Selection of Learning Rules: Theory and Experimental Evidence[1]

David J. Cooper
Dcooper@vms.cis.pitt.edu
Department of Economics
University of Pittsburgh
Pittsburgh, PA 15260

Nick Feltovich
Niklas@vms.cis.pitt.edu
Department of Economics
University of Pittsburgh
Pittsburgh, PA 15260

May 1997

Abstract: The goal of this paper is to formulate and apply a rule for mapping between experimental designs and learning rules. We propose the use of a hierarchy of learning rules, with rules higher on the hierarchy possessing a greater degree of cognitive sophistication, and suggest that the learning model with the least cognitive sophistication should be used unless certain well-specified criteria are met for moving to a more sophisticated model. These ideas are applied to data from limit pricing experiments (Cooper, Garvin, and Kagel, 1996, 1997, and in preparation). We compare the abilities to characterize the data of a reinforcement-based learning model (Roth and Erev, 1995) and a belief-based learning model (Cooper, Garvin, and Kagel, 1996 and 1997), and find that the belief-based model outperforms the reinforcement-based model. This is not due to some universal superiority of belief-based learning models. Rather, the belief-based model's greater cognitive sophistication makes it the appropriate model for the limit pricing data.

[1]We would like to thank Richard Boylan, John Duffy, Ido Erev, John Kagel, Al Roth, and Bob Slonim for their helpful comments.

1. Introduction

In recent years, experimental economics has witnessed a boom in papers on learning in games. A number of authors have presented results demonstrating that learning models are able to better characterize experimental data than more standard static approaches.[2] Because the focus has been on the refutation of existing theory, a plethora of learning models has been developed with little concern for their differing predictions. Experimental economists now need to begin sorting out among the various learning models, and to determine some sort of mapping between games and appropriate learning models.

In this paper, we compare the ability of several different learning models to capture the main features of a specific experimental data set. We do not intend to conduct a horse race between the models; although we reach strong conclusions about what model works best for the data set in question, there is no reason to believe that this is generically the best learning model (or that any such model exists). Instead, our goal is to present a general approach for assigning learning models to data sets, to apply this approach to a particular data set, and to explain what the superior performance of one particular learning model tells us about the behavior of subjects in this particular set of experiments.

We believe Occam's Razor should play a central role in determining which learning model is applied to some particular experimental data set. Commonly employed learning models fall into a clear hierarchy when ranked by the level of cognitive sophistication required of players. While the term "cognitive sophistication" can be given a specific definition, as will be done in Section 5, it is best understood in an intuitive sense: one learning rule is more cognitively sophisticated than another if it requires an individual to use more information or make more complex calculations. Our approach to

---

[2]For examples, see Miller and Andreoni (1991), Crawford (1991, 1995), Brandts and Holt (1994), McKelvey and Palfrey (1994, 1995), Cooper, Garvin, and Kagel (1996, 1997), Cheung and Friedman (1997), Gale, Binmore, and Samuelson (1995), Duffy and Feltovich (1996), and Roth and Erev (1995). There also exists a broad theoretical literature on learning. For a survey of this literature, see Fudenberg and Levine (1996). For a regularly updated list of learning references on the Worldwide Web, see http://www.pitt.edu/~alroth/alroth.html.

mapping between learning models and experimental data sets is that a small number of learning models should be employed, each representing a different level of cognitive sophistication. The simplest model in terms of cognitive sophistication should be used first, and more demanding models should be employed only as the simpler ones fail. The models to be considered and a well-specified criterion for moving up the hierarchy should be specified *before* considering any specific experiment.

Our approach differs quite significantly from the standard approach of most game theorists. When a game theorist predicts behavior in terms of equilibrium and equilibrium refinements, he/she is implicitly assuming that players will act with the maximum possible rationality. If the data does not support this assumption, the model is adjusted as little as possible in order to better characterize the data. We take an approach which is almost diametrically opposed. If a simple, cognitively undemanding model characterizes play of a game, we stop there. For many games, simple models give us all the insight we need into how subjects play. However, if a game has features which *ex ante* seem to call for greater cognitive sophistication, and if *ex post* a simple model is unable to characterize the data while a more sophisticated model can, we move up the hierarchy to more cognitively demanding models. In other words, a more sophisticated model is employed if the simpler model fails to describe what is actually played because it attributes too little reasoning ability to players. Instead of trying to find the maximum amount of reasoning ability which is consistent with subject behavior, we try to find the minimum amount of reasoning ability necessary.

It is critical that both criteria, *ex ante* and *ex post*, are fulfilled. For any data set, we can no doubt find some way of modifying a basic learning model which results in a better fit. The goal is not simply to find which model best fits the data, but rather to find which model best explains how subjects might reason about a game. When a feature is added to a learning model, it should imply some significant change in how subjects reason about games, this change should be relevant for the game in question, and adding this feature should allow us to better explain the major qualitative features of the data set.

To determine whether a learning model does a good job of characterizing data, our methodology is to run simulations and then visually compare simulation output with the experimental data. This works well for the data set in question; the data have a number of striking features, and the simulations differ in obvious ways. An alternative methodology is the use of maximum likelihood estimation (MLE) techniques. For example, Camerer and Ho (1997) present a general learning model for which reinforcement-based learning and belief-based learning are special cases.[3] By fitting parameters for this model, they are able to measure whether learning in a data set more closely resembles reinforcement-based learning, belief-based learning, or neither. We could use MLE techniques, but have chosen to not do so. While this is partially a matter of taste, our choice is mainly driven by the goals of this paper. Rather than trying to best fit the data, we are interested in trying to identify how much reasoning ability is needed to characterize subject's behavior. The models we consider represent distinctly different levels of cognitive sophistication. By selecting one particular model, we reach clear conclusions about the level of reasoning employed by subjects. With an MLE approach, the parameter estimates typically don't tell any clear story about how subjects think about the game. MLE analysis is a wonderful quantitative tool, but it is not the best tool for answering this particular question.[4]

---

[3]See also Boylan and El-Gamal (1993), Holt (1993), and Mookherjee and Sopher (1997).

[4]Roth, Erev, Slonim, and Bereby-Meyer (1997) criticize the MLE approach on two grounds. First, the results of any MLE are sensitive to the specification. If the model being fitted is misspecified (for example by neglecting individual effects) the results can be wildly inaccurate. Roth *et al* claim that the simulation approach is less sensitive to specification. Second, MLE picks the parameters which make the observed data the most probable. As such, unusual events can unduly affect the estimates, and subtle trends may be missed. Roth *et al* claim that the simulation approach does a better job of capturing the major qualitative features of data sets. Our feeling is that no conclusive evidence exists on any of these issues, but that serious problems may exist with the MLE approach. Since neither of us is an econometrician, we are happy to leave these difficulties to others by avoiding the MLE approach.

The simulation approach is often criticized as encouraging an ad hoc approach in which results depend strongly on the choice of parameters and/or specification. This is a valid criticism which can only be answered through careful sensitivity analysis. Roth and Erev (1997) have done extensive sensitivity analysis for the Roth-Erev model in other games, and Cooper, Garvin, and Kagel (1997) have done extensive sensitivity analysis for the CGK model with limit pricing games. For this project, we have considered numerous specifications and parameter

We apply our approach to data from two treatments of signaling game experiments conducted by Cooper, Garvin, and Kagel (1996, 1997, in preparation). These treatments differ primarily in how difficult it is for subjects to recognize the existence of dominated strategies. We consider the learning models developed by Roth and Erev (1995) and by Cooper, Garvin, and Kagel (1996 and 1997). (Henceforth referred to as the RE and CGK models.) These are good representatives of, respectively, reinforcement-based learning models and belief-based learning models. These particular models are chosen due to their previous use in exploring experimental data, but comparable conclusions can be reached using other models from the literature with similar cognitive sophistication.

The RE model makes quite low demands on the rationality and reasoning ability of players. Players are simply required to remember what worked well in the past, and to do it more frequently in the future. The CGK model is more cognitively demanding. Players are explicitly required to set up a probability distribution over opponents' actions, update their beliefs, and maximize given their beliefs. For each model, we consider unaugmented and augmented versions. The augmented versions allow for limited strategic foresight; some players are able to recognize their opponent's dominated strategies, and anticipate that these strategies will never be used. Given our preceding arguments, the RE model should be lower in a hierarchy of learning models than the CGK model, an unaugmented version of a model should be low than an augmented version, and there is no obvious ranking between an augmented version of the RE model and an unaugmented version of the CGK model. These statements are indeed true given the definition of cognitive sophistication introduced in Section 5.

The unaugmented RE model is able to characterize results from one of the treatments, but does poorly with data from the second treatment. Moreover, the unaugmented RE model is only partially able

values for our analysis. At this point in time there exists no reason to believe our results are dependent on our choices of specification for the learning models or of parameters.

It also must be noted that we are looking at the data on an aggregate level. There is obviously a great deal of heterogeneity among subjects. While we may capture the level of cognitive sophistication for an average subject, there will certainly be subjects who are either more or less sophisticated.

to capture differences between the two treatments, and does so for reasons which cannot possibly be affecting the experimental data. Augmenting the RE model does not substantially improve its performance. The unaugmented CGK model also characterizes the experimental data quite poorly. However, the augmented CGK model captures the major features of subject behavior in both treatments, and captures differences between data from the two treatments.

The superior performance of the augmented CGK model in characterizing the limit pricing data is driven by two factors closely related to cognitive sophistication. First, the dynamics in the limit pricing game data are driven by sharp changes in behavior. The CGK model includes maximization; small differences in payoffs between strategies can yield large changes in behavior. In contrast, the RE model does not include maximization. The relative frequency of strategies is determined by the relative accumulated payoffs for the strategies, and small differences in payoffs between strategies can only slowly lead to changes in behavior. Therefore, the CGK model captures the large swings in behavior observed in the limit pricing game data better than either version of the RE model.

Second, a key role is played in the limit pricing game by dominated strategies. In particular, the recognition of dominated strategies plays a key role in determining whether certain strategies will induce entry. Without augmentation, the CGK model treats dominated strategies identically in the two treatments, and is therefore unable to capture differences between the treatments. With augmentation, the CGK model is able to explicitly incorporate differing degrees of recognition of the dominated strategies, and thus able to capture the differences between treatments. Augmentation does not play an analogous role in the RE model. To get the RE model to track any of the limit pricing data, it is necessary to include experimentation. Since players in the RE model are not optimizing, experimentation induces persistent play of dominated strategies. This persistent play of dominated strategies, which is never observed in the experimental data, takes on the same role as limited strategic foresight in driving differences between treatments. Thus, augmentation is superfluous in the RE model.

We conclude that both the unaugmented CGK model, which includes maximization but not limited strategic foresight, and the augmented RE model, which includes limited strategic foresight but not maximization, fail to characterize the limit pricing data. Only the augmented CGK model, which includes both maximization and limited strategic foresight, is able to do so.

We do not argue the unaugmented RE model is the "wrong" model and that the augmented CGK model is the "right" model for describing learning. Rather, we argue that both of our criteria for moving to a more sophisticated learning model are satisfied by the limit pricing game. *Ex ante*, this game evokes a degree of cognitive sophistication which the unaugmented RE model is not designed to capture. Dominated strategies play an obvious role in the limit pricing game, and the ability to sharply differentiate among strategies with similar payoffs plays a more subtle role. This makes it reasonable *ex ante* to move up the hierarchy, and employ the more demanding CGK model. *Ex post*, this move is justified by the greater ability of the augmented CGK model to characterize the experimental data. By using a systematic approach to assigning learning rules to experimental treatments, we can give a coherent explanation of why this particular game should demand a more cognitively sophisticated learning rule while other games do not.

The paper begins by presenting the limit pricing game, and summarizing the results of Cooper *et al.*'s experiments with this game. We then present the alternative learning models and fit them into a proposed hierarchy of cognitive sophistication. We demonstrate that only the augmented CGK model is able to track the limit pricing game data. The paper concludes by discussing this result.

2. The Limit Pricing Game

The limit pricing game is loosely based on Milgrom and Roberts' (1982) model of entry limit pricing. This model is of a two-period game involving a market with a homogeneous good and a linear market demand curve. There are two players, a monopolist M and a (potential) entrant E. The monopolist is either a high-cost (M_H) or low-cost (M_L) type. The probability of either type is 50%. The monopolist observes her type before the game begins. The entrant does not know the monopolist's type, but the probability of each type is common knowledge.

In the first period, M chooses a production level which is seen by E. E then decides whether or not to enter in the second period. If E enters, the firms compete as Cournot oligopolists, while if E stays out, M is an uncontested monopolist. M is always better off with no entry. Payoffs are such that if E knew M's type, he or she would enter against M_H but stay out against M_L. Since E does not know M's type, there is an incentive for M to limit price; in other words M has an incentive to signal her type if she is low-cost and to conceal her type if she is high-cost. Both types of signaling involve outputs above full-information levels, and hence pricing below full-information levels.

For their experiment, Cooper *et al.* collapse this game into a standard signaling game by imposing the Cournot outcome in the second stage if entry occurs, and the profit-maximizing monopoly outcome in the second stage if no entry occurs. They also discretize the M player's set of available first-period actions to the set {1,2,3,4,5,6,7}. The E player's action set subject to M's choice becomes {IN,OUT}, and payoffs are according to Tables 1 and 2. Cooper *et al.* use multiple variations of the limit pricing game, but in this paper we limit ourselves to consideration of the two most basic treatments, the "Standard Payoff" (SP) treatment and the "Zero Anticipation" (ZA) treatment. The treatments differ solely in the payoffs assigned to 6 and 7 for M_Hs. These strategies give negative payoffs in the SP treatment and positive payoffs in the ZA treatment.

The equilibrium sets are identical for the SP and ZA treatments. There are two pure-strategy sequential equilibria, both of which are separating equilibria. These are the Riley (1979) outcome, in which M_H chooses 2 and M_L chooses 6, and an inefficient equilibrium in which M_H chooses 2 and M_L chooses 7. In both of these equilibria, E chooses IN except when seeing M_L's equilibrium choice. There are also many partial-pooling (mixed-strategy) equilibria, one of which has drawing power in both the experiments and the simulations; in this equilibrium, M_L chooses 5 with probability 1, M_H chooses 5 with probability 1/5 and 2 with probability 4/5, and E chooses IN with probability 1 after seeing anything but 5, and OUT with probability 1/9 after seeing 5. The Riley outcome is the natural equilibrium of the limit pricing game in the sense that it is the only equilibrium which fulfills the Cho-Kreps intuitive criterion (Cho and Kreps, 1987).

Plays of either 6 or 7 by M_H's are strictly dominated strategies in either treatment of the limit pricing game. These dominated strategies play a central role in our analysis of the game. If players perform iterated removal of dominated strategies on the game, the only surviving sequential equilibrium in the reduced game is the Riley outcome. This leads directly to the Riley outcome being the sole equilibrium to pass the intuitive criterion.

3. Experimental Results

Cooper *et al.* (1997) conducted two inexperienced-subject sessions with the SP treatment, one lasting 24 periods and one lasting 36 periods, using the payoffs in Tables 1 and 2. An additional four inexperienced-subject SP sessions lasting 36 periods were conducted in Cooper *et al.* (in preparation). A notable feature of the experimental design is that players were able to observe the outcomes of all players' games (with anonymity). Beyond this, procedures were completely standard.[5] Figure 1a reproduces pooled data from the six inexperienced subject SP sessions. We report probability distributions over M player strategies in twelve period cycles; rounds 1 - 12 are cycle 1, rounds 13 - 24 are cycle 2 and rounds 25 - 36 are cycle 3. Entry rates are given by the numbers above the M players' strategies. Play by M players starts off near the myopic maxima (the action an M would choose if she did not believe that her opponent would condition his response on this action -- 4 for M_L's and 2 for

[5]See Cooper *et al.* (1996) for details of the treatments. The treatments in Cooper *et al.* (in preparation) differ only slightly from those in Cooper *et al.* (1996 and 1997), and are included to increase the sample size. See Cooper *et al.* (in preparation) for details of how the treatments varied.

M_H's). After a few rounds, as the entry rate on 2 rises, M_H's attempt to pool with M_L's (even though this game has no pooling equilibrium) by choosing 4 with greater frequency. The inexperienced-subject sessions ended here, with 4 the modal choice by both types of M.

One session, lasting for 36 periods, was conducted with experienced subjects in Cooper *et al.* (1997). Figure 1b reproduces data from this session. The qualitative features of this session have been replicated in Cooper *et al.* (in preparation) using somewhat different treatments.[6] At the beginning of the experienced subject session, M_H's start out playing 2 and 4 with approximately equal probability, and M_L's start out playing 4 and 6 with approximately equal probability. The high proportion of M_H's playing 4 pushes up the entry rate on this strategy, giving M_L's an incentive to limit price. This moves play toward the Riley outcome. By the end of the session, 2 and 6 are the modal choices of M_H's and M_L's, respectively.

If we consider the inexperienced- and experienced-player sessions to be two halves of a longer session, we see players starting at their myopic maxima, attempting to pool at 4, and finally approaching a separating equilibrium at 2 and 6. Looking at some details which play important roles in the simulations, there are relatively few attempts by M_H's to limit price at 3 instead of 4, and the dominated strategies (6 and 7) are virtually never chosen by M_H's.

Cooper *et al.* (1996) conducted two inexperienced subject ZA treatment sessions lasting 36 rounds. Two additional inexperienced-subject ZA sessions lasting 36 periods were conducted in Cooper *et al.* (in preparation). Pooled data from these sessions are reproduced in Figure 2a. Cooper *et al.* (1996 and in preparation) also conducted two experienced subject ZA sessions lasting 36 rounds. Pooled data from these sessions are reproduced in Figure 2b.

As in the SP treatment, the ZA treatment experimental data show M_H's quickly attempting to

---

[6]In the experienced sessions from Cooper *et al.* (in preparation), players played the SP treatment for the first 12 - 18 rounds. During these periods, play converges steadily to the Riley outcome. Players were then switched to an alternative treatment of the limit pricing game to test their ability to transfer what they had learned.

pool at 4 and then gradually returning to the myopic maximum, 2. M_L's, on the other hand, play very differently than in the SP treatment. There is almost no play of 6, as they instead play the myopic maximum 4 with high frequency throughout the inexperienced and experienced sessions. Attempts at limit pricing are less frequent and occur much later in the ZA treatment than in the SP treatment. Also, most attempts at limit pricing are choices of 5 rather than 6.

The limit pricing data are a relatively complex data set to explain. It is this complexity which makes the limit pricing data an interesting case to consider. There exist simpler games in which similar features play an important role; for example, subjects' perceptions of dominated strategies play a central role in the experiments presented by Cooper, DeJong, Forsythe, and Ross (1991). However, the simplicity of these games make it unlikely that anything more complex than hill-climbing given by a simple reinforcement-based model is necessary to explain the observed dynamics; any effects due to cognitive sophistication are being observed solely in the initial distribution of strategies rather than in the dynamics of learning. Only in a more complex setting like the limit pricing experiments is behavior sufficiently rich that a learning model beyond simple reinforcement is likely to be required. Thus, the limit pricing data provides an interesting case for application of our approach to selecting learning rules.

4. The Learning Models

The data from Cooper *et al.* (1996, 1997, and in preparation) have a clearly identifiable dynamic pattern, and yield striking differences between the SP and ZA treatments. Using the approach to selecting a learning rule described in the introduction and the definition of cognitive sophistication introduced in Section 5, we will examine which of four learning models is best able to capture the features of the experimental data. In this section, we introduce the four candidates: basic and augmented versions of the RE model and the CGK model.

The Basic and Augmented RE Models: The basic RE model, adapted for the limit pricing game, is as follows: there are 16 simulated players, 8 M players and 8 E players. In round t, the i^{th} M player has propensities $q_{i,t}^M(j|L)$ for choosing action $j \in \{1,2,...,7\}$ in the event that she is of type M_L, and $q_{i,t}^M(j|H)$ for choosing action j in the event that she is of type M_H. The i^{th} E player has propensities $q_{i,t}^E(IN|j)$ and $q_{i,t}^E(OUT|j)$ for choosing IN and OUT, respectively, after seeing his opponent choose action j. We define the strength of propensities at an information set to be the sum of propensities for all actions at that information set:

$$S_{i,t}^M(X) = \sum_{k=1}^{7} q_{i,t}^M(k|X) \qquad (1)$$

for the i^{th} M player, where $X \in \{H,L\}$ is her type, and

$$S_{i,t}^E(j) = q_{i,t}^E(H|j) + q_{i,t}^E(L|j) \qquad (2)$$

for the i^{th} E player where $j \in \{1,2,...,7\}$ is his opponent's action. The probability of a player choosing a particular action at a given information set is obtained by dividing the associated propensity by the strength of propensities at that information set. For M's,

$$p_{i,t}^M(j|X) = \frac{q_{i,t}^M(j|X)}{S_{i,t}^M(X)} \qquad (3)$$

for each action $j \in \{1,2,...,7\}$ and type $X \in \{H,L\}$. An analogous formula is used for E's. In round t, each M is randomly assigned a type and an opponent, then chooses her action based on her round-t probabilities (given her type). Each E then chooses his action based on his round-t probabilities (given his opponent's action). Payoffs are determined by M's type, M's action, and E's action as in Tables 1 and 2.[7] Propensities are updated for each M by adding her payoff to her propensity for choosing her

[7]Because negative propensities result in probabilities outside [0,1], care must be taken in order to eliminate the possibility of negative propensities. This has been done in these simulations by adding 292 to every payoff in the M players' payoff table. This changes neither the game-theoretic predictions nor the CGK predictions, although it is

realized action given her type, and for each E by adding his payoff to his propensity for choosing his realized action given his opponent's action. Thus, if in round t, the i^{th} M player was of type X, chose action j, and received payoff $\omega_{i,t}^M$, her propensity for choosing j given that she is of type X would be modified:

$$q_{i,t+1}^{M}(j|X) = q_{i,t}^{M}(j|X) + \omega_{i,t}^{M} \quad (4)$$

E players' propensities are updated in a similar manner.

For our simulations, we use a variation of the basic model. This variation differs from the basic model in that three additional parameters are present. The first is a "global experimentation" parameter ϵ_1. At the end of a round, a M player's propensity for playing the action that she chose, given her type, is increased by the resulting payoff times $1-\epsilon_1$, and her propensities for playing all other actions (again, given her type) are increased by that payoff times $\epsilon_1/6$. This "global experimentation" differs from the "local experimentation" used in Roth and Erev (1995) in that the propensities for all possible actions are increased, rather than only those neighboring the action played in the previous round. We put off until Sections 6 discussion of why global experimentation is included in the model rather than local experimentation or no experimentation. The second parameter added to the basic model is a "forgetting" parameter ϵ_2: at the end of a round, all propensities are multiplied by a factor of $(1 - \epsilon_2)$. This ensures that propensities do not become arbitrarily large with time, so that learning is still possible even after many rounds. Finally, we added an "imitation" parameter, ϵ_3. This captures the ability of subjects in the limit pricing experiments to observe not only their own outcomes, but also the outcomes of other players. At the end of a round, each M's propensity for choosing an action that was played by a different M in the previous round is increased by the payoff that the observed M received, times ϵ_3. The primary

quite possible that experimental results based on this scaled game would differ from those in the unscaled game.

effect of adding imitation to the simulations is speeding up convergence to equilibrium.[8] With addition of these three parameters, the updating rule for the unaugmented RE simulations becomes: if in round t, the i^{th} M player was of type X, chose action j, and received payoff $\omega_{i,t}^{M}$, her propensities for the next round are

$$q_{i,t+1}^{M}(j|X) = (1-\epsilon_2)[q_{i,t}^{M}(j|X) + (1-\varepsilon_1)\omega_{i,t}^{M} + \sum_{\substack{\ell \neq i, \\ \ell \text{ was type X} \\ \text{and chose j}}} \varepsilon_3 \omega_{\ell,t}^{M}], \tag{5}$$

$$q_{i,t+1}^{M}(k|X) = (1-\epsilon_2)[q_{i,t}^{M}(k|X) + \frac{\varepsilon_1}{6}\pi_{i,t}^{M} + \sum_{\substack{\ell \neq i, \\ \ell \text{ was type X} \\ \text{and chose k}}} \varepsilon_3 \omega_{\ell,t}^{M}] \tag{6}$$

for k≠j, and

$$q_{i,t+1}^{M}(k|Y) = (1-\epsilon_2)[q_{i,t}^{M}(k|Y) + \sum_{\substack{\ell \text{ was type Y} \\ \text{and chose k}}} \varepsilon_3 \omega_{\ell,t}^{M}] \tag{7}$$

for Y≠X. Propensities for E players are updated analogously.

For each of the RE simulations, the initial strength of propensities equals 1000. Each simulation uses the same initial propensities for both types of players, which are estimated from early-round play by inexperienced players in the experimental sessions.[9] The value of ε_3 was set to .25, the value of ϵ_2 was

---

[8] See Roth and Erev (pp 173-176) for a description of some advantages of adding experimentation and forgetting to the model, as well as a few disadvantages.

[9] We used high-cost and low-cost M player actions from rounds 1-3 of Cooper *et al.*'s (1996, 1997, and in preparation) inexperienced-subject sessions to obtain M_H and M_L initial probabilities. Data were pooled from sessions with the SP treatment and the ZA treatment. To obtain E initial probabilities, we used E player responses from rounds 1-3 of Cooper *et al.*'s inexperienced-subject sessions, and pooled some responses (as was done by Roth and Erev (pp. 195) for their ultimatum game simulations) to ensure that initial entry rates were nonincreasing in monopolist quantity. We did not pool data from SP treatments and ZA treatments in obtaining E initial propensities.

set to .003, and the value of ε_1 decreased with time according to the following formula:

$$\varepsilon_1 = \text{Max}[.1 \times 2^{-t/10,000}, .00001] \qquad (8)$$

where t is the current round. Sensitivity analysis indicates that none of these parameter values are central to our primary conclusions.

Although it is not entirely natural, the RE model can be augmented to allow for some players possessing limited strategic foresight. We put off discussion of whether we should augment the RE model until Section 8, and concentrate on how to augment the RE model. In doing this, it is useful to first think about augmenting the CGK model. Because the CGK model involves optimization, players will never choose the dominated strategies (6 or 7 for a M_H). In augmenting the CGK model to allow for limited strategic foresight, we restrict the beliefs of E's to put zero weight on play of 6 or 7 by M_H players. Given that E's are always better off not entering versus a low cost monopolist, it follows from this restriction that E's will never enter following play of 6 or 7. Thus, limited strategic foresight could have been modeled as a restriction on the strategies available to M_H's and the strategies available to E's. This is the approach we follow in augmenting the RE model. In keeping with a single iteration of elimination of dominated strategies, all M players are modified to never use the dominated strategies. Using our preceding notation, $q_{i,t}^M(6|H) = q_{i,t}^M(7|H) = 0$ for all i and all t.[10] In addition, some percentage τ of the players possesses limited strategic foresight, and thus anticipates the elimination of dominated strategies. Such players put zero probability on entry following play of 6 or 7 as an E. If the i^{th} E player

---

The first three periods were used in order to decrease the effects of random noise on initial propensities. We also ran simulations using actions from only round 1 to obtain initial propensities and found no qualitative differences. Using the robust rank order test (Siegel and Castellan, 1988, pp. 137-144), we found no significant differences by either M_H or M_L players in rounds 1 - 3 between play in the ZA and SP treatments. Simulations which did not pool M data from the two treatments in forming initial propensities do not yield qualitatively different results. We did find a significant difference for E's in rounds 1-3 between play in the ZA and SP treatments (at the 5% level). Pooling E data from the two treatments in calculating initial propensities has little effect on simulation outcomes.

[10]We could have made a less extreme modification to the RE model by only restricting some (rather than all) M players to not use 6 or 7 as high cost types. The affect of making such an adjustment is discussed below.

has limited strategic foresight, $q_{i,t}^{E}(IN|6) = q_{i,t}^{E}(IN|7) = 0$ for all t. This is consistent with the E player anticipating that no M_H will ever use 6 or 7, making entry on 6 or 7 a dominated strategy.

The Basic and Augmented CGK Models: Cooper *et al.* (1996 and 1997) simulate a population playing the limit pricing game using a learning model loosely based on fictitious play. As such, the CGK model falls under the rubric of belief-based learning models. We provide only an abbreviated description of the CGK model; for more detail, see Cooper *et al.* (1997). There are 20 M players and 20 E players in the population.[11] At time t, the i^{th} M player has conjectures (beliefs) regarding the likelihood with which her opponent will enter subject to M's action. Her conjectures are represented by a 7 x 2 matrix of weights $v_{ijk}^{M}(t)$, where j is her action and $k \in \{1,2\}$ is her opponent's response; $v_{ij1}^{M}(t)$ is the weight she places on her opponent entering after she plays j, and $v_{ij2}^{M}(t)$ is the weight she places on her opponent staying out after she plays j. The probability she assigns to her opponent entering after play of j, $p_{ij1}^{M}(t)$, is given by (1). The probability she assigns to her opponent staying out after play of j, $p_{ij2}^{M}(t)$, equals 1 - $p_{ij1}^{M}(t)$.

$$p_{ij1}{}^{M}(t) = \frac{v_{ij1}{}^{M}(t)}{v_{ij1}{}^{M}(t) + v_{ij2}{}^{M}(t)} \tag{9}$$

In round t, the i^{th} E player has beliefs regarding the likelihood that his opponent is a high- or low-cost type conditional on her output level. These beliefs are represented by a 7 x 2 matrix of weights $v_{ijk}^{E}(t)$, where j is his opponent's action and k is her type; $v_{ij1}^{E}(t)$ is the weight he or she places on his opponent being high-cost conditional on her choosing action j, and $v_{ij2}^{E}(t)$ is the weight he or she places on his opponent being low-cost conditional on her choosing action j. Probabilities $p_{ij1}^{E}(t)$ and $p_{ij2}^{E}(t)$ are calculated analogously to those for M players.

In the CGK simulations, weights for the first round are assigned to each M player and each E

[11]In the actual experiments, there are fewer players and players rotate between roles. Adding these features to the CGK simulations has no significant effect on the results.

player randomly, with the initial distribution of weights fitted to the first period data. The procedure for fitting this distribution and generating first round weights is given in Cooper *et al.* (1997). We have followed this procedure exactly using data from Cooper *et al.* (1996, 1997, and in preparation). At the beginning of round 1, each M player is assigned a type (there is a 50-50 chance of either type being chosen). Each M player chooses the action that maximizes her expected payoff given her type and her conjectures about her opponent's response. The E players are randomly matched to the M players. Each E player observes his opponent's action, and then chooses the response that maximizes his expected payoff given his beliefs and his opponent's action. At the end of round 1, players update their expectations by adding 1 to the appropriate element of their weight matrix for each action observed. For example, suppose an M player was matched to an E player, the M player was low-cost and played 4, and the E player responded with OUT. Then $v_{i42}^{M}(2)=v_{i42}^{M}(1)+1$ and $v_{i42}^{E}(2)=v_{i42}^{E}(1)+1$ for all $i \in \{1, ..., 20\}$, and probabilities are recalculated accordingly. Note that all players observe each action taken, so all players update even if they were not directly involved in the pairing. This process is repeated, generally for 60 rounds, with new types and pairings being generated each round and beliefs updated at the end of each round.

Thus far, we have described the unaugmented CGK model. The CGK model is augmented by adding an additional iteration of elimination of dominated strategies. Due to the use of optimization, M_H's never use the dominated strategies in the unaugmented CGK model. This gives a single iteration of removal of dominated strategies. What augmentation adds is that (some) E's recognize that M_H's have two dominated strategies, anticipate that these strategies will never be used, and modify their strategies accordingly. (We refer to this as limited strategic foresight.) These E's realize that play of either 6 or 7 could only be coming from an M_L, and thus play of IN following either 6 or 7 is a strictly dominated strategy. Since the unaugmented CGK model explicitly contains players' conjectures about the behavior of their opponents, we incorporate the ability of some players to anticipate elimination of dominated strategies into the model by having some E players' initial conjectures put no weight on play of 6 or 7 by M_H's. M players' beliefs about the behavior of E players are not modified; this would imply an additional iteration of removing dominated strategies.

In keeping with this analysis, Cooper *et al.* (1997) add an additional parameter to their model, giving the augmented CGK model. They assume that some proportion of E players, π, recognize that if 6 or 7 is played, it must have been played by an M_L type, and thus do not enter; the parameter τ in the augmented RE model is analogous to the parameter π in the augmented CGK model. In our notation, if the ith E player has limited strategic foresight, $v_{i61}^E(t)=v_{i71}^E(t)=0$ for all t. The augmented model with π = 0 is identical to the unaugmented model.

5. Cognitive Sophistication and a Hierarchy of Learning Rules

Our approach to selecting a learning requires a ranking of the four models introduced in Section 4 by "cognitive sophistication." In the introduction, this term was loosely defined. The goal of the current section is to give a more concrete definition of what it means for one learning rule to be more cognitively sophisticated than another. The approach we take here is non-technical in nature, concentrating on examples to illustrate our meaning. While the same concepts can be discussed in a more technical setting, doing so adds a great deal of notation for little gain in understanding.[12]

Loosely, a learning rule is defined as a function mapping an individual's observed history of play into his/her current action. Virtually the only restriction implied by this definition is that an individual's behavior cannot be conditioned on actions or outcomes not observed by that individual.

In general, any learning rule which requires players to perform a more complex task should have higher cognitive sophistication. The ideal definition of cognitive sophistication would capture all

[12]For a technical treatment, see Cooper (1995).

possible dimensions of complexity, including what types of information and calculations were required to implement a learning rule. Such an all-encompassing definition is well beyond what is needed for our narrow purposes. We do not attempt to give a universal definition of cognitive sophistication, but instead concentrate on two factors which are likely to be central considerations for learning in a game-theoretic setting: the amount of information a player needs to know about his/her opponent to implement a rule, and whether or not a player needs to be able to solve an optimization problem. Suppose we are considering two learning rules, Rule A and Rule B. We apply the following definition.

Definition 1: Rule A is higher in the hierarchy of cognitive sophistication than Rule B if either of the following two criteria holds.

Criterion 1) *Rule A requires more information about the opposing player than Rule B for implementation. If Rule B requires optimization, Rule A also requires optimization.*

Criterion 2) *Rule A requires at least as much information about the opposing player as Rule B for implementation. Rule A requires optimization while Rule B does not.*

The application of Definition 1 is best illustrated by ranking the learning rules we have already defined. The unaugmented RE model is at the very bottom of our hierarchy of cognitive sophistication. Implementing this rule does not require any information about the opposing player -- players need not even realize that a game is being played. The only necessary information is the history of strategies played and realized payoffs. Optimization is not used. The unaugmented CGK model requires optimization and requires information about what strategies the opposing players have used. By either Criterion 1 or Criterion 2, the unaugmented CGK model is more cognitively sophisticated than the unaugmented RE model.

Ordering the augmented and unaugmented RE models requires a careful definition of augmentation. Cognitive sophistication can only make meaningful distinctions among learning rules if learning rules are defined in ways which are not game-specific. For example, an important feature of the augmented RE model is that it implicitly requires some E's to possess limited strategic foresight; these players recognize that play of 6 or 7 is dominated for M_H's, and thus could only come from M_L's. On a more concrete level, some E's never enter following play of 6 or 7. This augmentation could be made in a very game-specific manner, without implying any rationale behind the elimination of certain strategies or that play for any other game is altered. Thus, we could define the augmented RE model to be identical to the unaugmented RE model except that in the limit pricing game, for some unmodeled reason, 6 and 7 are never played by M's and some E's never enter on these strategies. While these happen to be the strategies which a player with limited strategic foresight would eliminate, a player need not know which strategies are dominated for his/her opponent in order to follow a learning rule which eliminates play of some strategies. Augmenting the RE model in this manner does not imply any increase in cognitive sophistication over the unaugmented RE model, nor does it imply any change in behavior for other games. We prefer to think of a learning rule as giving an algorithm for behavior in any game; a learning rule which is an arbitrarily different rule for every game is not really a rule at all. This does not mean that different games cannot elicit different learning rules, but rather that defining learning rules and assigning rules to specific games are two separate issues. Learning rules are defined for all games, but behavior in specific games may only be consistent with some learning rules. Returning to the augmented RE model, a more general method of augmentation than simply eliminating specific strategies in the limit pricing game is to specify that some players never use strategies removed by iterated elimination of dominated strategies *in any game*. Making this modification has implications not just for how the limit pricing game will be played, but how all games will be played. For the limit pricing game, this augmentation implies "never enter following 6 or 7 in the limit-pricing game," as did the less general augmentation proposed previously However, it also implies an increased level of cognitive sophistication -- eliminating dominated strategies requires knowledge of the opposing player's payoffs (and implies knowledge that the opposing player is rational). Thus, the augmented RE model is more cognitively sophisticated than the unaugmented RE model.

It is not possible to order the cognitive sophistication of the unaugmented CGK model and the augmented RE model. The unaugmented CGK model requires optimization which the augmented RE model does not, while the augmented RE model requires information about the opposing player's payoffs (to determine which strategies are dominated) which the unaugmented CGK model does not.

The augmented CGK model is the most cognitively sophisticated of the models we consider. It requires optimization, which the augmented RE model does not, and knowledge of which strategies are dominated for the opposing player, which the unaugmented CGK model does not. The augmented CGK model is thus more cognitively sophisticated than the augmented RE model by Criterion 2 and more cognitively sophisticated than the unaugmented CGK model by Criterion 1.

These criteria may seem a bit arbitrary. For example, consider ranking the unaugmented CGK and RE models. The first is more cognitively sophisticated than the second by both Criterion 1 and Criterion 2, yet it can be argued that these rules barely differ in the amount of information required. While the unaugmented RE model uses no information about the opposing player, it does require the collection of information about one's own history of play which is not required for the unaugmented CGK model. This gathering of information in not used in our ranking the models. The inclusion of one type of information and not the other is actually not arbitrary, but instead reflects our basic approach; we are not trying to find all differences between learning models, but rather to find those which imply a greater degree of reasoning ability. We argue that information about one's opponent is more relevant for determining cognitive sophistication. A key element of game theory is that players anticipate their opponents' actions and respond accordingly. In other words, a player realizes he/she is playing a game in which his/her payoffs are not determined solely by their own actions, but by an interaction between his/her choice and the choices of others. A critical step in making this leap from solipsism is gathering some information about one's opponent. Thus, we rank rules only on the amount of information gathered about other players.

Our criteria only give a weak ordering over the set of learning rules. We are agnostic as to whether optimization or greater information requirements is more important in determining the level of cognitive sophistication. Any answer is likely to depend on the situation being considered as well as personal taste.

6. Simulation Results -- The Unaugmented Roth-Erev Model

We begin the selection process by considering the least cognitively sophisticated of our four models, the unaugmented RE model. This section shows that while the unaugmented RE model characterizes broad features of the SP treatment data, it fails to capture some of the finer details and, more importantly, does a poor job of characterizing data from the ZA treatment.

Figure 3 gives a summary of results from 50 simulations of the SP treatment with the unaugmented RE model. We report probability distributions over strategies in periods 0, 1000, 10,000, and 50,000, with entry rates given by the numbers above the M players' strategies. There is little variation in results across simulations, so these graphs accurately depict play in a typical simulation.

On a broad level, the unaugmented RE simulations resemble the experimental results from the SP treatment. In particular, the simulations exhibit the emergence of limit pricing at 6 by M_L's seen in the experimental data. The driving force is the same as in the experimental data: M_H's limit price at 3 and 4 in response to high initial entry rates on 1 and 2. The entry rates on 3 and 4 (and to a lesser extent 5) rise in response to this shift by M_H's. This increase in entry rates pushes the M_L's to limit price at 6. The simulations converge to the Riley outcome, consistent with the experimental data.

While the unaugmented RE simulations capture the broad features of the experimental data, some of the finer details are substantially different. First, the intermediate pooling stage is relatively weak. While there is substantial movement toward 4, play of the myopic maximum, 2, remains the modal choice of M_H's in period 1000, and 3 is played with virtually the same frequency as 4. This occurs even though 4 is the best response for M_H's in period 1000, given the play of E's. The unaugmented RE simulations also show large amounts of play of 7 by M_L players and 6 and 7 by M_H players, strategies which are virtually never observed in the experiments.[13]

These problems highlight one of the great weaknesses (and strengths) of the RE model -- it does not contain optimization. Strategies which are second or third best on a consistent basis can be played quite frequently, as long as they aren't too much worse than the expected payoff maximizing action. Thus, play of 3, 6, and 7 by M_H's and 7 by M_L's persists even though these strategies are rarely optimal. The CGK model, with optimization, is more sensitive to small differences in payoffs. Turning to the ZA treatment, this difference between the two models plays a critical role.

We ran 50 unaugmented RE simulations with the ZA treatment. These simulations converge to a variety of final outcomes. Of the 50 simulations, 9 converge to the Riley outcome and 20 to converge to the partial pooling equilibrium discussed in Section 2. The 21 remaining simulations do not converge to any equilibrium; rather, they appear to converge to myopic maximum play (that is, M_H's play 2 and M_L's play 4).[14] Figures 4, 5, and 6 show probability distributions of M player strategies in periods 0, 1000, 10,000, and 50,000, averaged over the simulations that converged to myopic-maximum play, the mixed-

---

[13]The RE simulations are very slow to converge to equilibrium, much slower than, for example, ultimatum game simulations. (Ultimatum game simulations typically converge to a Nash equilibrium within 300 rounds, while the limit pricing simulations take in excess of 10,000 rounds.) While this is curious, it is not in itself great grounds for concern. Changing parameter values can have dramatic impact on the speed of convergence -- the correct parameters for the ultimatum game may simply be very different then those for the limit pricing game.

[14]Because of the nature of the RE model and the specific parameters used, true convergence to myopic-maximum play (or in fact, any non-Nash equilibrium) is a zero-probability event. However, it is possible for simulations to remain close to this outcome for arbitrarily long lengths of time. We took one simulation that went to myopic maximum play and ran it for 500,000 periods, and there was no apparent movement away from myopic-maximum play. The limit pricing game has a mixed strategy equilibrium in which M_L's always choose 4, M_H's choose 4 with probability 1/5 and 2 with probability 4/5, and E's enter always on everything but 4 and with probability 20/23 on 4. It can be shown that the simulations which end with myopic maximum play will, with very high probability, eventually converge to this mixed strategy equilibrium.

strategy equilibrium, and the Riley outcome, respectively. Entry rates are given by the numbers above the M players' strategies.

It is only after period 10,000 that the ZA simulations begin to diverge significantly from each other en route to either the myopic maximum outcome, partial pooling equilibrium, or Riley outcome. These different outcomes are driven by subtle differences in the simulations which begin to emerge as early as round 1000. At this point, the simulations which eventually converge to myopic maximum play show relatively little limit pricing by M_L's, especially at 5. Even though 6 and 7 are dominated strategies for M_H's, global experimentation induces play of these actions in the unaugmented RE model. Because the payoffs to M_H's on 6 and 7 are not much lower than the payoffs from undominated strategies, play of 6 and 7 by M_H's build up to significant percentages in round 1000, almost as high as play of 5 by M_H's (5.1%, 4.5%, and 4.0% on 5, 6 and 7 respectively). The high percentages of play of 5, 6, and 7 by M_H's (compared to 11.9%, 7.9%, and 6.8% for M_L's) make entry a best response to these output levels. In simulations which converge to myopic maximum play, the relatively low amount of limit pricing leads to virtually 100% entry on 5, 6, and 7 by round 10,000. This pushes play toward the myopic maxima. In simulations which go to the partial pooling equilibrium or the Riley outcome, there is enough play of 5 and 6 by M_L players to keep the entry rates below 100% following these actions. Which equilibrium emerges depends on small differences in how frequently M_L's play these strategies, and hence on the entry rates. Thus, very small differences in entry rates which only emerge around round 10,000 are driving the equilibrium selection in the unaugmented RE model.

On a broad level, the unaugmented RE model does a mixed job of characterizing the data in the ZA treatment. It does capture the most important differences between the SP and ZA treatments. Comparing experimental data for M_L's from the ZA and SP treatments, there is less limit pricing in the ZA treatment at any point in time, and relatively more play of 5 versus 6. While play by M_L's in the first 1000 periods of the ZA simulations is similar to play in the SP simulations, this breaks down between period 1000 and period 10,000. M_L's in the SP treatment begin to limit price extensively, with 6 being the most frequently chosen strategy by period 10,000. In contrast, play of 4 is still the modal choice in period 10,000 for M_L's in the ZA treatment, with 5, 6, and 7 each being played with decreasing frequency. While it is mildly disturbing that differences in rates of limit pricing take such a long time to develop, the differences are eventually quite strong, as in the experimental data. Other aspects of the simulations are less consistent with the experimental data. It is troubling that such a high percentage of the simulations go to the myopic maximum. Nothing resembling this has ever been observed in the data, and it strains credulity to state that such a pathological outcome could actually occur.[15] Looking at the play of E's raises another disturbing point. In the experimental data, the ZA treatment yields somewhat higher entry rates of 5 and substantially higher entry rates on 6 than the SP treatment. This is borne out in the RE simulations to an extreme -- the almost 100% entry rates observed on these strategies by round 10,000 resemble nothing seen in the experiments. Finally, as in the SP treatment, we observe that the intermediate pooling stage is relatively weak with frequent play of 3 rather than 4 by M_H's.

These anomalies follow again from the lack of optimization in the RE model. Without optimization, there is no sharp switch by M_H's from 2 to 4 and there is persistent experimentation by M_H's with 6 and 7. Given that pooling by M_H's is relatively weak, M_L's are slow to limit price at 5, 6, and 7. Combined with extensive use of these strategies by M_H's, this leads to the surprisingly high entry rates on 5, 6, and 7. When these entry rates are exceptionally high, the myopic maximum emerges in the long run.

Under closer scrutiny, the unaugmented RE simulations characterize the experimental data quite poorly. This is especially clear when the features which drive differences between SP and ZA simulations are examined. The unaugmented RE simulations capture the lower rate of limit pricing in

[15]Making the global experimentation parameter die out more slowly reduces, but does not eliminate, the likelihood of simulated play becoming "stuck" at myopic maximum play.

general, and of choices of 6 in particular, in the ZA treatment (versus the SP treatment) because entry rates on 6 and 7 are higher than in the SP treatment. However, these higher entry rates are driven by M_H players' persistent experimentation with 6 and 7, something which is observed *virtually never* in the actual experiments. Because the unaugmented RE model does not include optimization, it allows dominated strategies to be played extensively. This leads to the anomalous play of 6 and 7 by M_H's which is so important in differentiating SP and ZA treatments. Thus, the unaugmented RE simulations mimic the experiments for reasons which could not possibly be driving behavior in the experiments. Instead, the change in behavior by E players on 6 and 7 in the ZA treatment must be driven by some change in their perceptions of these strategies. Experimentation with 6 and 7 makes concrete what is actually a cognitive effect.

The positive results generated by the unaugmented RE model are sensitive to the form of experimentation used. Our version of the RE model employs global experimentation; experimentation increases the probability of all available strategies. This differs from the local experimentation employed by Roth and Erev, in which experimentation only affects strategies adjacent to the one most recently employed. With local experimentation, play does not converge to the Riley outcome, instead exhibiting extensive play of 5 and 7 by M_L's even in period 50,000. Moreover, the intermediate pooling stage is even weaker if local rather than global experimentation is used. It is troublesome that the RE model's ability to characterize the limit pricing data is so sensitive to details of the specification. Since there is no real reason to believe *ex ante* that the limit pricing game ought to generate global experimentation as opposed to local experimentation, this change is difficult to justify.[16] If we could

[16]If global experimentation better characterized the data in all experiments, we would be less worried. However, this is not true. We have run ultimatum game simulations using the parameters of Roth and Erev (μ=0, ϕ=.001, initial propensities estimated from American data), along with a global experimentation parameter of .05 (they use a local experimentation parameter of .05), for 20,000 periods. Demands increase steadily over time, and by period 20,000, about two-thirds of demands are for the largest possible amount, and these are accepted over 90% of the time. This stands in marked contrast to standard ultimatum game results, in which there is little, if any, tendency for demands to increase over time.

show that the RE model was unaffected by eliminating experimentation altogether, the natural approach would be to simplify the model in this way. However, the need for some type of experimentation in the RE model is highlighted by comparing the ZA and SP treatments. Without adding experimentation to the simulations, entry rates for 6 and 7 are very low in both types of simulation, and play converges to the Riley outcome in either case. It is only with experimentation that the unaugmented RE model is able to track (even partially) differences between the SP and ZA treatment data. Our approach has been to choose the formulation which gives the RE model the best chance. As noted in Cooper *et al.* (1997), Appendix C, the augmented CGK model's results are robust to changes in how initial weights are generated, weights are updated, and the addition of various noise terms.[17]

7. Simulation Results -- The Unaugmented CGK Model

The unaugmented RE model's inability to characterize major features of the experimental data suggests the need for a more cognitively sophisticated learning model. The unaugmented CGK model uses optimization and requires information about what strategies the opposing player has used, making it more cognitively sophisticated than the unaugmented RE model on two counts. In this section we examine whether our *ex ante* and *ex post* criteria are met for moving between these models.[18]

Good reasons exist *ex ante* for considering a learning model with optimization. Dominated strategies play a central role in the limit pricing game. Eliminating dominated strategies implies that

---

[17]The RE model is designed to incorporate the power law of practice -- learning curves become flatter over time. A side effect of this is that for games where convergence is slow, such as the limit pricing game, changes in strategies due to differences in payoffs are overwhelmed by changes due to experimentation. Thus, the RE model is sensitive to how experimentation is incorporated into the model. For a detailed discussion of this issue, see Binmore and Samuelson (1995).

[18]Our material on CGK model simulations replicate results from Cooper, Garvin, and Kagel (1996, 1997, and in progress). As such, we have kept our description of the results brief. It should be noted that the CGK model was designed to characterize the limit pricing game data. Because of this, it would be surprising if it did not do a good job. What is interesting is the differences between the RE and CGK model which allow the latter to outperform the former in this case.

subjects' are optimizing at least to the extent of not using strategies which can't be optimal.[19] More generally, M's have a relatively large number of strategies which are not sharply differentiated in payoffs. To get sharp predictions about which strategy will actually be used, optimization is needed. Since subjects tend to use only a small subset of the available strategies, a model which mimics this feature is likely to do better than one which does not.[20]

Unfortunately, the unaugmented CGK model does not fulfill our *ex post* condition for moving to a more cognitively sophisticated model. Figure 7 summarizes data from 500 simulations of the SP treatment with the unaugmented CGK model.[21] Like the experiments, the simulations move at first toward an attempt to pool at 4. Unlike the experiments, where behavior proceeds to the Riley outcome, play in most (80%) of the simulations converges to the partial pooling equilibrium described in section 2. The remainder converge to the Riley outcome. The unaugmented CGK model does not capture a central feature of the SP treatment data.

8. Simulation Results -- The Augmented CGK and Augmented Roth-Erev Model

Cooper *et al.* (1997) assert that experimental subjects possess a degree of cognitive sophistication that is captured by neither unaugmented learning model. Namely, some E players are able to recognize that play of 6 or 7 by a high-cost monopolist is strictly dominated:

> "When no Es recognize the existence of dominated outcomes, there is no difference in the way Es treat 5 or 6 initially. In the early stages of play, only low cost Ms are observed playing 5 or 6, with 5 typically chosen more often than 6, due to its more favorable payoff. Due to its greater initial frequency of play, Es learn that 5 represents low cost Ms

---

[19]Yaw Nyarko shows that any learning model which does not allow play of dominated strategies can be rewritten as a Bayesian model with maximization versus some prior beliefs (personal conversation).

[20]Considering the other players' behavior is justified *ex ante*, since this is the essence of playing a game.

[21]We typically ran 500 CGK simulations in each cell, but only 50 RE simulations. This should not affect any of our results, and merely reflects the different run times of the two types of simulation; running 500 CGK simulations takes less time than running a single RE simulation.

faster than they learn that 6 does. The net result is that Es play OUT on 5 earlier than on 6, which reinforces M_Ls natural bias towards 5. As a result, in many simulations 6 ceases to be played before Es learn that only low cost Ms ever choose it."[22]

This is a textbook case for moving to a more cognitively sophisticated model. The limit pricing game evokes a cognitive element which cannot be captured by the unaugmented CGK model, and the unaugmented CGK model is unable to fully characterize the experimental results. Our *ex ante* criterion for moving up the hierarchy of cognitive sophistication to the augmented CGK model is met. We now show that the *ex post* criterion is also met.

Figures 8 and 9 summarize 500 simulations of the augmented CGK model with $\pi = .5$ and $\pi = 1.0$ respectively for the SP treatment. Recall that the unaugmented CGK model is identical to the augmented CGK model with $\pi = 0$. The results of the simulations show a strong dependence on π, not only in the amount of limit pricing by M_L's, but also the speed with which it emerges and the form which it takes. As π increases, limit pricing is more prevalent and occurs sooner. Furthermore, this limit pricing is more often accomplished by choosing 6 (as opposed to 5). This last difference has a strong effect on which outcome is seen when convergence occurs. As previously mentioned, when $\pi = 0$, only 20% of the simulations converge to the Riley outcome. However, when $\pi = 0.5$, 65% of the simulations converge to the Riley outcome, and when $\pi = 1.0$, 89% of the simulations converge to the Riley outcome (the rest of the simulations converge to the partial pooling equilibrium). As in the $\pi = 0$ case, simulated players in the $\pi = 0.5$ and $\pi = 1$ cases attempt to pool at 4 before eventually converging to equilibrium.

In general, the augmented CGK model with $\pi = 0.5$ does a good job of characterizing the main patterns in the SP treatment experimental data. Play starts at the myopic maxima, followed by attempts to pool at 4, and then eventual convergence to the Riley outcome. These simulations also match up well with the details of the experimental data highlighted in Section 4 -- M_H's attempt to limit price at 4 much

[22] Cooper *et al.* (1997) p. 14

more frequently than at 3, and there is no play of either 6 or 7 by M_H's.

The experimental results from the ZA sessions are also consistent with predictions of the augmented CGK model. Cooper *et al.* hypothesize that since the proportion of subjects with limited strategic foresight is likely to be lower in the ZA treatment than in the SP treatment, differences in the experimental data between these treatments should be similar to the differences between the π=0 CGK simulations and the π=0.5 CGK simulations. The differences between the ZA experimental results and the SP experimental results are indeed analogous to the difference between the π=0 CGK simulations and the π=0.5 CGK simulations; as π decreases from 0.5 to 0, the likelihood of limit pricing by M_L players in the form of choices of 5 rather than 6 increases, and this limit pricing takes longer to emerge. We therefore conclude that the augmented CGK model does a good job of characterizing data in both the SP and ZA treatments, and captures the differences between these two treatments. Both our *ex ante* and *ex post* criteria for moving from the unaugmented CGK model to the augmented CGK model are met.

Before we can conclude that the augmented CGK model is best suited to the limit pricing data, the augmented RE model must be considered. Like the unaugmented CGK model, the augmented RE model is more cognitively sophisticated than the unaugmented RE model and less cognitively sophisticated than the augmented CGK model. If the augmented RE model is able to characterize the major features of the data, the *ex post* criterion for moving to the augmented CGK model is not met. Such a result would suggest that only limited strategic foresight, not optimization, is needed to capture subjects' behavior in the limit pricing game.

Note that augmenting the RE model implies a fundamental change to the model. When we refer to limited strategic foresight, this is shorthand for a player performing two iterations of removal of dominated strategies. Not only does the player not use any of his/her own strategies which are dominated, but he/she also anticipates that the opposing player will never use a dominated strategy. The second part of this statement implies that the player knows something about how the opposing player is making choices. While this may not seem significant, it is actually a quite dramatic leap in the cognitive sophistication of players. One of the most striking features of the CGK data is the inability of players to anticipate their opponent's play. Numerous subjects never choose 6 or 7 when they are high cost M's, and yet enter on play of 6 or 7 (recall that in the CGK design, players are randomly assigned roles in every round, and so acquire experience as M_H, M_L, and E players). Along similar lines, play by M's is initially clustered around the myopic maxima, even though entry rates are far lower on 4 than 2 from the beginning of the experiments. This implies that the same players who realize as E's that they should enter on 2 and not enter on 4 fail to anticipate this behavior when they are M_H's.

Because of this striking increase in the level of cognitive sophistication, augmenting the RE model with limited strategic foresight yields a model which is inconsistent with the spirit of the original. The nicest feature of the RE model and of reinforcement-based models in general is how little is required of players. A player need not know anything about his or her opponent to act in a manner consistent with the RE model; he or she need not even realize that a game is being played. This nice feature vanishes with the addition of limited strategic foresight.

Cooper *et al.* (1996) model the differences between the SP and ZA treatments as a change in π, the percentage of players with limited strategic foresight in the augmented CGK model. They run sets of simulations with $\pi = .5$ and $\pi = 0$, postulating that differences between these sets of simulations should mirror differences in the experimental data from the SP and ZA treatments. This hypothesis is borne out by the simulation results. If the augmented RE characterizes the Cooper *et al.* data, we should be able to perform a similar exercise; differences between the augmented RE model with $\tau = .5$ and $\tau = 0$ should parallel differences between the SP and ZA treatments.

As reported in Section 6, the unaugmented RE model does a reasonably good job of tracking SP data. In considering the augmented RE model, we first confirm that augmentation does not affect the RE model's positive performance with the SP treatment. We ran 50 simulations of the SP treatment using the augmented RE model with $\tau = .5$. These simulations are identical to the unaugmented RE simulations, except for the restrictions that no M_H can use 6 or 7 and half of the E's never enter on 6 or 7. Average outcomes from these simulations are summarized in Figure 10. Comparing with Figure 3, which summarizes simulations of the SP treatment using the unaugmented RE model, we see that convergence is faster with augmentation, but general patterns of play are unaffected.

Given that the augmented RE model does a good job of tracking the experimental data in the SP treatment, we turn to the ZA treatment. We ran 50 simulations of the ZA treatment using the augmented RE model with $\tau = 0$. These simulations only differ from the unaugmented RE simulations by restricting M_H's to never play 6 or 7. Average results for these simulations are summarized in Figure 11. Augmenting the RE model strongly affects simulation results for the ZA treatment. The wide variety of outcomes observed with the unaugmented RE model vanishes with augmentation. The only outcome observed in the augmented RE simulations of the ZA treatment is the Riley outcome -- no simulations converge to either the myopic maxima or the mixed strategy equilibria. This by itself is not cause to reject the augmented RE model, as we cannot argue that the long run outcome is observed in the Cooper *et al.* experiments. Even though the ZA data show much more play of 5 than 6, if play had continued indefinitely, the Riley outcome might have emerged. However, the augmented RE simulations do a poor job of capturing differences between the ZA and SP treatment data. This is seen most easily by comparing Figure 7 and Figure 8. By period 1000, these figures are virtually identical, and remain so throughout the remainder of the periods. In the augmented RE simulations, limit pricing does not emerge more slowly in the ZA treatment than in the SP treatment, and the relative frequency of limit pricing at 5 versus 6 is not higher in the ZA treatment. Given that these are the most striking features of the experimental data, we conclude that the augmented RE simulations fail to track the data.

It is not surprising that augmenting the RE model does little to improve tracking of the experimental data. The only way augmentation could help the RE model is if it improved performance in the ZA treatment. However, the reasons for the RE model's poor performance in the ZA treatment have little to do with limited strategic foresight -- increased play of 6 and 7 by M_H's in the ZA simulations *vs.* the SP simulations proxy for the presumably greater difficulty of recognizing dominated strategies. The unaugmented RE model cannot characterize the ZA data because it lacks optimization. Augmentation will not solve this problem.[23]

Thus, both the *ex ante* and *ex post* criteria for moving from the unaugmented RE model to the more cognitively sophisticated augmented CGK model are met. The augmented CGK model's superior ability to characterize the limit pricing data is driven by its greater cognitive sophistication. Optimization strengthens attempts to pool by M_H's and eliminates persistent play of dominated strategies. Limited strategic foresight generates differences between ZA and SP treatment simulations of the augmented CGK model in a manner more consistent with the experimental data than the combination of experimentation and non-optimization which drives differences between ZA and SP treatment simulations of the unaugmented RE model.

9. Conclusion

The goal of this paper is to formulate and apply a rule for mapping between experimental data

[23]By augmenting the RE model, we remove the source of difference between the SP and ZA treatments -- M_H players are restricted to never choose 6 and 7. Given that we have removed the one source of difference, it follows naturally that the results for the two treatments look virtually identical.

It can be argued that the RE model has been augmented too much. In particular, the RE model is augmented so that all M_H's avoid play of 6 or 7. For the ZA treatment, where it is less obvious that these are dominated strategies, we might wish to have only some M_H's eliminate play of 6 and 7. Note that such a modification has no effect on the cognitive sophistication of the model as defined in Section 5 -- no knowledge of one's opponent is required to eliminate one's own dominated strategies. Thus, we can formulate an alternative augmented RE model in which some proportion γ of M_H players never use 6 or 7. The model is otherwise identical to the unaugmented RE model. Extensive simulations show that this alternative augmentation adds little to the RE model. For $\gamma > .25$, the simulation results are nearly identical to results with the augmented RE model. For $0 < \gamma \leq .25$, we see extensive play of 5 by M_L's and the emergence of the mixed strategy equilibrium. Unfortunately, we also see all of the flaws observed in unaugmented RE simulations of the ZA treatment. In particular, we observe the myopic maximum as a long run outcome in a high frequency of trials (13 of 50 with $\gamma = .25$) and see significant departures from the experimental data at the individual level (play of 6 and 7 by M_H players, too much play of 3, little attempt to pool at 4 by M_H players , etc.).

and learning rules. Selection takes place over a hierarchy of learning rules, with rules higher on the hierarchy possessing a greater degree of cognitive sophistication. We select the model with the least possible cognitive sophistication. Movement up the hierarchy of cognitive sophistication is conditional on two criteria. First, *ex ante* the game should have some feature which demands a greater degree of cognitive sophistication for analysis. *Ex post*, a higher degree of cognitive sophistication should be necessary to characterize the experimental data.

The bulk of the paper applies this approach to a particular experimental data set, the CGK limit pricing data. We establish that both the *ex ante* and *ex post* criteria hold for moving up the hierarchy of cognitive sophistication from the least sophisticated model, the unaugmented RE model, to the most sophisticated model, the augmented CGK model. This indicates that both optimization and limited strategic foresight are necessary to characterize the limit pricing data.

Our purpose is not to conduct a horse race between reinforcement-based models and belief-based models, but rather to present a method for systematically determining which learning model best organizes a data set. Simple reinforcement-based models work wonderfully for many data sets, and in such cases there is no reason to move up to a more cognitively sophisticated model. The limit pricing game is not one of these situations. Even picking the most favorable specification of the RE model, it lacks the cognitive sophistication needed to capture major elements of the limit pricing data.

The augmented CGK model hardly exhausts the limits of cognitive sophistication in learning models. However, given that the augmented CGK model does an excellent job of characterizing the experimental data, it is hard to see the *ex post* criterion for moving to a more cognitively sophisticated model being satisfied.

While we have concentrated on the limit pricing game, our results suggest that certain types of games will require greater degrees of cognitive sophistication. Games which are very simple or have sharp differences in payoffs are natural candidates for reinforcement based learning. Games in which forward induction and similar reasoning play an important role and games in which differences in payoffs are less extreme are likely to require the greater cognitive sophistication embodied by a belief-based learning rule.

This paper is fairly limited in its goals. Our point is simple. It has already been established that learning models have an important role to play in the interpretation of experimental data. What is needed is some way to sort among the available learning models. Instead of engaging in an unproductive search for the canonical learning model, we feel it is far more important to think about how the most appropriate learning model for describing an experimental data set should be selected. Our primary contribution is to present a systematic approach to this problem.

BIBLIOGRAPHY

Binmore, K. and L. Samuelson (1995), "Evolutionary Drift and Equilibrium Selection," University of Wisconsin, Madison, mimeo.

Boylan, R. and M. El-Gamal (1993), "Fictitious Play: A Statistical Study of Multiple Economic Experiments," Games and Economic Behavior, 5, 205-222.

Brandts, J. and C. Holt (1994), "Naive Bayesian Learning and Adjustment to Equilibrium in Signaling Games," forthcoming Journal of Economic Behavior and Organization.

Camerer, C. and T. Ho (1996), "Experience-Weighted Attraction Learning in Games: A Unifying Approach," California Institute of Technology, mimeo.

Cheung, Y. and D. Friedman (1997), "Learning in Evolutionary Games: Some Laboratory Results," Games and Economic Behavior, 19, 46-76.

Cho, I. and D. Kreps. "Signaling Games and Stable Equilibria," Quarterly Journal of Economics 102, 1987, 179-221.

Cooper, D., S. Garvin, and J. Kagel (1996). "Adaptive Learning vs. Equilibrium Refinements in an Entry Limit Pricing Game," forthcoming The Economic Journal.

Cooper, D., S. Garvin, and J. Kagel (1997). "Signaling and Adaptive Learning in An Entry Limit Pricing Game," forthcoming Rand Journal of Economics.

Cooper, D., S. Garvin, and J. Kagel, "The Effects of Context on Learning in an Entry Limit Pricing Game," in preparation.

Cooper, D. (1995), "Reinforcement-Based Learning *vs.* Bayesian Learning: Technical Notes," University of Pittsburgh, mimeo.

Cooper, R., D. DeJong, R. Forsythe, and T. Ross (1990), "Selection Criteria in Coordination Games: Some Experimental Results," American Economic Review, 80,1, 218-233.

Crawford, V (1991), "An Evolutionary Interpretation of Van Huyck, Battalio, and Beil's Experimental Results on Coordination," Games and Economic Behavior, 3, 25-59.

Duffy, J. and N. Feltovich (1996), "The Effect of Information on Learning in Strategic Environments: An Experimental Study," mimeo.

Erev, I. and A. Roth (1996), "On the Need for Low Rationality, Cognitive Game Theory: Reinforcement Learning in Experimental Games with Unique Mixed Strategy Equilibria," University of Pittsburgh, mimeo.

Fudenberg, D. and David Levine (1996), Theory of Learning in Games, forthcoming MIT Press: Cambridge.

Gale, J., K. Binmore, and L. Samuelson (1995). "Learning to be Imperfect: The Ultimatum Game," Games and Economic Behavior, 8, 56-90.

Holt, D. (1993), "An Empirical Model of Strategic Choice with an Application to Coordination Games," Queen's University, mimeo.

McKelvey, R. and T. Palfrey (1994), "Quantal Response Equilibrium in Extensive Form Games," California Institute of Technology, mimeo.

McKelvey, R. and T. Palfrey (1995), "Quantal Response Equilibria for Normal Form Games," Games and Economic Behavior, 10, 6-38.

Milgrom, P. and J. Roberts (1982), "Limit Pricing and Entry Under Incomplete Information: An Equilibrium Analysis," Econometrica 50, 443-459.

Miller, J. and J. Andreoni (1991), "Can Evolutionary Dynamics Explain Free Riding in Experiments?" Economic Letters, 36, 9-15.

Mookherjee, D. and B. Sopher (1997), "Learning and Decision Costs in Experimental Constant Sum Games," Games and Economic Behavior, 19, 97-132.

Riley, John, "Informational Equilibrium," Econometrica 48, 1979, 331-59.

Roth, A. and I. Erev (1995). "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," Games and Economic Behavior 8, 164-212.

Roth, A., I. Erev, B. Slonim, and Y. Bereby-Meyer (1997), "On the Boundaries of Bounded Reinforcement Learning Models," mimeo, University of Pittsburgh.

Siegel, S. and Castellan, N.J. (1988), Nonparametric Statistics for the Behavioral Sciences, New York, McGraw-Hill.

Table 1:
Monopolist Payoffs (Standard Treatment)

High Cost Monopolist		
Monopolist Action	Entrant Response	
	IN	OUT
1	150	426
2	168	444
3	150	426
4	132	408
5	56	182
6	-188 *(38)*	-38 *(162)*
7	-292 *(20)*	-126 *(144)*

Low Cost Monopolist		
Monopolist Action	Entrant Response	
	IN	OUT
1	250	542
2	276	568
3	330	606
4	352	628
5	334	610
6	316	592
7	213	486

Note: Italicized numbers represent changes in payoffs made for the ZA treatment.

Table 2:
Entrant Payoffs

Entrant's Strategy	Monopolist's Type	
	High Cost	Low Cost
IN	500	200
OUT	250	250

Figure 1a: Experimental Data -- Cooper *et al.* SP Sessions (Inexperienced Subjects)

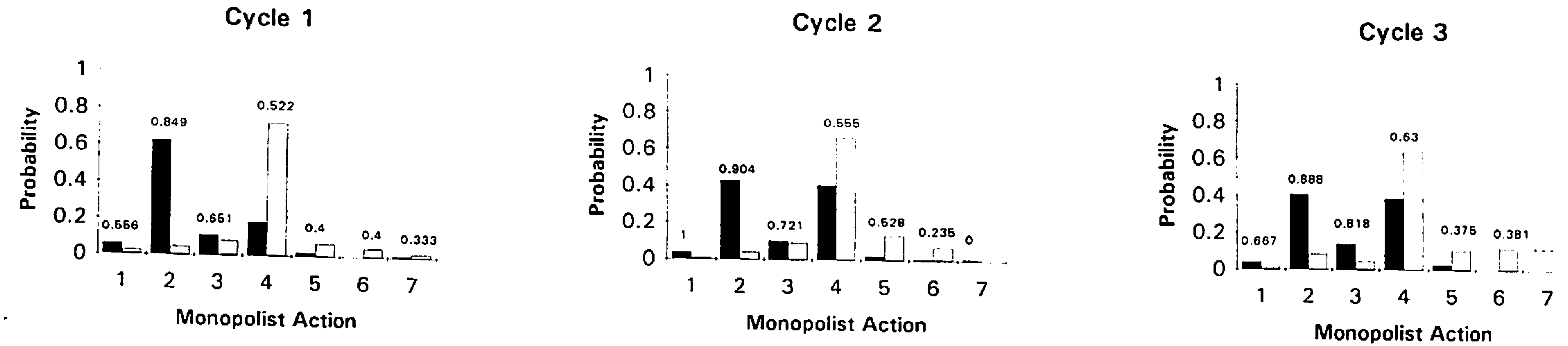
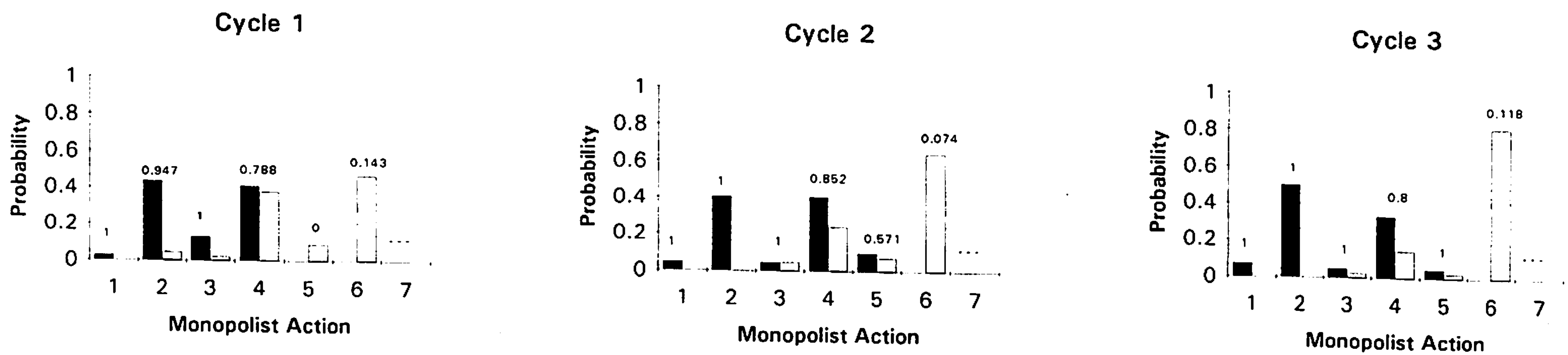


Figure 1b: Experimental Data -- Cooper *et al.* SP Sessions (Experienced Subjects)



■ High-Cost Monopolist □ Low-Cost Monopolist Entrant Probability of IN

Figure 2a: Experimental Data -- Cooper *et al.*
ZA Sessions (Inexperienced Subjects)

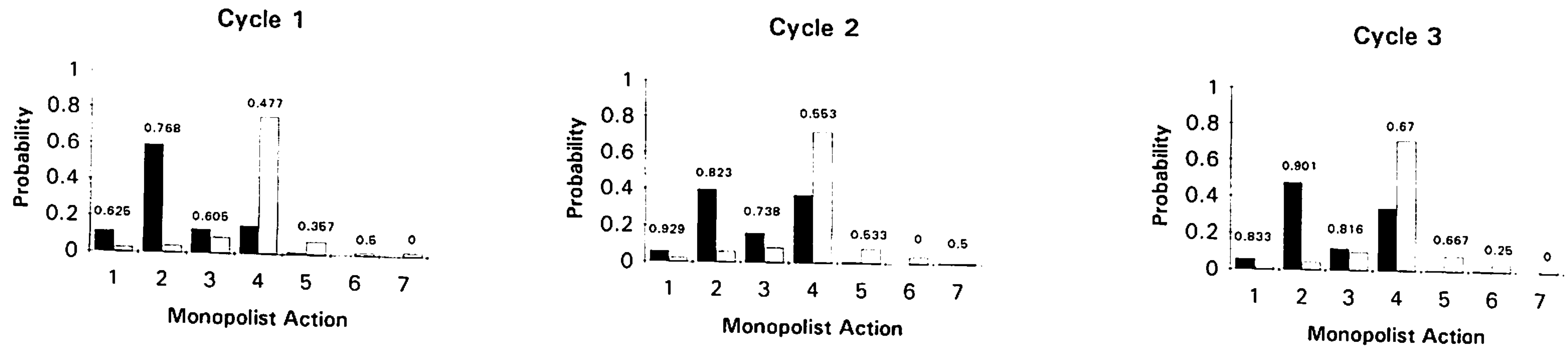
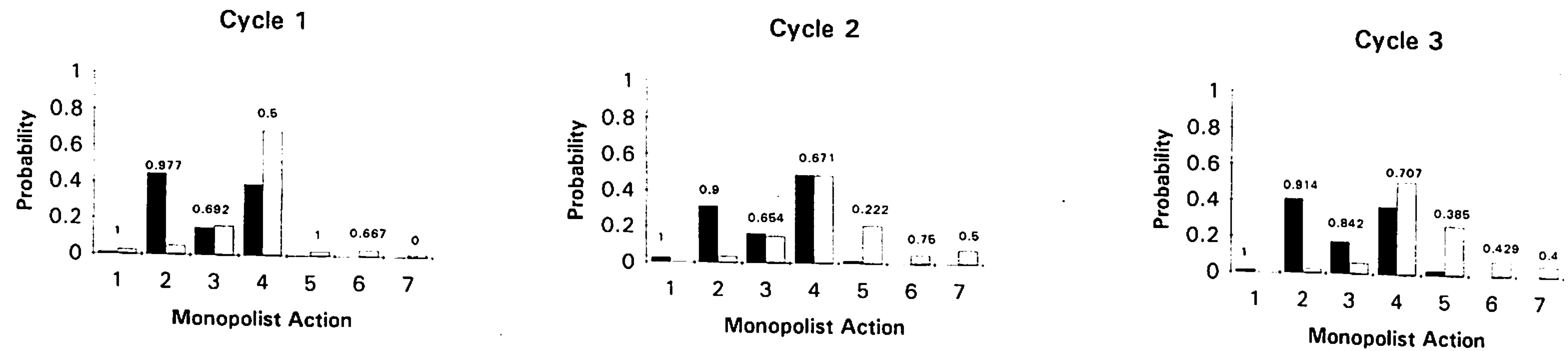


Figure 2b: Experimental Data -- Cooper *et al.*
ZA Sessions (Experienced Subjects)



■ High-Cost Monopolist □ Low-Cost Monopolist Entrant Probability of IN

Figure 3: Simulation Data -- Unaugmented CGK Model

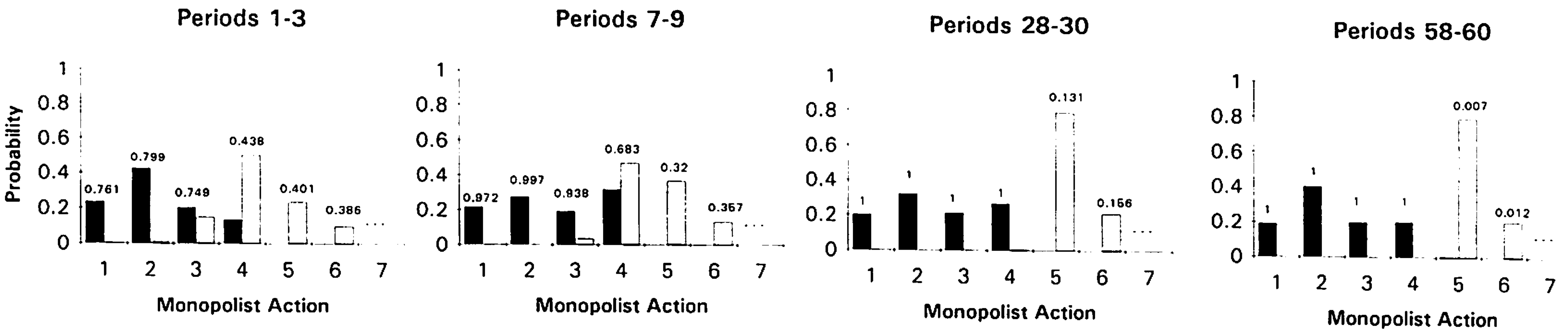
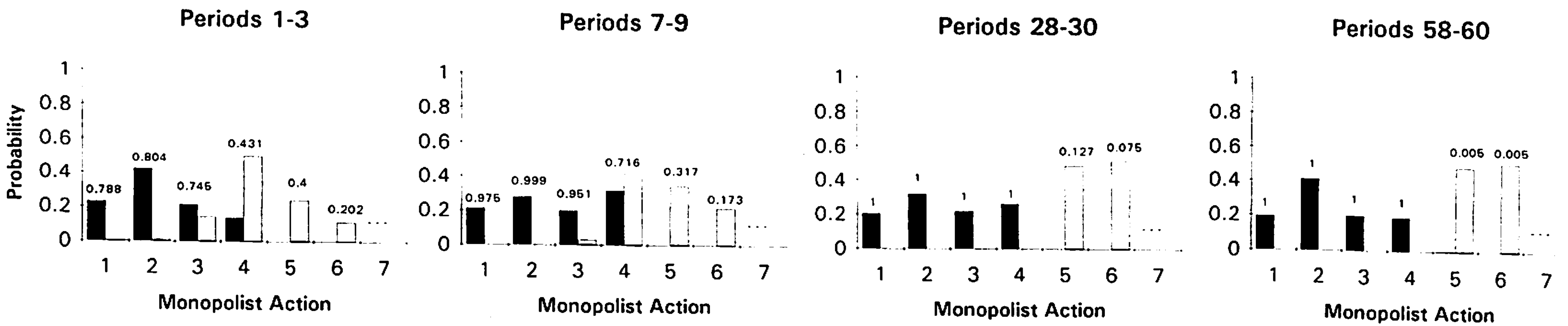


Figure 4: Simulation Data -- Augmented ($\pi = 0.5$) CGK Model



■ High-Cost Monopolist □ Low-Cost Monopolist ··· Entrant Probability of IN

Figure 5: Simulation Data -- Augmented ($\pi = 1.0$) CGK Model

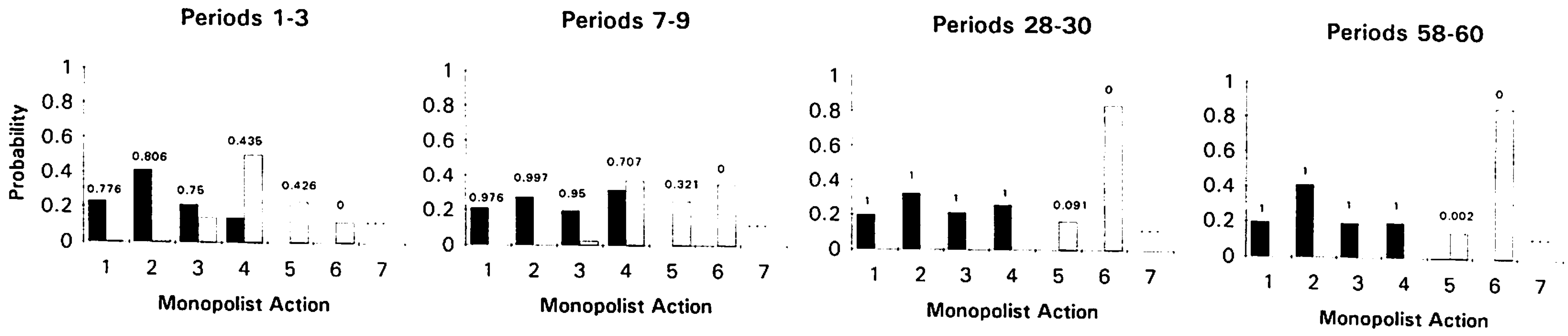
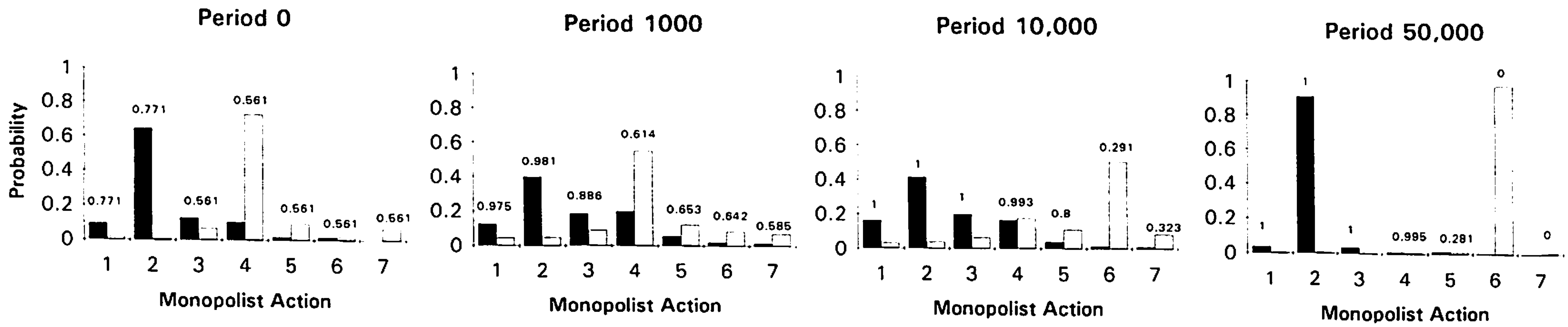


Figure 6: Simulation Data -- RE Model (SP)



■ High-Cost Monopolist ☐ Low-Cost Monopolist ... Entrant Probability of IN

Figure 7: Simulation Data -- RE Model (ZA, Converging to Myopic Maximum)

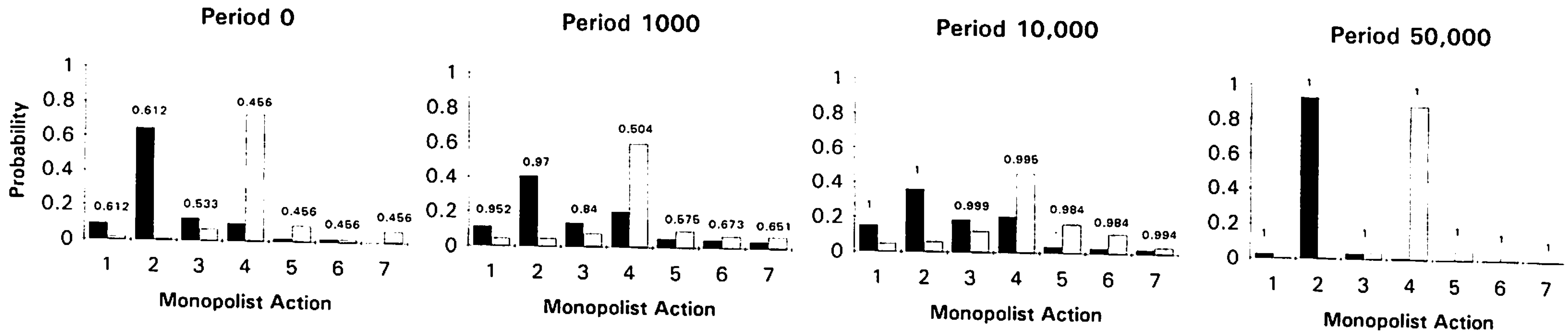
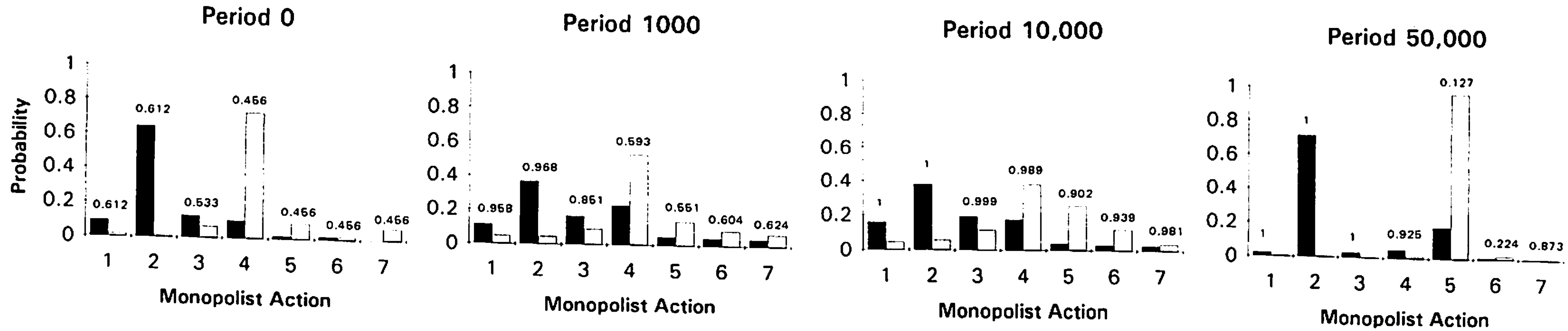


Figure 8: Simulation Data -- RE Model (ZA, Converging to Partial Pooling)



■ High-Cost Monopolist □ Low-Cost Monopolist Entrant Probability of IN

Figure 9: Simulation Data -- RE Model (ZA, Converging to Riley Outcome)

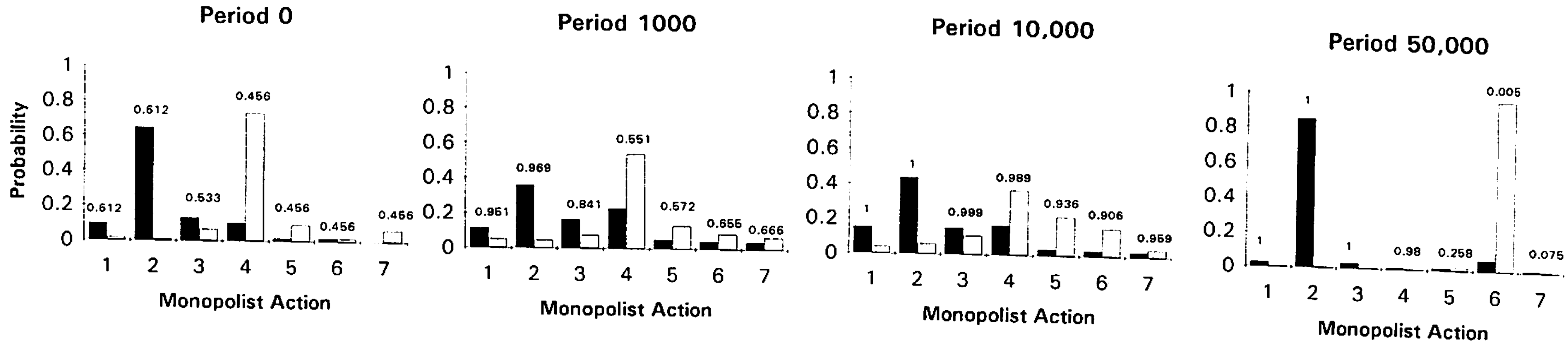
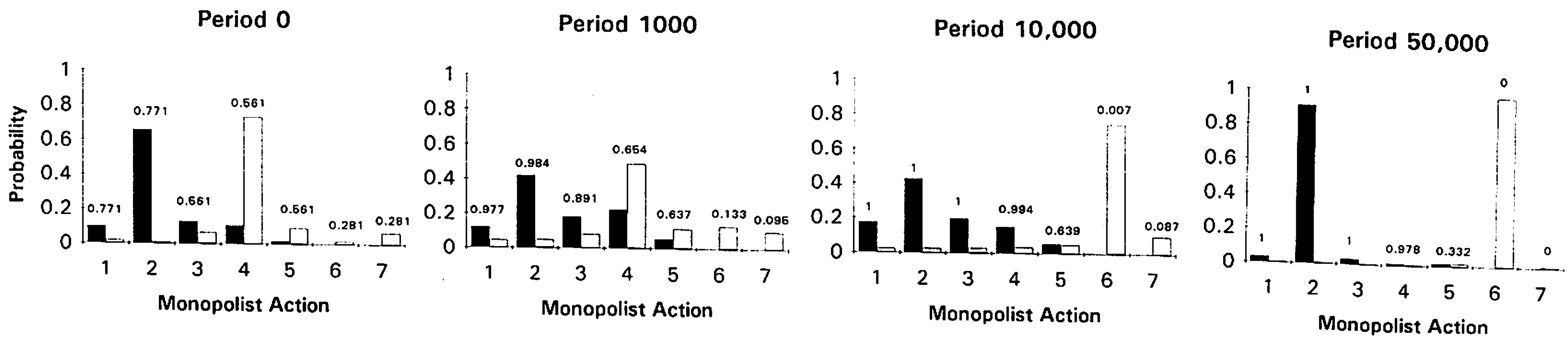


Figure 10: Simulation Data -- Augmented RE Model (SP)



■ High-Cost Monopolist □ Low-Cost Monopolist .771 Entrant Probability of IN

Figure 11: Simulation Data -- Augmented RE Model (ZA)