Reprinted from
INTERNATIONAL ECONOMIC REVIEW
Vol. 3, No. 1, January 1962

EXPERIMENTAL ANALYSIS OF A DUOPOLY SITUATION FROM THE STANDPOINT OF MATHEMATICAL LEARNING THEORY*

BY PATRICK SUPPES AND J. MERRILL CARLSMITH[1]

1. INTRODUCTION

WE ASSUME that readers of this journal are familiar with the standard theories of duopoly of Cournot, Edgeworth, and more recent writers. We propose, with certain reservations discussed below, a new theory of duopoly which is derived from recent work in mathematical learning theory. It is very likely the case that economists will not take the theory too seriously as an *economic* theory. This we are prepared to accept. The evidence for the theory comes from a highly structured, highly simplified experimental situation. On the other hand, we do maintain that it is a serious theory of behavior, and in so far as the experimental situation bears resemblance to a real economic situation, the theory is directly relevant to economic behavior. The most important feature of our theory is to provide a specific dynamic mechanism for duopolistic behavior. Such a specific mechanism is notably lacking in the theories of Chamberlain and Mrs. Robinson and is only partly provided in recent game-theoretic approaches like that of Shubik [8]. Moreover, the mechanism given here is formulated in terms of the psychological notions of stimulus, response and reinforcement, and the psychological processes of stimulus sampling and conditioning. These concepts yield a theory of oligopoly that has a considerably different flavor from that of the theories based on concepts of maximization and rational behavior.

In the next section we discuss the stimulus sampling formulation of mathematical learning theory employed in this paper and also consider a simple choice situation to which the theory has often been applied. In the third section, we briefly describe the specific duopoly situation with which we deal and then apply stimulus sampling theory in detail. The fourth section discusses the experimental setup. The

* Manuscript received February 28, 1961.

[1] This research was supported by the Rockefeller Foundation and the Group Psychology Branch of the Office of Naval Research. Discussions with Kenneth J. Arrow, Leonid Hurwicz, Paul A. Samuelson, and Hirofumi Uzawa have been most helpful. The paper will appear in the reprint series of the Stanford Institute for Mathematical Studies in the Social Sciences.

final section consists primarily of an analysis of the experimental results.

2. STIMULUS SAMPLING THEORY

Learning theory based on the notion of stimulus sampling was first formulated as a quantitative theory in a basic paper by Estes [3]. A more detailed formulation was given subsequently in several papers by Estes and Burke [4], [2]. The version to be considered here is most closely connected to that given in Suppes and Atkinson [10]. Before giving an exact verbal statement of the assumptions or axioms from which we derive our results, it will be appropriate to give an intuitive description of stimulus sampling theory.

We think of the individual's making a sequence of responses or decisions. In a simple experiment it might be a response which consists of pressing one of two keys in order to predict which one of two lights will flash. In an economic context a response might be the quarterly decision fixing the prices of all commodities produced by a firm. Stimulus sampling theory postulates that these responses are controlled by a set of stimuli in the environment from which the organism samples subsets. Each stimulus in this set is connected ("conditioned") to a possible response the organism may make. The experimenter typically requires that an overt response be chosen from a finite set of possible responses at fixed intervals. In this case, the probability that the i-th response in the set is made by the organism is given by the proportion of stimuli, sampled in a short interval prior to that response, which are conditioned to that response. In the simple two-choice key experiment it may be realistic to assume that there is but one stimulus, say the signal light which announces the beginning of a new trial. The stimuli relevant to the quarterly pricing decision could be very heterogeneous: the quarterly gross profit of the firm, the quarterly net profit of the firm, pricing policies of competitors during the past quarter, etc.

After some particular response is made, an outcome occurs which is interpreted by the organism as reinforcing one of the possible responses. In the simple two-choice experiment it is reasonable that the outcome of the left light flashing reinforces the left response with probability one. The objective outcomes which follow in the wake of a quarterly decision on prices have no such simple relations as reinforcers of possible responses, that is, possible pricing decisions for the next quarter. In the experimental situation we consider in this paper, there are varying amounts of reward, representing units of gross profit, as outcomes. Let $O_{j,n}$ be the j-th outcome on trial n, that is, the amount of gross profit received on trial n; let $A_{i,n}$ be the i-th response, $i = 1, 2$, on trial n; and let $E_{k,n}$ be the k-th reinforcing event, $k = 1, 2$, on trial n. Then

$$c_{ijk} = P(E_{k,n} | O_{j,n} A_{i,n}) \ . \tag{1}$$

Put into words, c_{ijk} is the conditional probability that profit O_j and response A_i on trial n will reinforce response A_k. Note that $E_{1,n}$ reinforces response $A_{1,n}$, and $E_{2,n}$ reinforces response $A_{2,n}$. The notation of (1), and the concepts behind it, have been discussed extensively in Estes and Suppes [5], [6] and in Suppes and Atkinson [10]. For simplicity of notation, since we have only two responses, we define

$$c_i(j) = P(E_{i,n} | O_{j,n} A_{i,n}) \ . \tag{2}$$

It should be noted that the set of stimuli available for sampling may vary systematically from trial to trial, in which case a discrimination learning problem is introduced. A simple example is the two-choice key pressing experiment with the following modification. Suppose that instead of a single signal light we introduce two lights, one red and one green, and arrange the outcomes so that the left side is always correct following the green signal light and the right side is always correct following the red light. Then the two stimuli, red and green lights, will become conditioned to their respectively correct responses. In this case, of course, the sampling of some of the stimuli is not under the organism's control. Similarly, in the quarterly pricing context, the prices set by other firms may well serve as stimuli between which it is important to discriminate. In the duopoly experiment reported here, one of the two groups of subjects was given information about the prices set by their competitors on each trial, and the other group was not.

If these preliminary remarks are borne in mind, the axioms of stimulus sampling theory, which we now give, should be easily understood. It must be realized, of course, that understanding of the axioms does not necessarily entail understanding how to apply them in detail to any behavioral situation, for extensive empirical study may be required to identify the stimuli, outcomes, and responses which are important for predicting behavior in the situation.

The first group of axioms deals with the conditioning of stimuli, the second group with the sampling of stimuli, and the third with responses. For reasons of simplicity we have restricted ourselves to the assumption that exactly one stimulus is sampled on every trial and, in the actual application of the axioms, to the still stronger assumption that exactly one stimulus is available for sampling. However, these restrictive assumptions are not an essential part of the theory.

Conditioning Axioms

C1. *On every trial each stimulus is conditioned to exactly one response.*

C2. *If* $A_{i,n}$ *and* $O_{j,n}$ *are the response and outcome respectively on trial* n, *then a sampled stimulus becomes conditioned to response* A_i *with probability* $c_i(j)$.

C3. *Stimuli which are not sampled on a trial do not change their conditioning on that trial.*

C4. *The probability* $c_i(j)$ *that a sampled stimulus is conditioned to response* A_i *is independent of the trial number and the occurrence of preceding events.*

Sampling Axioms

S1. *Exactly one stimulus is sampled on each trial.*

S2. *Given the set of stimuli available for sampling on a given trial, the probability of sampling a particular element is independent of the trial number and the occurrence of preceding events.*

Response Axiom

R1. *On each trial that response is made to which the sampled stimulus is conditioned.*

It may be remarked that Axioms C4 and S2 are "independence of path" assumptions which are necessary to prove the theorem that when exactly one stimulus is available the sequence of response random variables $A_1, A_2, \cdots, A_n, \cdots$ is a Markov chain. Of course, in the duopoly situation it is the sequence of *pairs* of response random variables which forms the chain. Detailed proof of the theorem may be found in Estes and Suppes [6]. The developments of the next section in terms of Markov chains are all based on the theorem. It may also be noted that an exact deductive development of the particular Markov processes we consider may be given from these axioms, although detailed proofs will not be presented here.

For those subjects who were told the price set by their competitor at the end of each trial, it is natural to identify this price as the most important stimulus for the next trial. In this case the stimulus available for sampling may change from trial to trial, and the experimental situation may be analyzed as a discrimination problem, involving learning to make a differential response to different stimuli. Once the stimuli are identified as the prices set by the competitor, the derivation of the appropriate Markov process is a straightforward matter. Unfortunately, it is also very tedious, for the process has 64 states. For this reason we do not analyze this discrimination setup in greater detail. Fortunately, in the present experiment the evidence indicates that there was little difference between the pairs of subjects who were told their competitors' prices and those who were not; detailed analysis relevant to this point is given in the final section.

3. APPLICATION TO DUOPOLY

The simplified duopoly situation described to the subjects is as follows. There are two firms which completely control the output of a given commodity. Their standards of quality are identical. Each firm is able to produce 6 units per quarter. Production costs are assumed constant, independent of the number of items produced. It is assumed that the market consists of eight consumers who are willing to pay different amounts for the commodity.[2] The firms are restricted to two prices, "high" and "low," which we often abbreviate H and L. (Arbitrarily we set $H = 6$ and $L = 4$.) Six of the consumers are willing to pay the high price and all eight are willing to pay the low price. When both firms set the same price, the available consumers (six or eight as the case may be) are distributed randomly between them. Thus if both firms set the high price, the distribution of consumers is binomial with $n = 6$, $p = 1/2$. If both set the low price, $n = 8$ and $p = 1/2$.[3] If one firm sets the high price and the other the low price, then the latter sells six items and the former sells 0, 1 or 2 with the appropriate hypergeometric distribution, that is, with probability 1/28, 12/28, or 15/28 respectively. In terms of expected gross profits, the situation may be represented by the following 2×2 matrix for a two-person, nonzero sum game:

(3)

	H	L
H	(18, 18)	(9, 24)
L	(24, 9)	(16, 16)

The outcome is (18, 18) when both firms set the monopoly price of H.

[2] The description of the market given to subjects, which is described in the next section, differs on this point, but not in such a way as to affect the actual market transactions.

[3] For mathematical simplicity, we permitted each firm the possibility of selling 8 units when both set the low price. However, this event occurs with low probability and no subject noted the discrepancy between this and the production quota of 6 units. The implications of a different market are worked out in Suppes and Atkinson [10].

On the other hand, the equilibrium point in the sense of Nash is the pair of strategies (L, L), yielding the payoff (16, 16). The game defined by (3) is a form of the prisoner's dilemma, which is extensively discussed in Luce and Raiffa [7].

The learning theory analysis proceeds along entirely different lines, as should be clear from the preceding section. The four states of the Markov process are the ordered pairs (L, L), (L, H), (H, L), (H, H). The interpretation of being in the state (L, L), for instance, is that each subject's single stimulus is conditioned to response L. From considerations of symmetry it is possible to combine the states (L, H) and (H, L) to obtain a three-state process, but then the markedly different predictions of the transitions from (L, H) to (L, H), as opposed to those from (L, H) to (H, L) are obliterated, and similarly for transi-

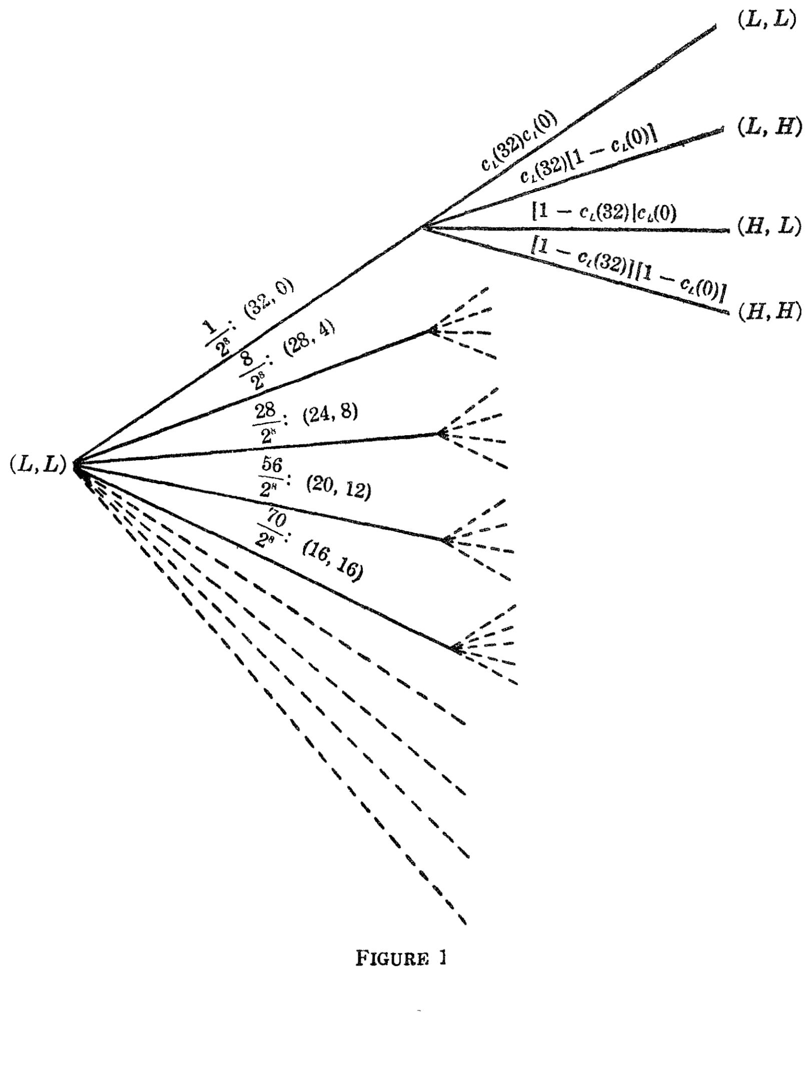

FIGURE 1

tions from the state (H, L).

From what has been said so far, it is easy to construct the four trees that show the possible transitions from trial n to trial $n+1$, one tree for each of the four possible states. A typical part of the tree for (L, L) is shown in Figure 1. Because of the large number of branches, the tree has not been completed. If, for example, one firm (i.e., subject) receives 32 units of gross profit and the other receives 0 when both have set the low price, then there are four conditioning possibilities, indicated by the four branches following the outcome $(32, 0)$. Thus $c_L(32)c_L(0)$ gives the probability that the single stimulus element of each subject remains conditioned to response L when the outcome is $(32, 0)$. If no functional relationship is assumed between the $c_i(j)$, the 4×4 transition matrix will have 18 free parameters to be estimated. This means that identifiability of the individual conditioning parameters $c_i(j)$ is not possible from the transition data for the 4×4 matrix. Also, it is neither reasonable nor interesting to assume no constraints on the $c_i(j)$. We shall, in fact, adopt the postulate that for each $i = H, L$, the $c_i(j)$ have a linear relationship,[4] which we define as follows:

$$c_L(j) = a + \frac{bj}{4}\,, \qquad j = 0, 4, 8, \cdots, 32\,, \tag{5}$$

and

$$c_H(j) = \alpha + \frac{\beta j}{6}\,, \qquad j = 0, 6, 12, \cdots, 36\,. \tag{6}$$

On the basis of (5) and (6), the number of free parameters in the transition matrix is reduced to 4, and the matrix assumes a rather simple form in terms of these four parameters. Before writing down the complete matrix, we compute some of the transition probabilities to indicate the methods. For this purpose we simplify (5) and (6) by letting $i = j/4$ in (5) and $i = j/6$ in (6). We then have, for example,

$$\begin{aligned} P((L, L) \mid (L, L)) &= \frac{1}{2^8} \sum_{i=0}^{8} \binom{8}{i} c_L(i) c_L(8 - i) \\ &= \frac{1}{2^8} \sum_{i=0}^{8} \binom{8}{i} (a + bi)(a + 8b - bi) \\ &= a^2 + 8ab + 14b^2\,. \end{aligned} \tag{7}$$

[4] Certainly it is plausible to assume that the coefficients are monotonically increasing functions of gross profit j, and a linearity postulate yields the mathematically simplest function having this property. Empirical support for this assumption is given in the final section.

$$
\begin{aligned}
P((H, L) \mid (L, L)) &= \frac{1}{2^8} \sum_{i=0}^{8} \binom{8}{i} c_L(i)[1 - c_L(8 - i)] \\
&= \frac{1}{2^8} \sum_{i=0}^{8} \binom{8}{i} (a + bi)(1 - a - 8b + bi) \qquad (8) \\
&= [a + 4b - (a^2 + 8ab + 14b^2)] .
\end{aligned}
$$

Applying the same methods to the other transition probabilities, we obtain the following transition matrix in terms of the four parameters a, b, α and β, where

$$A = a^2 + 8ab + 14b^2$$

$$B = \alpha^2 + 6\alpha\beta + \frac{15}{2}\beta^2$$

$$C = a + 6b$$

$$D = \alpha + \frac{3}{2}\beta ,$$

(9)

	(L, L)	(L, H)	(H, L)	(H, H)
(L, L)	A	$a + 4b - A$	$a + 4b - A$	$A + 1 - 2a - 8b$
(L, H)	$C(1 - D)$	CD	$(1 - C)(1 - D)$	$(1 - C)D$
(H, L)	$C(1 - D)$	$(1 - C)(1 - D)$	CD	$(1 - C)D$
(H, H)	$B + 1 - 2\alpha - 6\beta$	$\alpha + 3\beta - B$	$\alpha + 3\beta - B$	B

There are three important observations to be made about the kind of economic behavior implied by the transition matrix (9). In the first place, unlike the theory of Cournot and many other writers, the theory does not predict the eventual adoption of a pure strategy by either firm. Secondly, unlike the Edgeworth theory and others which predict cyclic behavior, the present theory predicts random fluctuations in behavior for both firms. Equilibrium exists, but in a statistical rather than deterministic sense. Thirdly, the present theory is more detailed than the classical economic theories of duopoly, for innumerable predictions about sequential behavior can be made from (9) and even more from the full theory embodied in the axioms given earlier. As already remarked, from the standpoint of behavior theory, much of classical economic theory is defective in not providing a detailed analysis of the dynamic process by which equilibrium is reached. (This same defect is characteristic of many psychological theories of behavior.)

4. EXPERIMENTAL METHOD

The subjects were 40 volunteers from undergraduate courses in the summer session at Stanford University. They were informed that the experiment was about "economic decision-making" and that they would be paid between $1.00 and $2.50 for an hour's time, the pay to depend on their decisions. They were heterogeneous with regard to their knowledge of economics, ranging from total naiveté to the equivalent of an A.B. degree in economics. The subjects were run in pairs and randomly assigned to the two experimental groups.

The experimental apparatus was a table partitioned so that the subjects could not see each other, but could both see the experimenter, who sat facing them. They had cards in front of them with which to indicate their responses, and poker chips to use for currency. The experimenter seated them and said:

> This experiment is one in which each of you plays the part of a firm in a simplified economic market. The market we have in mind is the following: There are exactly two firms, who between them control all of the goods to be presented. Their outputs are identical, so that the buyer has no reason to choose one over the other except for price. For example, we may think of two producers of electronic calculating machines, assuming that these two firms are the only manufacturers of these machines, and that their outputs are so close to equivalent that price provides the only reason to choose one over the other. We further assume that each of these firms is able to produce up to, but no more than, six calculators in the time period we are considering, say every three months.
>
> Below is the specific demand curve we have postulated for this market. (The linear demand curve was shown to subjects.) Each of you can set the price of your product independently at either four or six units per calculator. You will note that if both of you set price four, there is total demand eight, while if both set price six, there is demand for six machines. Your costs in producing these calculators are assumed to be constant; that is, so far as cost is concerned, it is immaterial to you whether you produce two or six calculators.
>
> To be more specific about the demand curve and the behavior of the consumers, we may imagine that it is made up of twelve consumers, the first of whom is willing to pay up to twelve units for a machine, the second, up to eleven units, etc. Thus, if both of you set price four, there are eight consumers who will pay four units or more, and all will attempt to buy a machine. If you both set the same price, the consumers are indifferent as to whom they buy from, and so decide essentially at random. Thus the number of customers buying from each of you will vary randomly from

period to period in the event that you both set the same price.

If one of you sets price four and the other sets price six, then all consumers will attempt to buy at the lower price, and the first six to get there (recall that you can only produce six machines) will buy. The remaining two will buy at the higher price if and only if they were among those who were originally willing to pay the high price. That is, among the eight consumers who were willing to buy at the cheap price, there were two who would only buy at that price and if they were not among the six served, they will not pay the high price.

As strict competitors, you are not allowed to get together and form a monopoly. In fact, no communication will be allowed between the two of you. Each of you will set his price independently; tell me that price and I will calculate resulting sales and profits for each. The experiment will then consist simply of a large number of these decisions. As I mentioned before, costs are fixed and will be fourteen units per quarter. Thus, we will subtract fourteen from your gross sales each period, and the resulting figure is your profit or loss for that period. We will use these chips to keep track of your profits and losses, and pay you at the end for the chips you have.

Questions were answered by paraphrasing the instructions, and the subjects were then run for 200 trials. The amount of sales for each subject for a fixed pair of prices was determined in advance from a table of random numbers in accordance with the distributions described above.

Ten pairs of subjects were run in each of the two experimental groups. In both groups, the gross profits on each trial of each firm were announced to both members of the pair. In addition, in Group T ("Told") the price that each player had set was announced. In Group N ("Not Told"), only gross profits were announced, and prices set by the opponent were unknown.[5]

5. RESULTS

Figure 2 shows the first result of interest, the changes in probability of the various pairs of responses over the 200 trials, plotted in 10 trial blocks. The response pairs (H, L) and (L, H) have been collapsed into a single pair (L, H). These learning curves are remarkably orderly (see also Figure 4 below), but of course show none of the cyclic or pure strategy behavior predicted by classical economic theories of

[5] It is, of course, possible in most, but not all, cases to determine the opponent's price on the previous trial from the total sales. However, only one subject in Group N gave any indication of doing this.

duopoly. There seems to be some evidence that the asymptotic response patterns have not yet been attained.

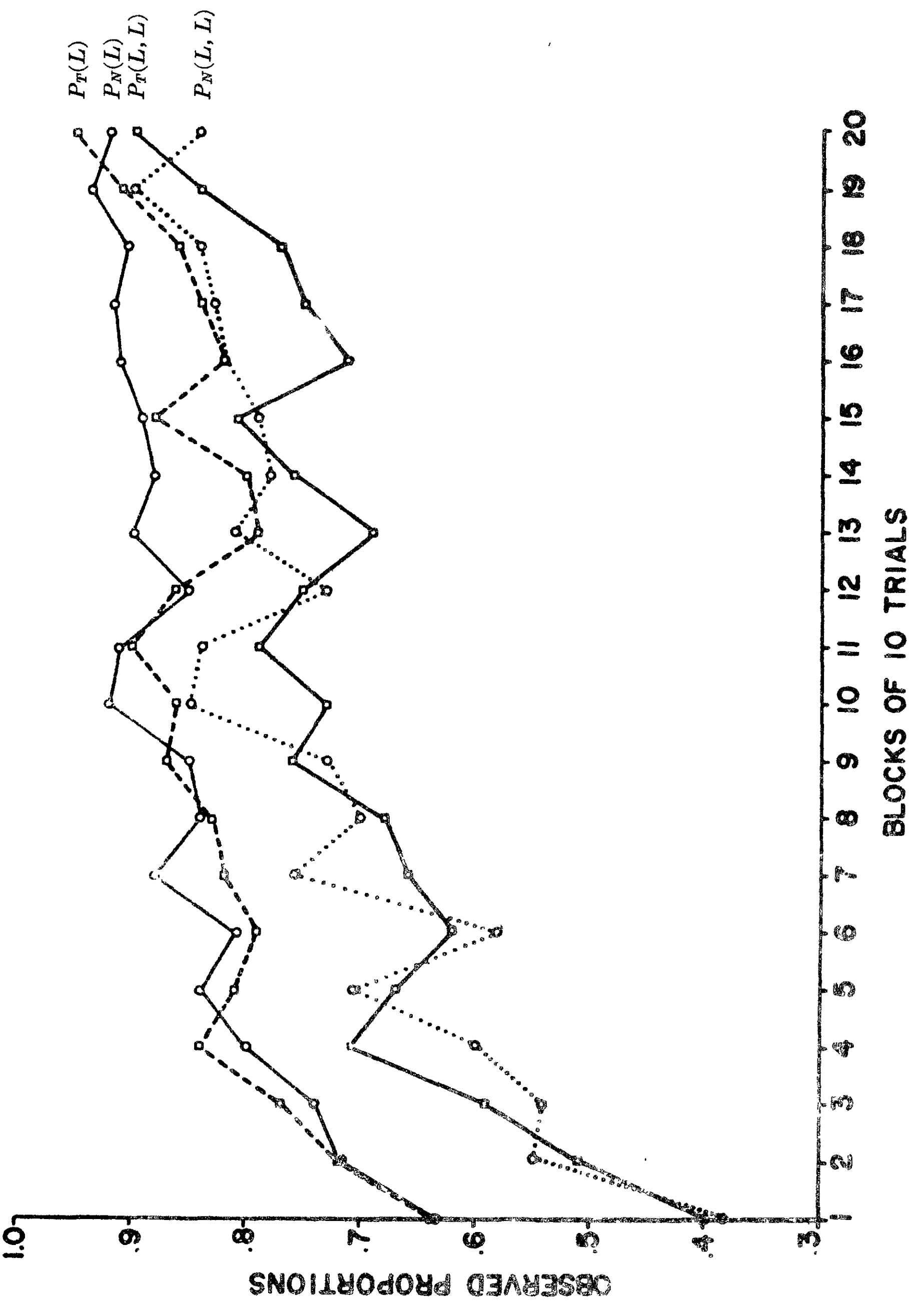

FIGURE 2

OBSERVED RESPONSE PROPORTIONS IN SUCCESSIVE 10-TRIAL BLOCKS FOR THE TWO EXPERIMENTAL GROUPS

Several other comments may be made about these curves. First, it is clear that there are no major differences between the two experimental groups. A more exact test of this hypothesis will be presented later, but it is apparent to the eye that knowledge of the opponent's price on the previous trial does not seriously affect the average behavior of the subjects. Second, it is clear that the subjects are in no sense maximizing their joint utility. There is, with the exception of one pair of subjects, no evidence of collusion at the monopoly price. One pair of subjects in the "Told" group did both set the high price for approximately 40 trials, but no other subjects were able to maintain this collusive behavior for more than a few trials.

The next question we ask is whether, as the theory demands, the behavior of the subjects is in fact Markovian in the responses. To test this hypothesis, (a test which incidentally is free of any restriction to a model, but which tests the general hypothesis of the Markovian character) we use a χ^2 test proposed by Anderson and Goodman [1]. The null hypothesis is that $p_{ijk} = p_{jk}$ for $i, j, k = 1, 2, 3, 4$. That is, the null hypothesis corresponds to the theoretical statement that knowledge of responses made before trial $n - 1$ does not change the probability of a given response on trial n. For example, knowledge that a person has made two consecutive H responses instead of an L followed by an H does not affect the likelihood that he will make an H response on the next trial. The results confirm the hypothesis that the responses do form a first order Markov chain. For the first 100 trials, the value of χ^2 for Group T does not approach significance ($\chi^2 = 35.83$, df $= 36$, $P = 0.49$), and for Group N it is just significant at the 0.05 level ($\chi^2 = 51.51$). The test was not run for the last 100 trials because too many of the cell entries were close to zero. (See Table 1 below.)

Another demand of the specific model we have proposed here (which does not hold for the multi-stimulus model) is that the process is stationary in the observable response pairs, i.e., that the transition probabilities p_{ij} do not change over time. This hypothesis can also be tested independently of the model, again using a χ^2 test given by Anderson and Goodman [1]. The test can be made separately for each row of the matrix, and in both groups. The differences in the first row (L, L) are highly significant, while in the other three rows they are not. Thus there is some evidence that the process is not stationary, although in our opinion it is not serious enough to force a rejection of the model. Readers who do wish to reject the one-element model on this evidence will still find it expedient to use it in computations as an approximation to a multi-stimulus model.

When the transition probabilities of a Markov process are independent of the trial number, the transition numbers n_{ij}, which summarize the number of transitions from state i to state j, form a set of sufficient statistics for the process. Table 1 shows the transition

TABLE 1

TRANSITION NUMBERS FOR BOTH GROUPS AND FOR FIRST AND LAST 100 TRIALS

n \ $n+1$	First 100 trials (L, L)	(L, H)	(H, L)	(H, H)	Last 100 trials (L, L)	(L, H)	(H, L)	(H, H)
(L, L)	456 T 456 N	67 79	95 87	10 14	698 708	40 52	45 54	1 8
(L, H)	80 83	55 53	10 14	14 12	34 57	34 23	7 5	6 3
(H, L)	90 90	22 14	44 40	5 7	44 50	7 7	33 13	3 0
(H, H)	6 10	14 16	11 9	11 6	1 4	1 6	9 1	47 4

numbers for both experimental groups, split into the first 100 and last 100 trials. It is clear that with the exception of the last row in the last 100 trials, the two groups are essentially identical. The distorting entry $n_{HH} = 47$ is due to the one pair of subjects mentioned above who did adopt collusive strategies for approximately 40 trials. Applying a χ^2 test of homogeneity to test the null hypothesis that the two samples are drawn from the same Markov chain (Anderson and Goodman [1]), we find that for the first 100 trials, the resulting χ^2 (7.51, df = 12, $P = 0.90$) is not significant, and for the second block of 100 trials the χ^2 is also not significant when cells with entries of less than 5 are ignored. Because of this relatively close agreement, we felt justified in combining the two groups in much of what follows.

Turning now to more specific predictions of the theory proposed here, Figure 3 shows the functions $c_i(j)$, plotted separately for the first 100 and last 100 trials. Because of the evidence presented above for the homogeneity of the two groups, we have pooled the data from the two groups in plotting these curves. We may note first that on a qualitative level, the curves are completely in accord with our conceptual model. That is, the probability of making the same response on the next trial is, with the exception of unstable end points, a

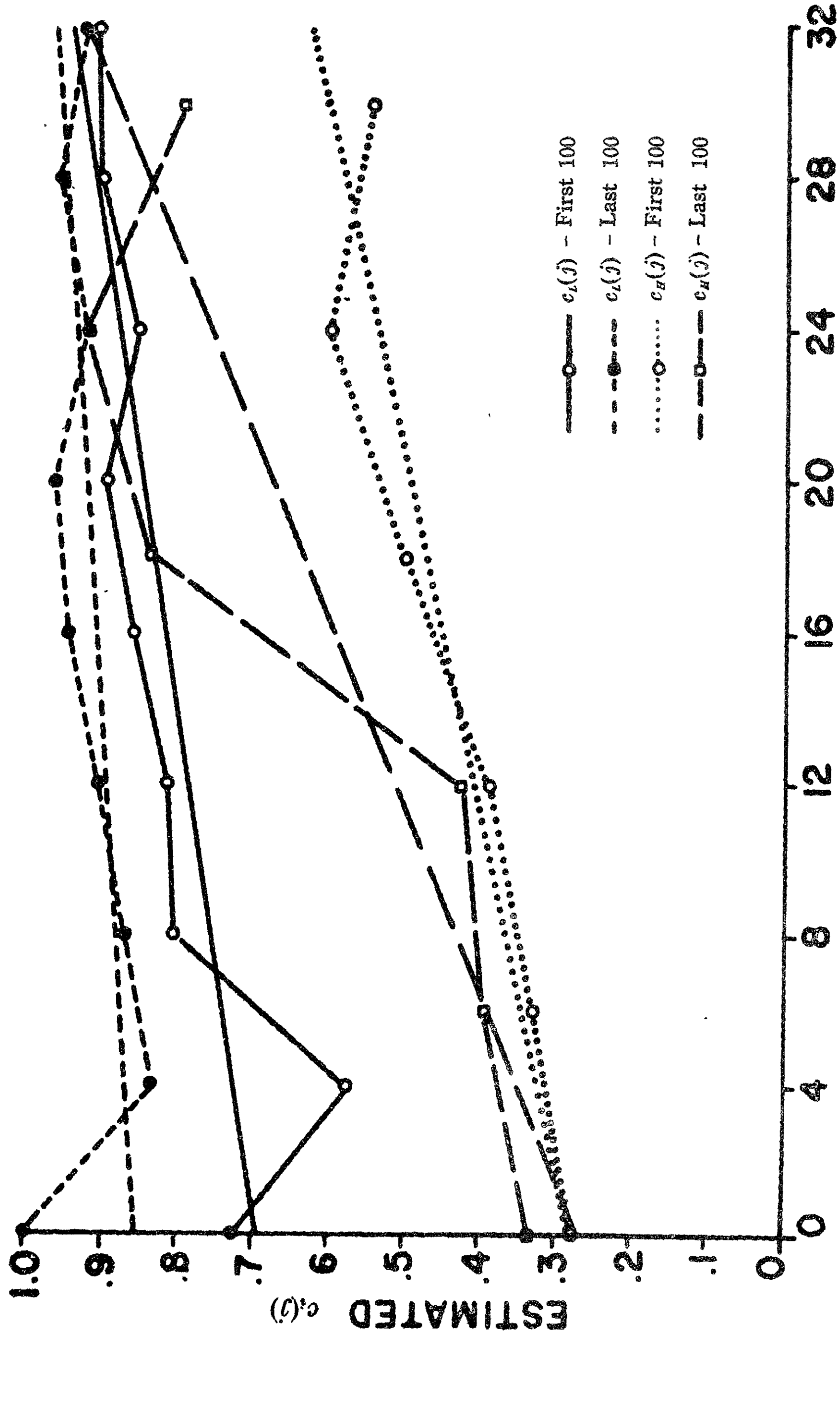

FIGURE 3

ESTIMATED CONDITIONING PARAMETERS AND LEAST SQUARES FITTED LINES strictly increasing monotonic function of the reinforcement given.[6] This result gives strong support for the general learning and reinforcement framework within which we are attempting to describe duopoly behavior. Whether or not the specific model we propose is appealing, this finding would seem to show that the approach to the problem via reinforcement ideas is a valid one.

One straightforward method of estimating the parameters a, b, α and β of (5) and (6) is by fitting straight lines to the data of Figure 3 by the method of least squares. The results are shown in the figure; the numerical values of the parameters are given in Table 2. The small values of b and β are to be expected. This is evident from Figure 3. The abscissa runs from 0 to 36 whereas the ordinate runs from 0 to 1. Meaningful results will thus necessarily be lines with relatively flat slopes. The data for $\hat{c}_L(0)$ were not used on the last 100 trials because the estimated probability is based on only 13 observations and is decidedly out of line with the other estimated probabilities $\hat{c}_L(j)$ for the last 100 trials. It is interesting to note that $\beta > b$ for both the first and last 100 trials, which we take to indicate that when subjects set the rather risky high price they are more sensitive to the exact amount of reward than in the case of setting the conservative low price.

TABLE 2

ESTIMATED PARAMETERS FOR FIRST AND LAST 100 TRIALS

	First 100 trials	Last 100 trials
a	.690	.856
b	.008	004
α	.280	.274
β	.011	.021

Of the four linear fits, the one for the high price on the last 100 trials is clearly the worst. The evident departure from linearity warrants further study in other experimental situations, for it is to be emphasized that the exact character of the goodness-of-fit results discussed below depend heavily upon the assumption that the conditioning parameters $c_i(j)$ are linear in the amount of reward. It is apparent from inspection of Figure 3 that the linearity assumption is not a bad one in three of the four cases.

[6] The instability of the endpoints is due to the small number of observations available to estimate $c_L(0)$. No subjects received gross income $j = 36$, and thus this point does not occur on the abscissa.

It is to be remarked that the assumption of linearity at this point is not equivalent to the assumption of a utility function which is linear in money. The conditioning parameters $c_i(j)$ mainly describe the changes in conditioning from trial to trial, given the reinforcement and response. The new response occurs on an all or none basis of conditioning and is not due to an estimation of expected utility. Even if a utility interpretation were given to the response probabilities (see Suppes [9]), the conditioning parameters would essentially describe changes in utilities and not the levels of utilities.

Having estimated the parameters a, b, α and β we may obtain the predicted transition matrix shown in Table 3. The observed transition probabilities $\hat{p}_{ij}$, based on the data from both groups, are also shown

TABLE 3

PREDICTED AND OBSERVED TRANSITION PROBABILITIES
(Observed values in parentheses)

	First 100 trials				Last 100 trials			
	(L, L)	(L, H)	(H, L)	(H, H)	(L, L)	(L, H)	(H, L)	(H, H)
(L, L)	.67 (.72)	.15 (.12)	.15 (.14)	.03 (.02)	.83 (.88)	.08 (.06)	.08 (.05)	.01 (.01)
(L, H)	.55 (.51)	.34 (.34)	.07 (.07)	.04 (.08)	.51 (.54)	.43 (.34)	.03 (.07)	.03 (.05)
(H, L)	.55 (.58)	.07 (.11)	.34 (.27)	.04 (.04)	.51 (.58)	.03 (.09)	.43 (.31)	.03 (.02)
(H, H)	.26 (.19)	.26 (.36)	.26 (.24)	.22 (.21)	.10 (.07)	.25 (.09)	.25 (.14)	.40 (.70)

for comparison. The agreement between observed and predicted values seems reasonably good for both the first and last 100 trials. We may use a χ^2 test again; in this case the null hypothesis is that the data (the observed transition probabilities) fit the model (the predicted probabilities). The χ^2 distribution under the null hypothesis has 12 degrees of freedom, but 4 parameters have been estimated; thus the significance level should be interpreted in terms of 8 df. For both blocks of trials the results are highly significant ($P < 0.001$). In interpreting these results it is important to realize that this test is extremely sensitive when the number of observations is large—here 2000 for each block. Few of the learning experiments reported in the psychological literature are interpreted in terms of models that do not yield highly significant results on overall goodness-of-fit tests. (For extensive discussion of such tests, see Suppes and Atkinson [10].)

The theoretical matrix may be used to predict the mean asymptotic probability of response. In Figure 4, the learning curves in 20 trial

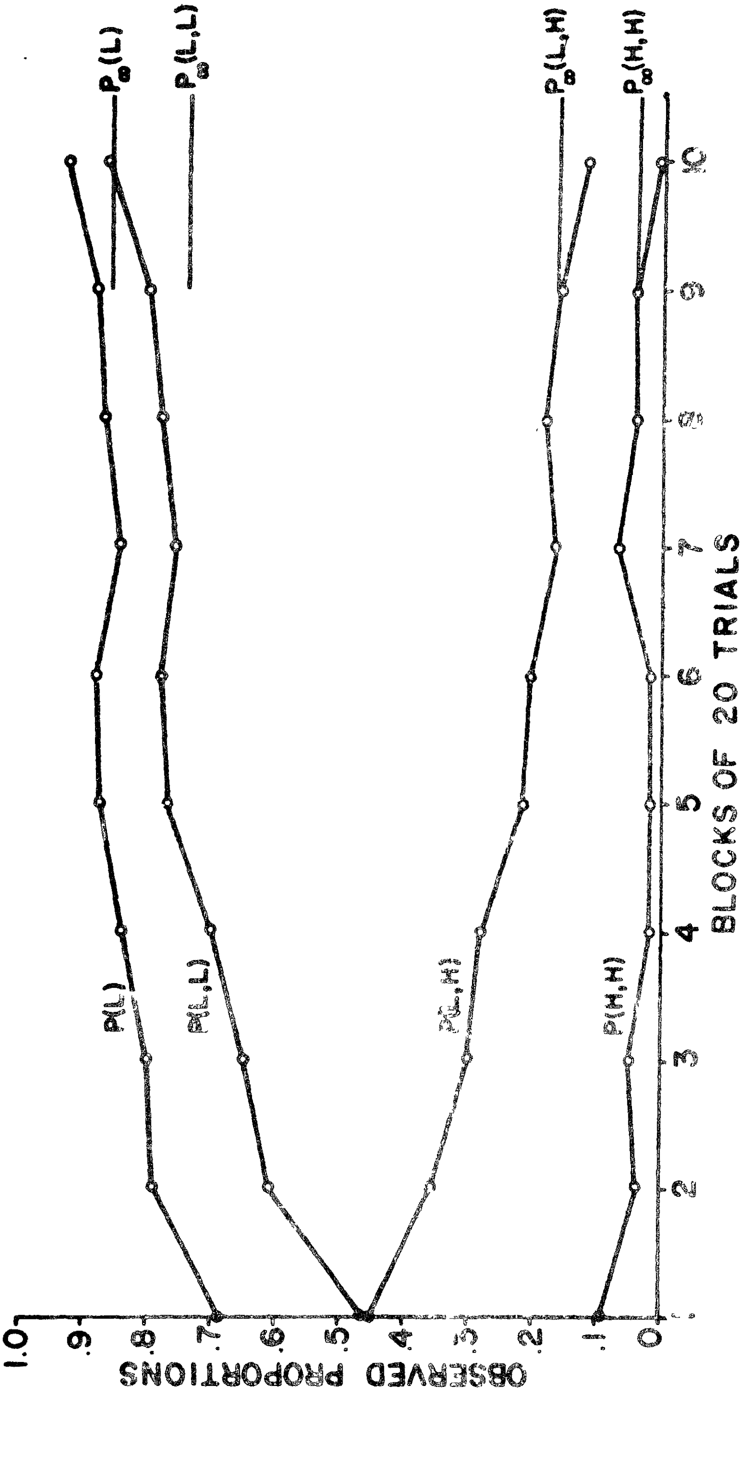

FIGURE 4

COMPARED OBSERVED [illegible] PROPORTIONS IN SUCCESSIVE 20-TRIAL BLOCKS AND PREDICTED ASYMPTOTIC VALUES blocks, with the data pooled from both groups, are shown for the probabilities $P(L)$, $P(L, L)$, $P(L, H)$ and $P(H, H)$, where $P(L, H)$ is the probability of either (L, H) or (H, L). The predicted asymptotes are shown as horizontal lines on the right. In discussing Figure 2 it was remarked that the learning curves seem not yet to have reached asymptotes. This is reflected in the discrepancy between the predicted asymptotes and the curves for the last two or three blocks of trials. This same discrepancy is found when we compute the asymptotic mean probabilities on the basis of the observed transition matrix for the last 100 trials given in Table 3. Toward the end of the experiment the subjects were tending more and more to set the competitive low price, with the one exception already noted. It would be desirable to determine if this tendency would be sustained over a much larger number of trials. If so, in order to maintain the present theory it would be necessary to postulate systematic changes in the conditioning parameters $c_i(j)$ over trials. (This seems necessary for other learning experiments as well; see Suppes and Atkinson [10, (Chap. 10)].)

On the other hand, it should be noticed that consideration of more than 200 trials is not realistic from an economic standpoint. When each trial is interpreted as a quarterly decision, 200 trials already cover fifty years. The external environment of no duopoly would be reasonably constant for any longer period of time, and most likely not this long.

Several other remarks may be made about the experiment. It may be argued that the data were collected in such a highly structured and oversimplified situation that they have little relevance to economic behavior. This is possibly true, and it would of course be desirable to collect similar data in the field. But then one runs into the problem of lack of control of many variables which are obviously relevant, but which are not as yet incorporated into the theory. It was primarily for this reason that we chose to do the experiment in the laboratory, where it was possible to control many more of these variables. This is, of course, common practice in psychology, but economists have not generally adopted this view.

It is also true that we were using naive subjects. It is likely that two sophisticated economists or businessmen would have reacted differently in this setting, perhaps settling on the advantageous monopoly price. Although this statement should be verified before it is accepted, it is a not unreasonable hypothesis, since the situation is extremely simple, and the results of any behavior can be easily computed without performing the behavior. However, in a more complicated market, the businessman cannot so easily compute the result of any given behavior. He would seem rather to be in the position of our subjects, who must perform the behavior, receive the reinforcement and react accordingly. Thus, we would expect the businessman to behave in a manner describable by the kind of reinforcement and conditioning ideas we have tried to isolate (and we would be willing to discount most of his verbal rationalizations derived from his knowledge of economic theory).

Finally, it is certainly true that the theory is at present far too simple to describe much actual economic behavior. The direction of generalization seems clear, although problems, both experimental and mathematical, are present. For example, one obvious need is the continuous analogue of the theory, where the possible prices the firm can set form a continuum. The mathematical problems involved are substantial, although not insuperable. A more direct generalization of interest is to the case of oligopoly with more than two firms and with more than two possible prices. There is in fact no conceptual obstacle to applying the theory to perfectly competitive markets by considering a large number of firms.

Stanford University, U.S.A.

REFERENCES

[1] ANDERSON, T. W. AND L. A. GOODMAN, "Statistical Inference about Markov Chains," *Annals of Math. Stat.*, XXVIII (1957), 89-110.

[2] BURKE, C. J. AND W. K. ESTES, "A Component Model for Stimulus Variables in Discrimination Learning," *Psychometrika*, XXII (1957), 133-45.

[3] ESTES, W. K., "Toward a Statistical Theory of Learning," *Psychological Review*, LVII (1950), 94-107.

[4] ESTES, W. K. AND C. J. BURKE, "A Theory of Stimulus Variability in Learning," *Psychological Review*, LX (1953), 276-86.

[5] ESTES, W. K. AND P. SUPPES, "Foundations of Linear Models," *Studies in Mathematical Learning Theory*, ed. by R. R. Bush and W. K. Estes (Stanford: Stanford University Press, 1959), ch. 8.

[6] ———, "Foundations of Statistical Learning Theory, II. The Stimulus Sampling Model," Technical Report No. 26, Contract Nonr 225(17), Institute for Mathematical Studies in the Social Sciences, Stanford University, 1959.

[7] LUCE, R. D. AND H. RAIFFA, *Games and Decisions* (New York: John Wiley and Sons, 1957).

[8] SHUBIK, M., *Strategy and Market Structure* (New York: John Wiley and Sons, 1959).

[9] SUPPES, P., "Behavioristic Foundations of Utility," *Econometrica*, XXIX (1961), 186-202.

[10] SUPPES, P. AND R. C. ATKINSON, *Markov Learning Models for Multiperson Interactions* (Stanford: Stanford University Press, 1960).