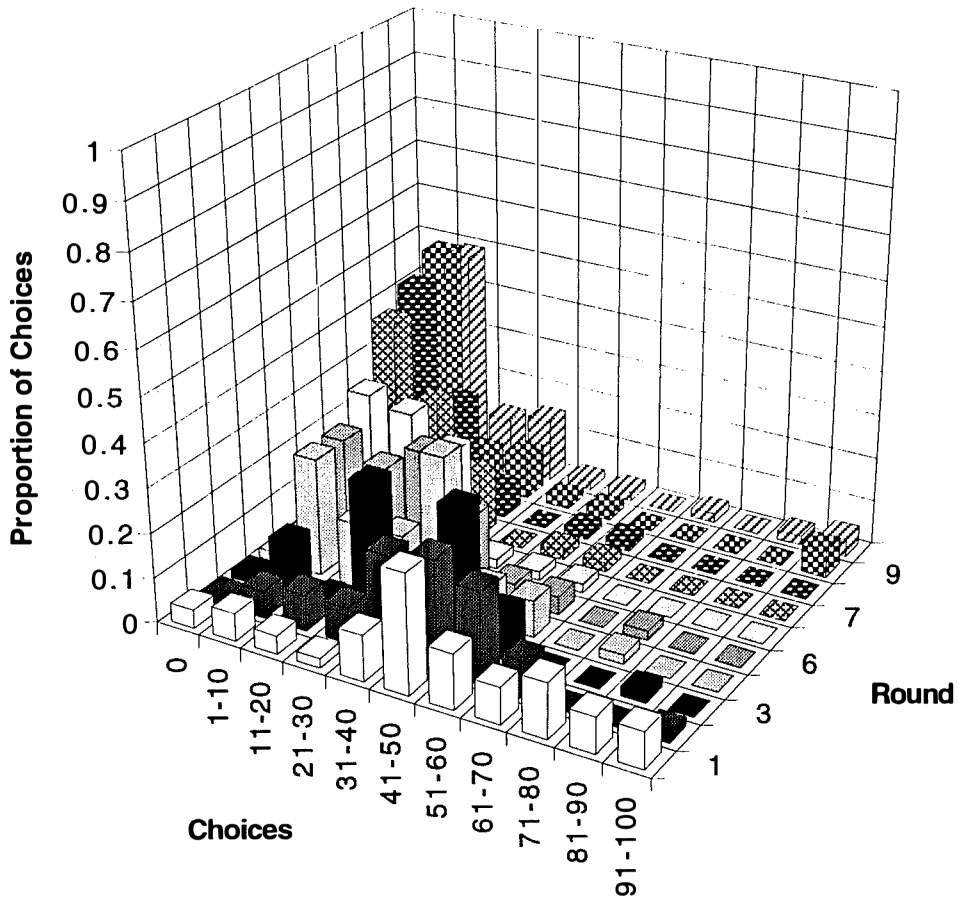


Figure 2e: Subjects' Choices Over Round in IT(0.7, 3)

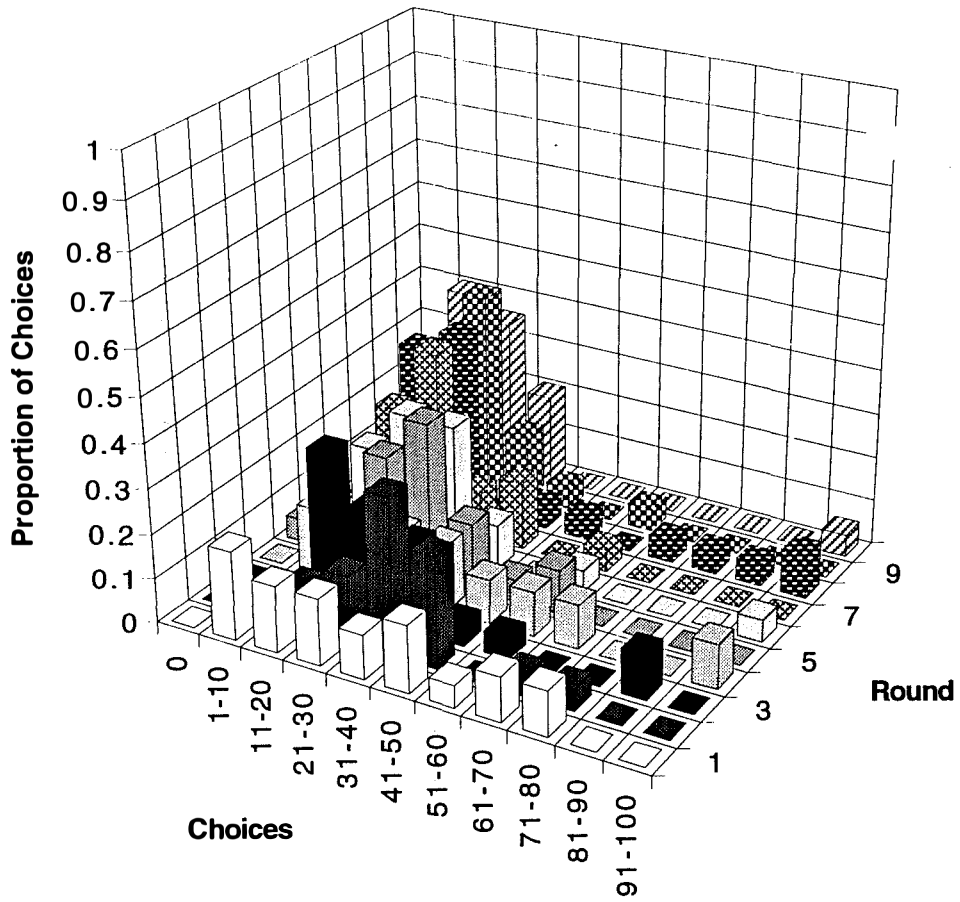


Figure 2f: Experienced Subjects' Choices Over Round in IT(0.7, 3)

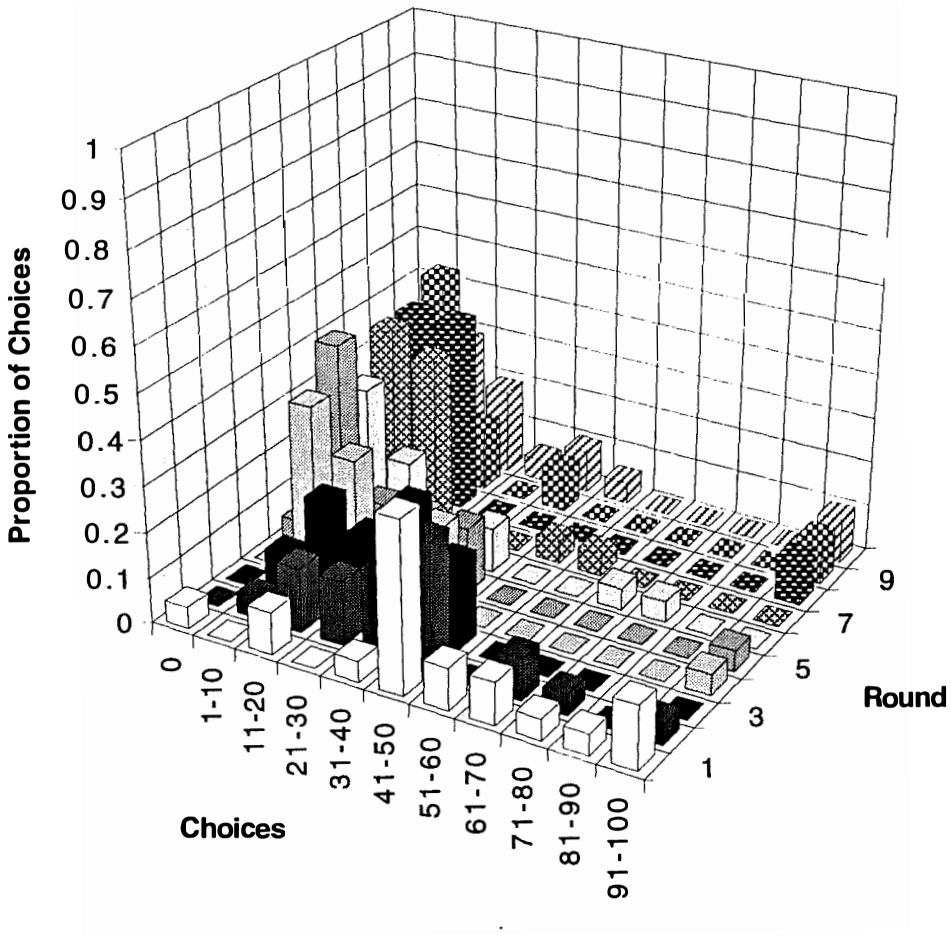


Figure 2g: Inexperienced Subjects' Choices Over Round in IT(0.9, 3)

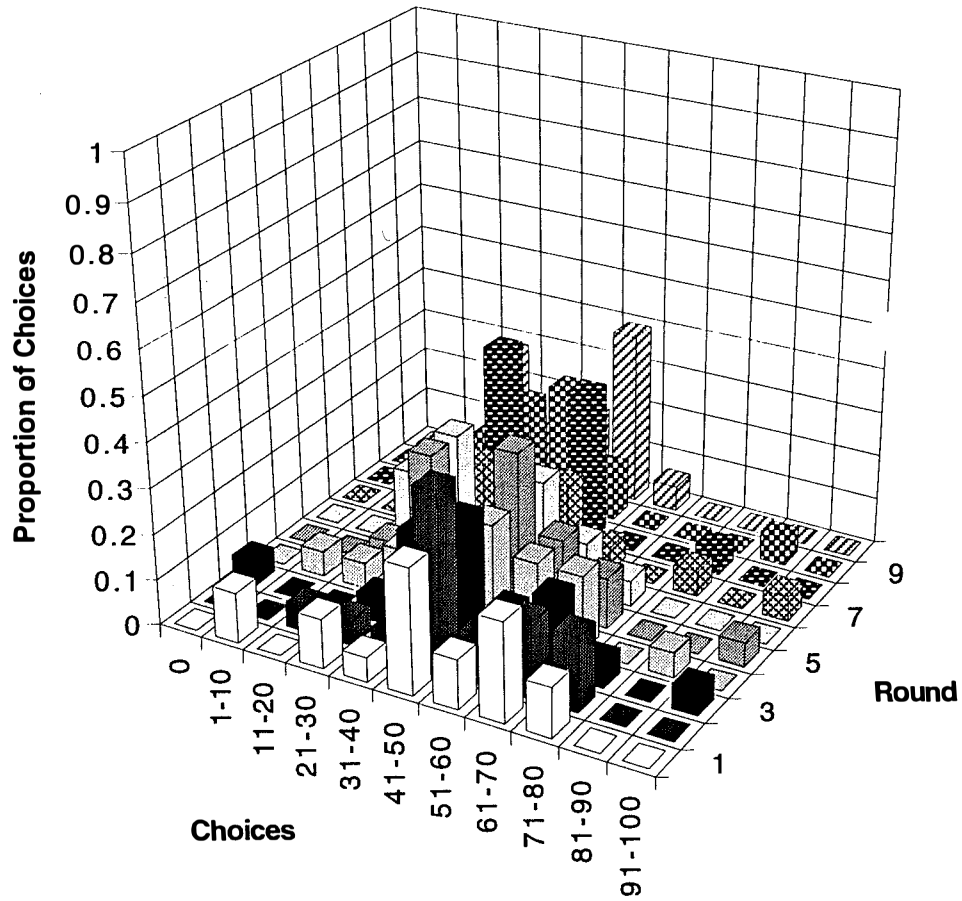


Figure 2h: Experienced Subjects' Choices Over Round in IT(0.9, 3)

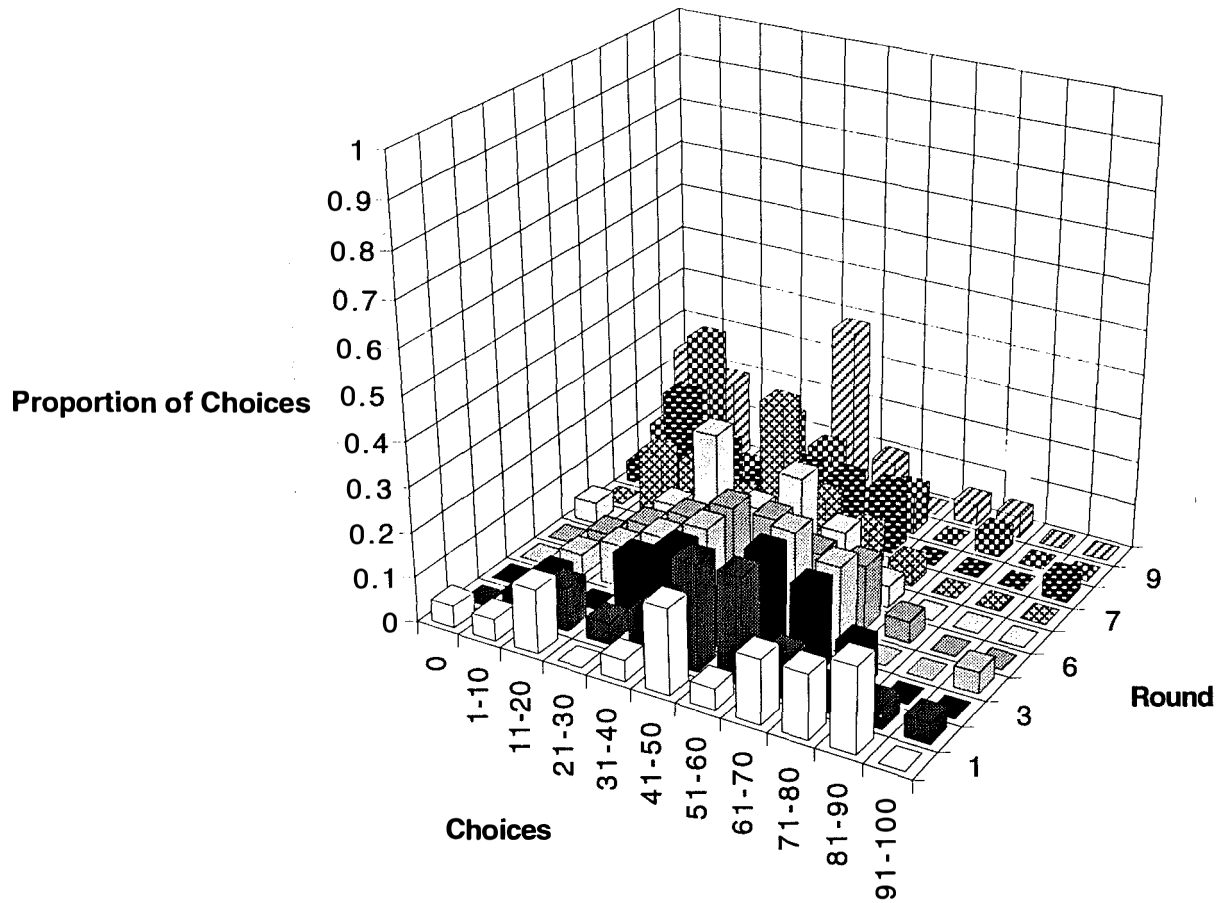


Figure 3a: Inexperienced Subjects' Choices Over Round in FT(1.3, 7)

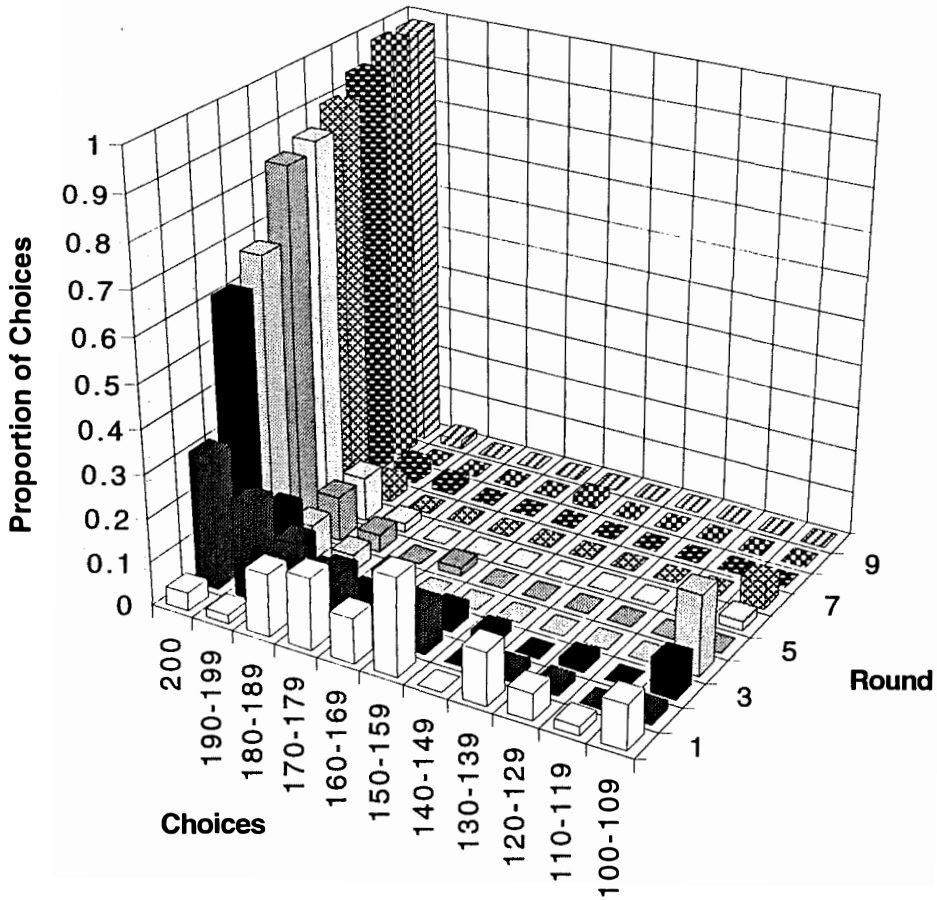


Figure 3b: Experienced Subjects' Choices Over Round in FT(1.3, 7)

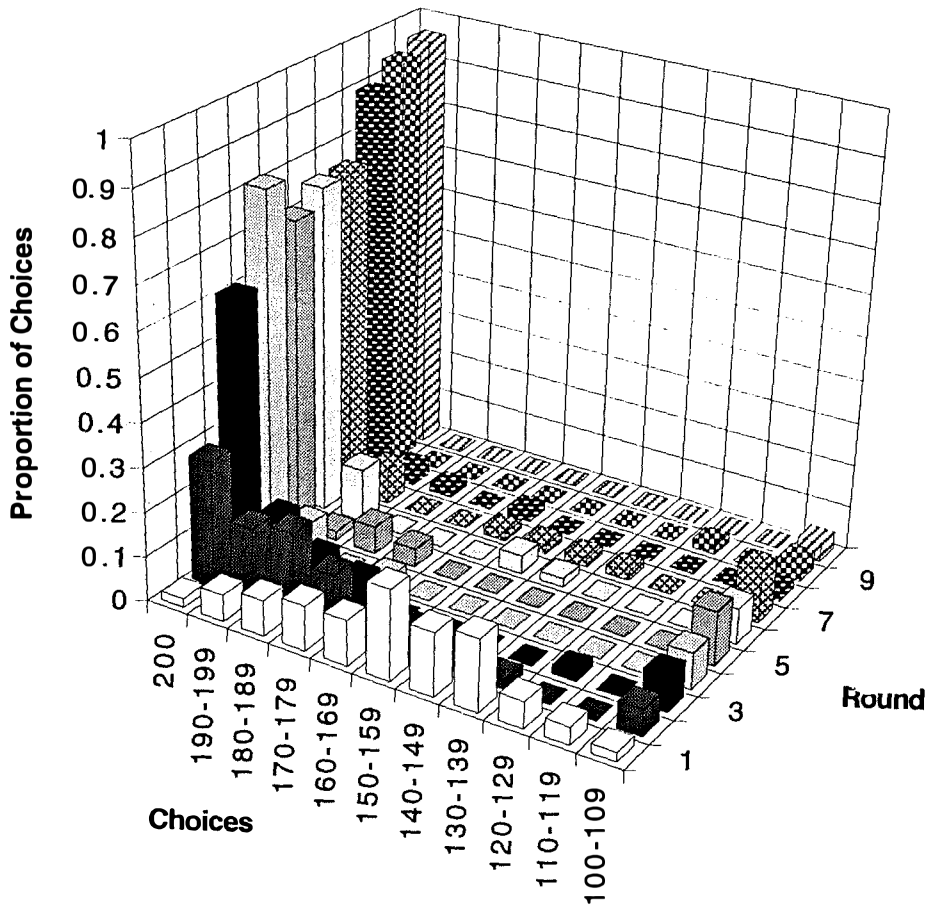


Figure 3c: Inexperienced Subjects' Choices Over Round in FT(1 7)

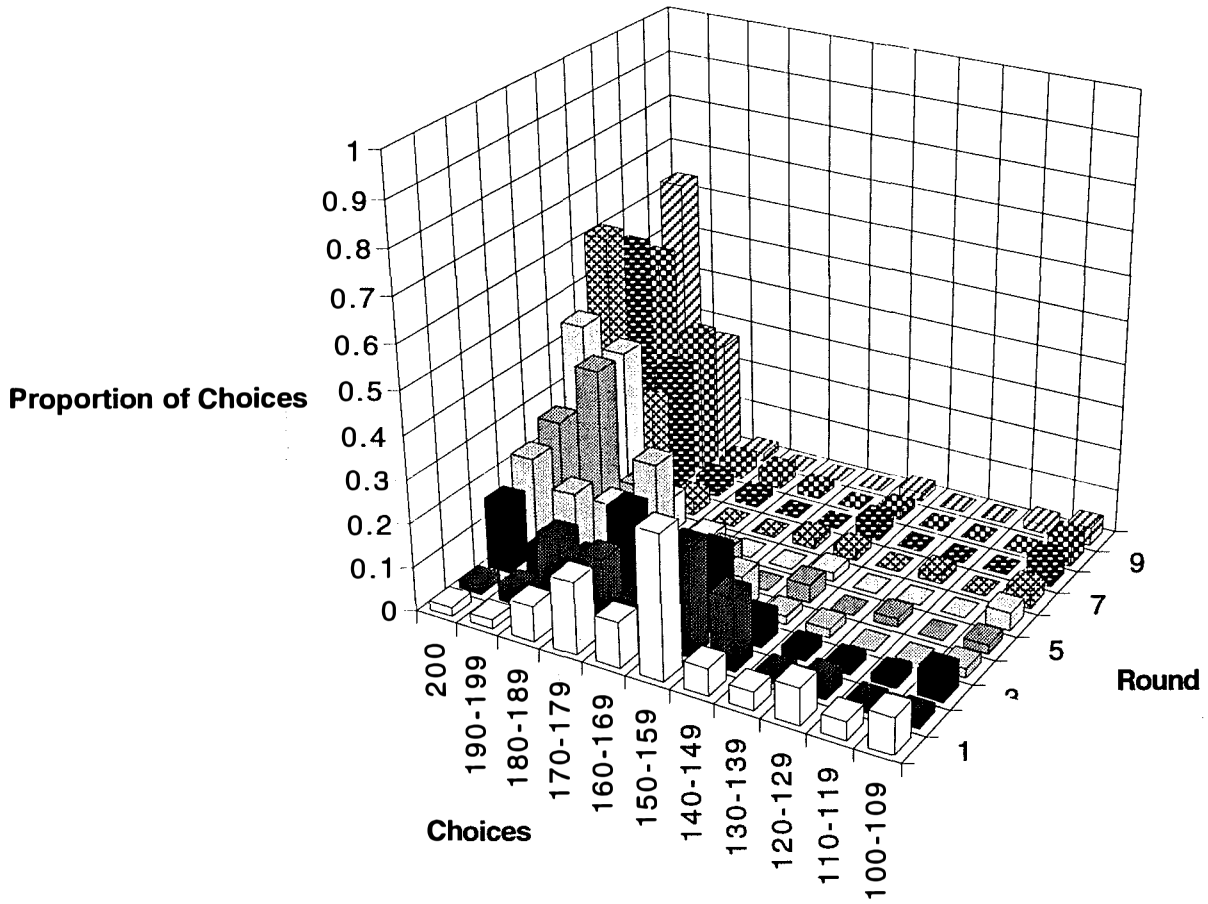


Figure 3d: Experienced Subjects' Choices Over Round in FT(1.1, 7)

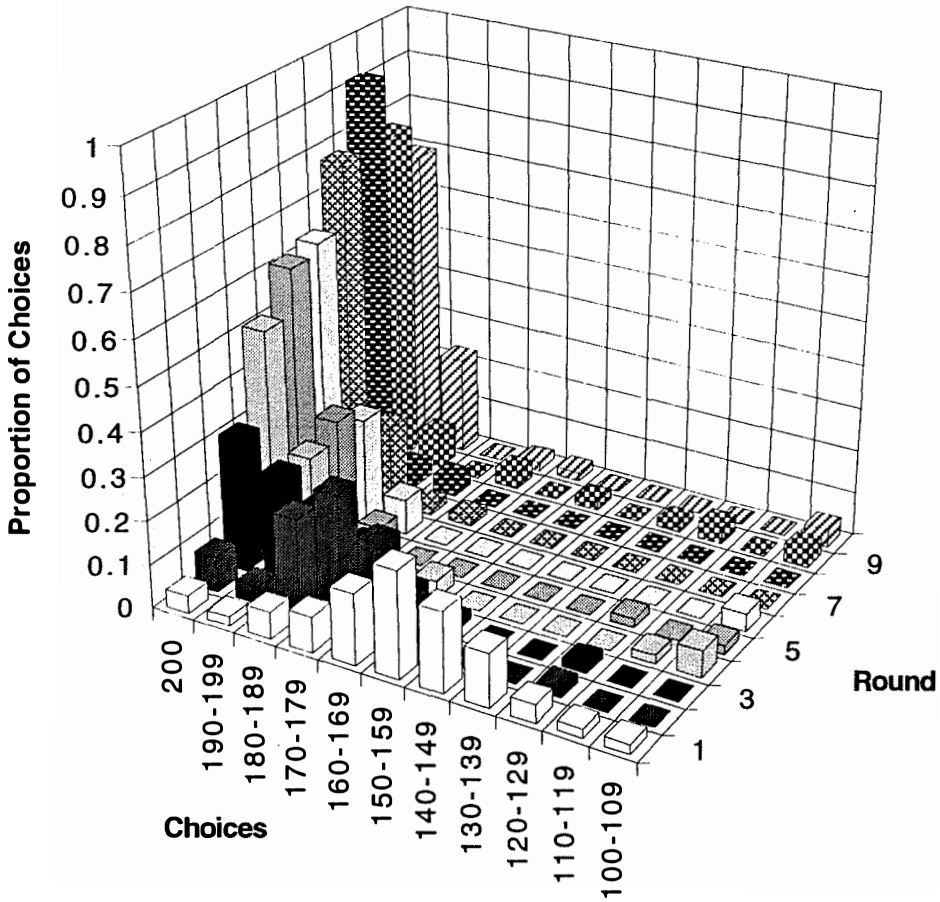


Figure 3e: Inexperienced Subjects' Choices Over Round in FT(1.3, 3)

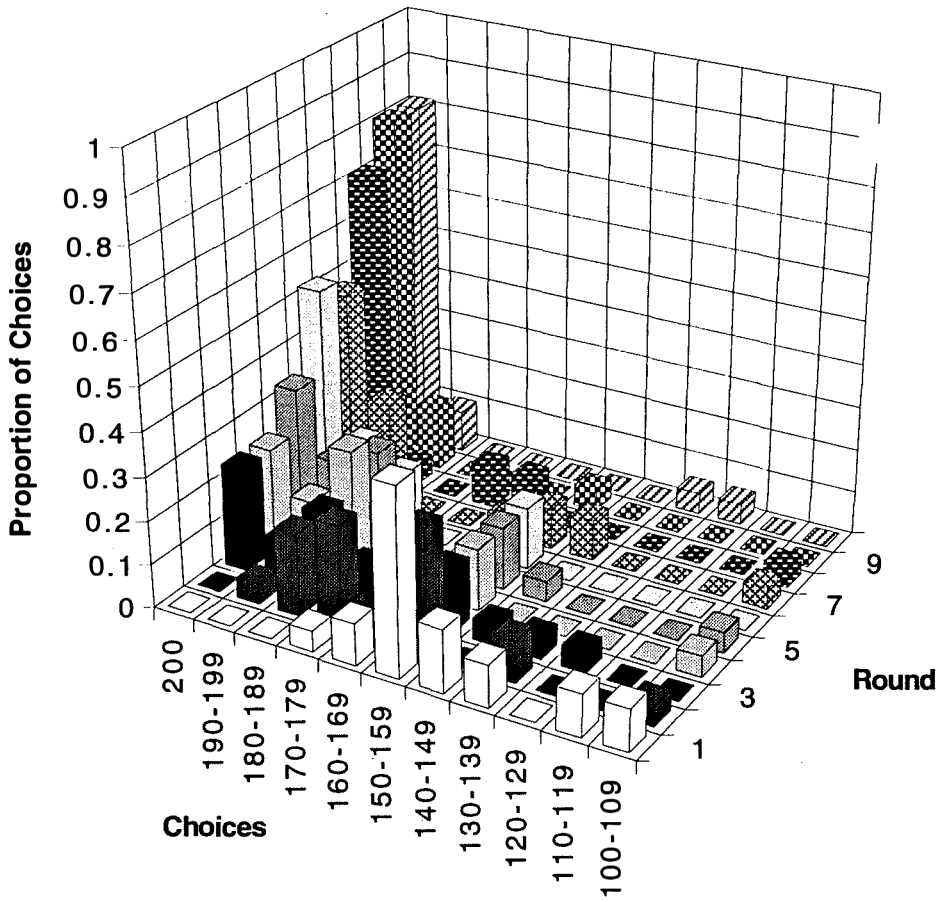


Figure 3f: Experienced Subjects' Choices Over Round in FT(1.3, 3)

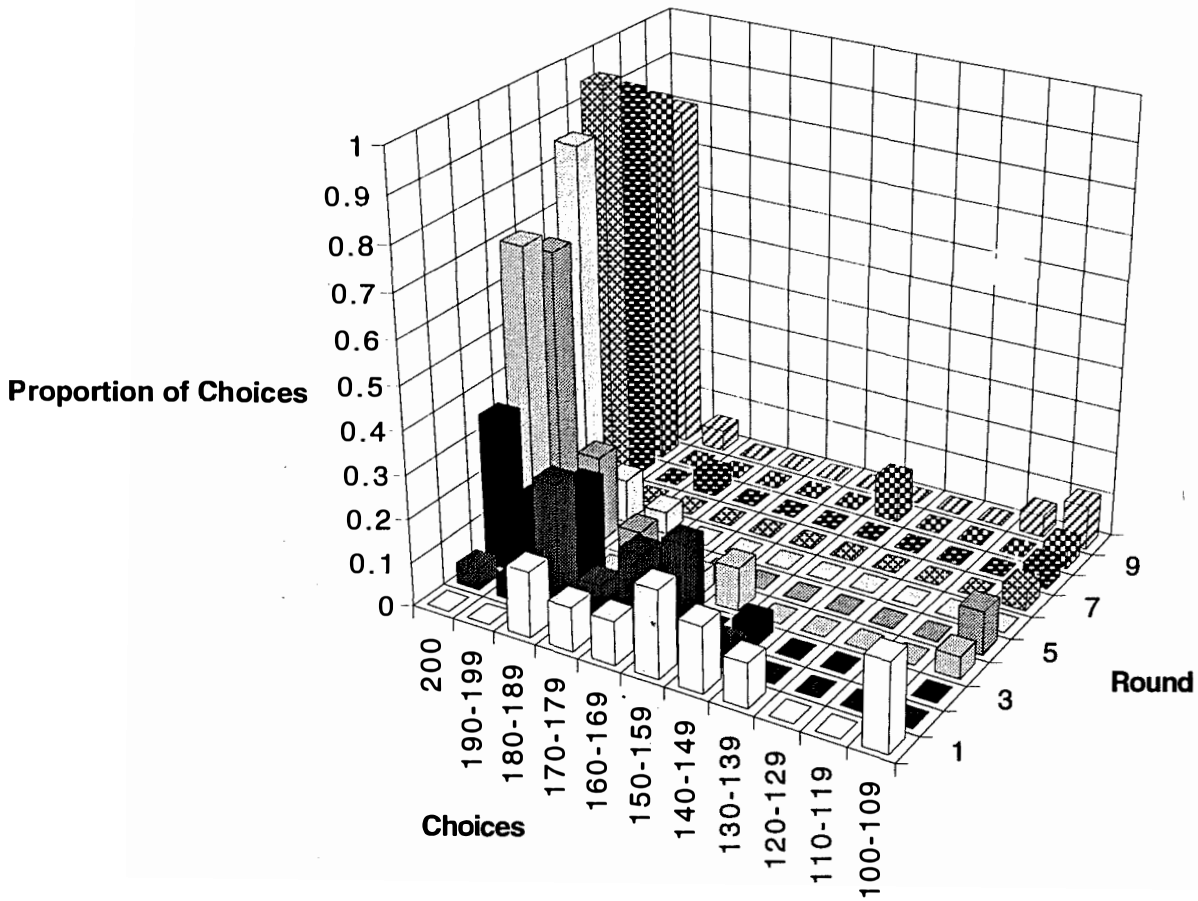


Figure 3g: Inexperienced Subjects' Choices Over Round in FT(1.1, 3)

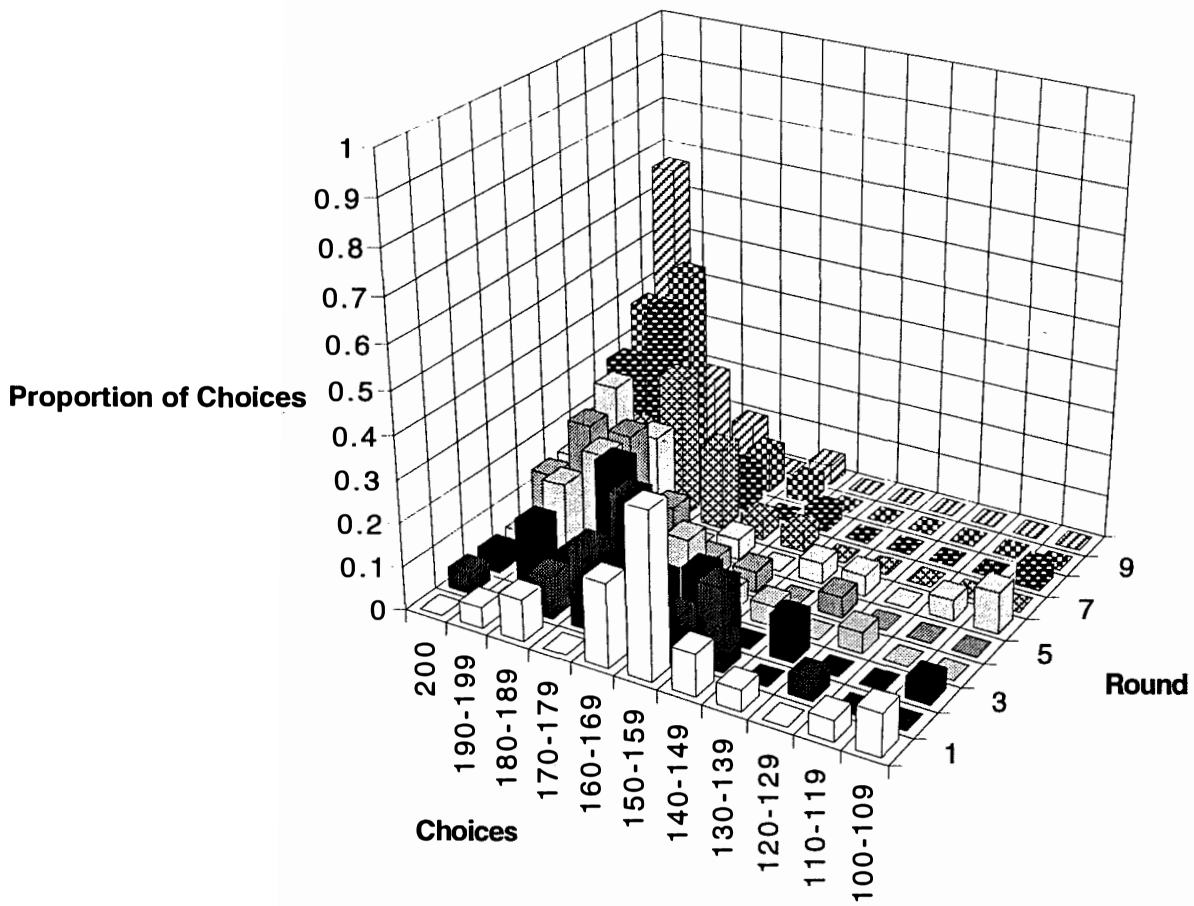
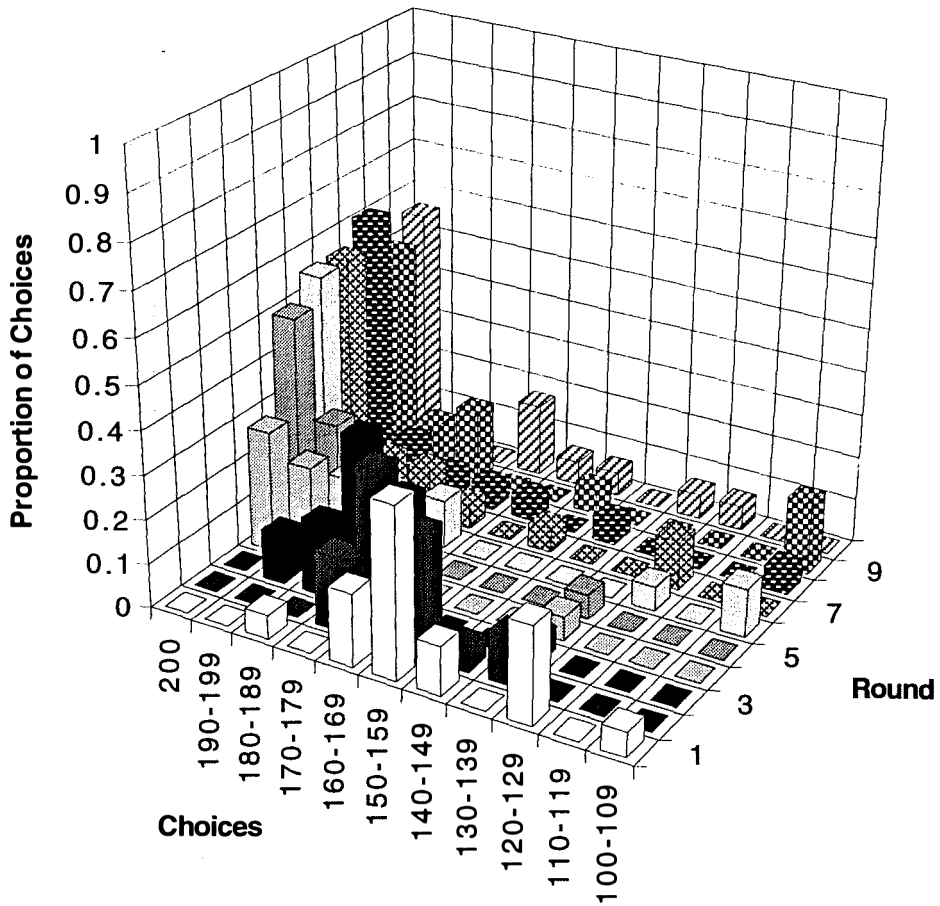


Figure 3h: Experienced Subjects' Choices Over Round in FT(1.1,3)



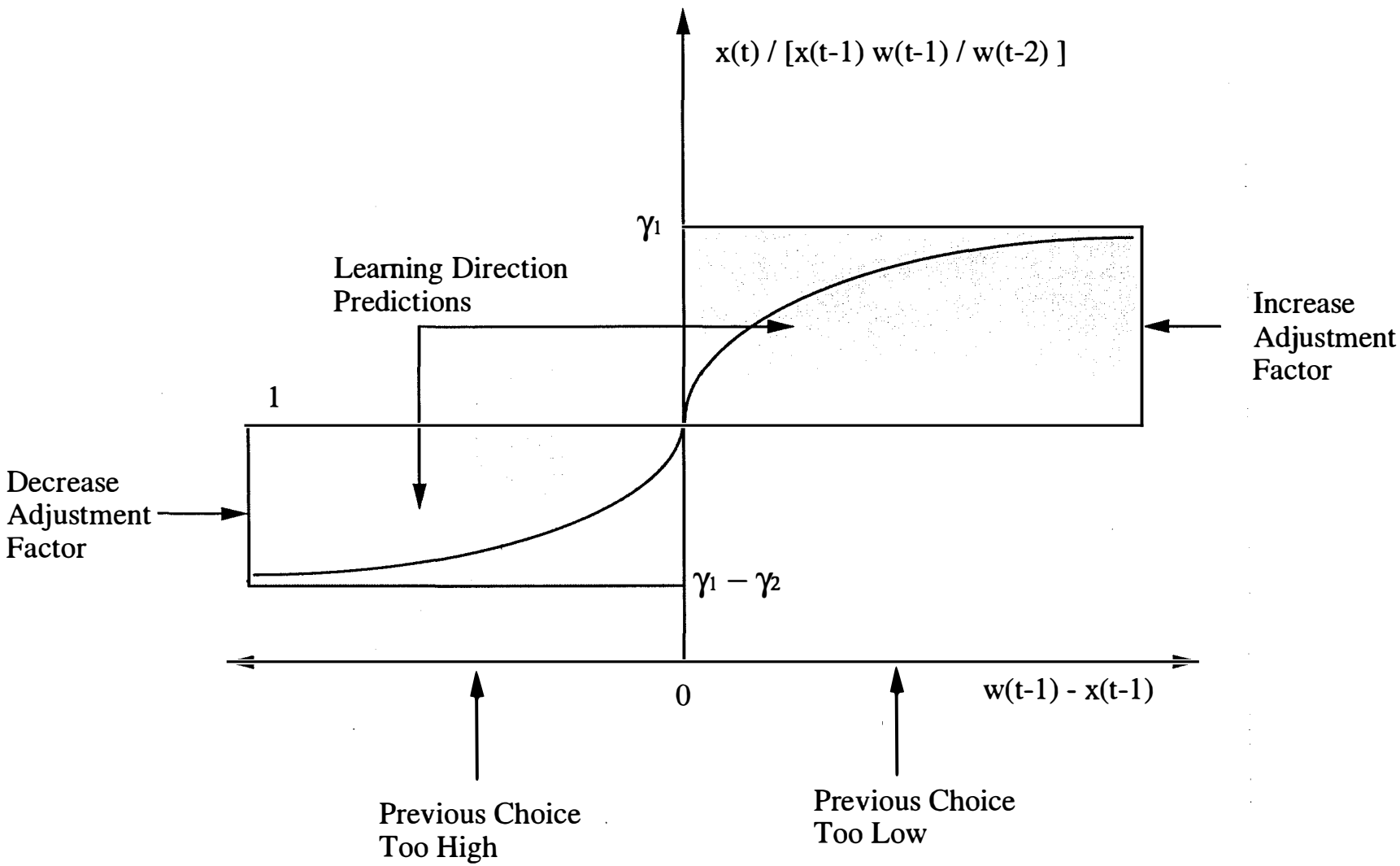


Figure 4: Learning Direction Theory Predictions