WORKING
PAPERS

EQUILIBRIA IN THE LABORATORY:

EXPERIMENTS WITH OLIGOPOLY MARKETS WHERE GOODS ARE MADE-TO-ORDER*

Dan Alger

WORKING PAPER NO. 121

December 1984

BUREAU OF ECONOMICS
FEDERAL TRADE COMMISSION
WASHINGTON, DC 20580

EQUILIBRIA IN THE LABORATORY:
EXPERIMENTS WITH OLIGOPOLY MARKETS WHERE GOODS ARE MADE-TO-ORDER*

by

Dan Alger

Bureau of Economics

Federal Trade Commission

November 1984

A B S T R A C T

This paper is primarily concerned with examining one common practice used in previous studies of laboratory markets. This practice is to test equilibrium models using a data set that includes data from markets that may not have reached an equilibrium. The laboratory markets examined here demonstrate that equilibrium data may be quite different from that data achieved after approximately the number of time periods used in previous studies. For these markets, tests of equilibrium models using equilibrium data would yield quite different results than using the data from the truncated sequence that includes some disequilibrium behavior.

1. Introduction

This paper is primarily concerned with examining one common practice used in previous studies of laboratory markets. This practice is to test equilibrium models using a data set that includes data from markets that may not have reached an equilibrium. The data used typically has consisted of data from all laboratory markets run and may be from the last period, an average of the last few periods, an average of all periods, or from the nth period for some fixed n. I know of no attempt, at least formally, to assure that all data was from markets in equilibrium. The primary question here is whether the use of data from markets that may not have reached equilibrium can significantly affect the results of tests of equilibrium models.

This practice would be useful, if the "disequilibrium" behavior could provide a useful estimate of the behavior that would have been observed if an equilibrium had been obtained. This would include cases where either the disequilibrium distribution centers around an equilibrium outcome or the disequilibrium outcomes gradually converge to an equilibrium so that this limit could be estimated. On the other hand, if for some markets some disequilibrium behavior is qualitatively different from the behavior that would be observed in an equilibrium, this practice could produce misleading results.

Within this paper, the effect of this practice will be examined in markets similar to those in two previous experimental studies, one by Ketcham, Smith, and Williams [KSW] and one be

Fouraker and Siegel [FS]. All of these laboratory markets were oligopoly markets for a homogeneous product where sellers post prices and goods are made-to-order. The primary comparison for this paper only uses the data from my laboratory markets and is between equilibrium behavior and the behavior observed after approximately the same number of market periods used in these two previous studies.

The data from the markets reported on here suggest that using all data available after approximately the same number of market periods as in KSW and FS would lead to significantly different equilibrium test results than restricting tests to equilibrium data. If the sequences of data generated here are truncated to approximately the same number of market periods used in these two studies, significantly different behavior is observed than in equilibrium.

Also, "collusive" behavior occurs much more frequently in the equilibrium data here than in these previous studies, and it appears that achieving equilibrium takes much longer to accomplish than had been anticipated. The equilibrium outcomes here seem to be concentrated around the monopolistic outcome rather than near the competitive outcome, which is what was observed in these two previous studies. The cause(s) of these observed differences could be any of the design differences between these sets of experimental markets, but given the results comparing my equilibrium data to my data from the truncated sequence, the difference in the number of time periods seems suspect. More experimental markets need to be run that control for the different design differences before the cause(s) can be determined.

2. Experimental Design and Procedures

The experimental markets reported on here consisted of a sequence of trading periods with identical market structures. The desired characteristics for each subject were induced by the profit structure of these markets, and these profits were paid to the subjects in cash (see Smith, 1976). Each period started with each seller choosing a price to charge for his good and the maximum quantity he was willing to offer to the market. Before these decisions had to be made, each seller had access to his personal cost information and the past history of the market, which consisted of the decisions made by all sellers, the quantities ultimately sold by all sellers, and his own profit for all previous periods. After the price and quantity choices were made by each seller, the buyers' behavior was simulated to follow that of a perfectly competitive buyer. These responses from the buyers determined the quantities actually sold and produced, and ultimately the profits for each seller. We say that the goods are made-to-order as sellers are assessed the costs only for those units they sell, and not those that are offered to the market but are unsold. The results for this period were then revealed to the sellers as the next period started.

The physical setting for these experimental markets was the PLATO computer lab at the University of Arizona. All of the experiments reported on here were conducted using the PLATO computer system with a program developed by me to run many different oligopoly experiments. The computer was used to store and control the relevant market information and to simulate buyer

behavior. Subjects sat in front of individual terminals, and all interactions during the experiments were limited to those with the terminal. There was no direct communication between any of the subjects. The subjects themselves were University of Arizona students.

As a primary interest for this research would be with equilibrium data, two alternative rules for terminating each experimental market immediately suggested themselves. One was to run each market a fixed number of time periods or until a fixed real time limit was met, and later screen out the data from those markets that did not satisfy the operational definition of an equilibrium. Another was to terminate the operation of any market as soon as it satisfied the definition, and assume that if this market had continued operating, this same "equilibrium" behavior would have been observed indefinitely. To reduce any potential bias from the termination rule[1] and to reduce the amount of subject payments needed to generate a given amount of useful data,[2] the second approach was selected.

Given this approach, an operational definition of an equilibrium would be needed to determine when to terminate an experiment. The ideal in an ideal world would be a definition that indicated when each buyer and seller would repeatedly make the same choices after being given sufficient experience with the same decisionmaking environment, but this is not possible as an infinite number of trials cannot be observed. To make a definition of an equilibrium operational, an approximation must be used and some elements of it must be arbitrarily chosen. One of these arbitary elements is the number of time periods the same behavior

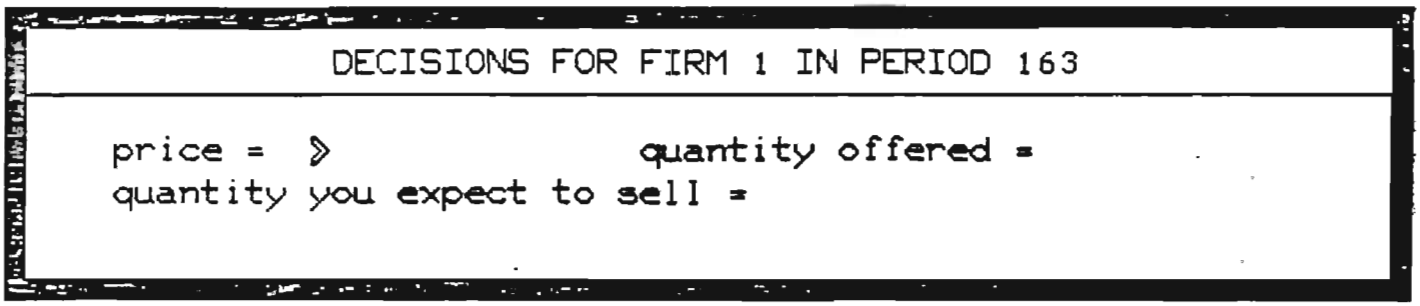
needs to be observed before one assumes it will last indefinitely. Another, since incentives are not perfectly controlled, concerns how close choices have to be before they are considered the "same."

In any event, these markets were terminated (and usually the subjects were re-assigned to a new market) when the following operational definition of an equilibrium was satisfied. The primary requirement was to exhibit some constancy of behavior across time periods (see Alger, 1984). This was measured by the largest deviation in profits over the last five periods and over the last ten periods.[3] For the experiments reported on here, the markets were terminated only if this deviation was zero for at least five periods (usually ten periods) or if an obvious cyclic pattern developed. A secondary requirement was to have each subject show that he had some understanding of the consequences of his choices. To measure this "understanding" each subject was asked to enter, along with his market decisions, the quantity he expected to sell given his market decisions, and the difference from the expected profit implied by this entry and the actual profit was calculated. If this difference between the expected profit and actual profit was zero, a certain amount of understanding was shown. These markets were terminated only if this difference was zero for at least five periods. The subject's response on the expected quantity sold, necessary to calculate this difference, was not well-motivated (i.e. it did not affect the cash reward), but it was thought this measure might be useful if changing behavior was observed after many periods.

The use of this stopping rule means some of the experimental markets could continue, as some did, for very many periods. These rules, with the large number of time periods possible when using them, could not have been used in the designs chosen by KSW and FS, primarily because of the real time constraint for running one session comfortably. They could not significantly increase the number of time periods run within three hours.

Allowing for a large number of time periods meant that either the length of the sessions had to increase or the real time needed for each market period had to decrease. The former is difficult since fewer students are willing to participate beyond three hour intervals, and running a market across several days is usually not desirable as any communication among subjects within this break in the session is uncontrolled. Here, the time needed for each market period was shortened substantially by simulating the behavior of the buyers, so that in effect the buyers' choices were made instantaneously. Here the buyers were assumed to be acting as perfect competitors. The cost of using this procedure is the possibility that actual buyers may not behave in the way that has been assumed and the sellers may react differently because of it. Fortunately, this cost now seems acceptable as data from previous experimental markets suggest that in this market environment the buyers do act competitively, even with a relatively small number of them.[4]

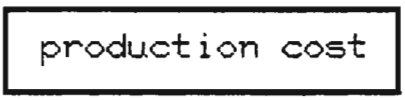
Figure 1 illustrates the screen display that was seen by a seller in these markets. The decision box at the top indicates the decisions that had to be entered for each time period. Each seller had to enter the price he was to charge, the quantity he



Market History

period	firm	price	qty offered	qty sold	profit
162	1	$ 1.24	12	3	$ 0.72
	2	1.24	12	3	
161	1	$ 1.24	12	3	$ 0.72
	2	1.24	12	3	
160	1	$ 1.24	12	3	$ 0.72
	2	1.24	12	3	
159	1	$ 1.24	12	3	$ 0.72
	2	1.24	12	3	
158	1	$ 1.24	12	3	$ 0.72
	2	1.24	12	3	
157	1	$ 1.24	12	3	$ 0.72
	2	1.24	12	3	
156	1	$ 1.24	12	3	$ 0.72
	2	1.24	12	3	
155	1	$ 1.24	12	3	$ 0.72
	2	1.24	12	3	

earlier periods



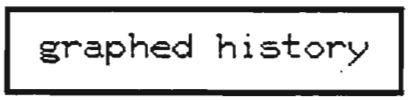


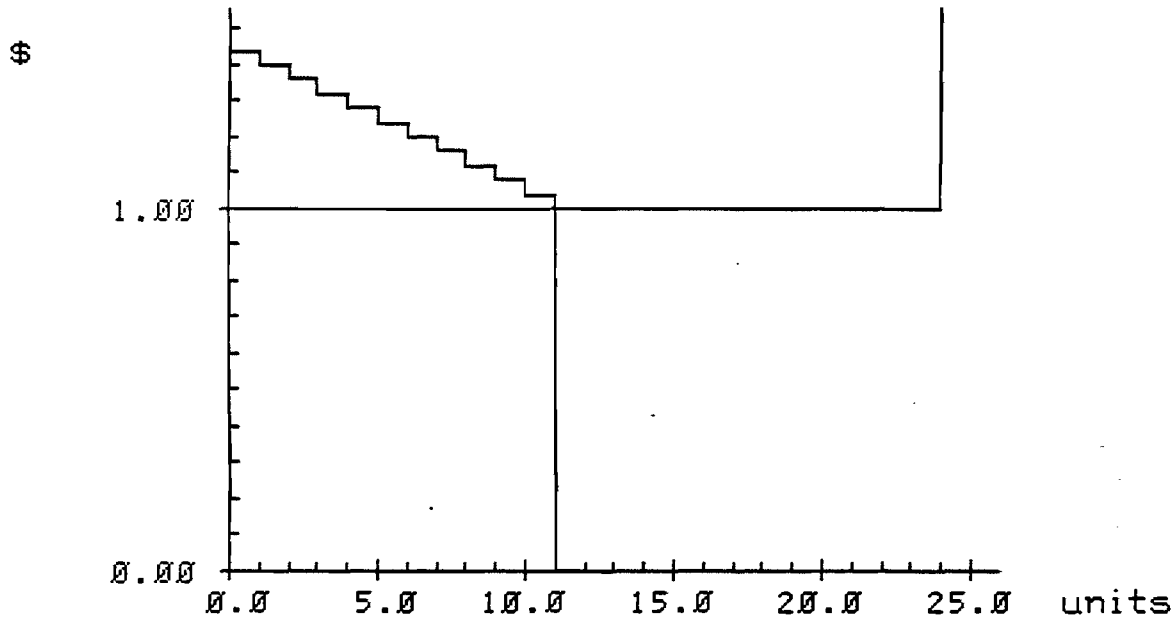
Figure I: Sample Screen Display for a Market Period

would offer to the market, and the quantity he expected to sell given the previous choices. The remainder of the display indicated the information available to the seller before these decisions had to be made. This consisted of cost information and information on the market history. Cost information could have been obtained directly by touching the cost box (PLATO has touch sensitive screens), which resulted in going to another display that gave a graph of the cost function and allowed queries about cost for specific quantities. Cost information could also have been obtained indirectly as the expected profit was calculated and shown to the seller once the expected quantity sold was entered. New choices could have been entered at any time prior to confirming them, updating the expected profit each time, so that the subject could consider the possible effect of alternative choices. Choices were made final by touching the decision box twice to confirm them. The market history was given in the table in the center of the display and in graphical form, by pressing the graphed history box. Thus, each seller knew all past market choices, the resulting sales from these choices, and his own profits. He was not given the costs or profits for any rivals. He was not given any information on market demand or on how the buyers might react to these choices. This had to be learned from experience.

The market parameters used for these markets are shown in Figure 2. Each of two firms had a marginal cost of $1.00 for all units up to a capacity constraint of 12 units. The market demand was linear but with units restricted to whole numbers. The

Summary of Market Parameters

1. 2 sellers
2. dynamic environment--1 period information lag
3. goods made-to-order
4. high-to-low rationing rule



5. Supply & 6. Demand

(Press DATA for more on costs & demand.)

7. parameter shift of $ Ø.ØØ
8. entry bonus of $ Ø.ØØ

"2slrs;mto;d;Ø1" already has some data.
Its parameters are not to be changed.

Figure 2

market demand curve intersects (0 units, $1.48) and (12 units, $1.00). The amount demanded from an individual firm was calculated in the following way: if the firm was offering the lowest price, he faced the entire market demand; if both firms were offering the same price, then market demand was split in the same proportions as the amounts offered for sale; if the firm was offering a price higher than his rival, then he received any unsatisfied demand assuming buyers with higher reservation prices purchased first. The quantity actually produced and sold by any firm (i.e. the amount that determined the total cost) was either the quantity demanded from the firm or the quantity offered, whichever was smaller.[5]

3. Experimental Results

Compared to the results of KSW and FS, one notable pattern emerges from the data. For some markets it appears that there is an initial period where firms gain some information about market demand, typically with one firm undercutting the price of the other in each period, and prices fall rather gradually to somewhere near the competitive level. Later some firm signals a willingness to go to a higher price with a relatively large jump in its price choice, and this is sometimes followed by a higher price from the rival. This jump to a higher price level sometimes breaks down with price cutting, sometimes falling back to a near-competitive level, but sometimes it does not, resulting in an equilibrium. If the price level does fall after a price jump, this process is often repeated. Many equilibria started with one

of these jumps to a higher price level. Thus, for markets that follow this pattern we observe some behavior prior to an equilibrium that is drastically different from that observed in equilibrium. This pattern is illustrated by the price histories in Figure 3 for copy 1 of the market, where in 3(a) prices which accompanied a positive quantity being sold are graphed, "o" for firm 1 and "x" for firm 2, while in 3(b) all price offers are graphed. (A higher price usually meant no sales.) The data in KSW and FS looks like the initial period of this pattern, where the market price falls rather gradually to near the competitive level.

The equilibrium data here appear to be quite different from the data in KSW and FS. The outcomes here are concentrated around the monopolistic outcome while those in the previous two studies are concentrated around the Nash equilibrium outcome. One possible explanation to reconcile these seemingly large differences is suggested by the price pattern noted above. If the markets in the previous studies were terminated in some period prior to an equilibrium and the pattern above was followed, prices might be observed that are much lower than prices that would eventually be obtained if the markets were allowed to continue to an equilibrium.

To test this hypothesis we need to consider the outcomes that would be observed if the markets here were truncated to approximately the same number of time periods as in these two previous studies, and then compare this data to the equilibrium data. The markets in KSW with the most similar design are those in their Design II, and these had a last time period that varied

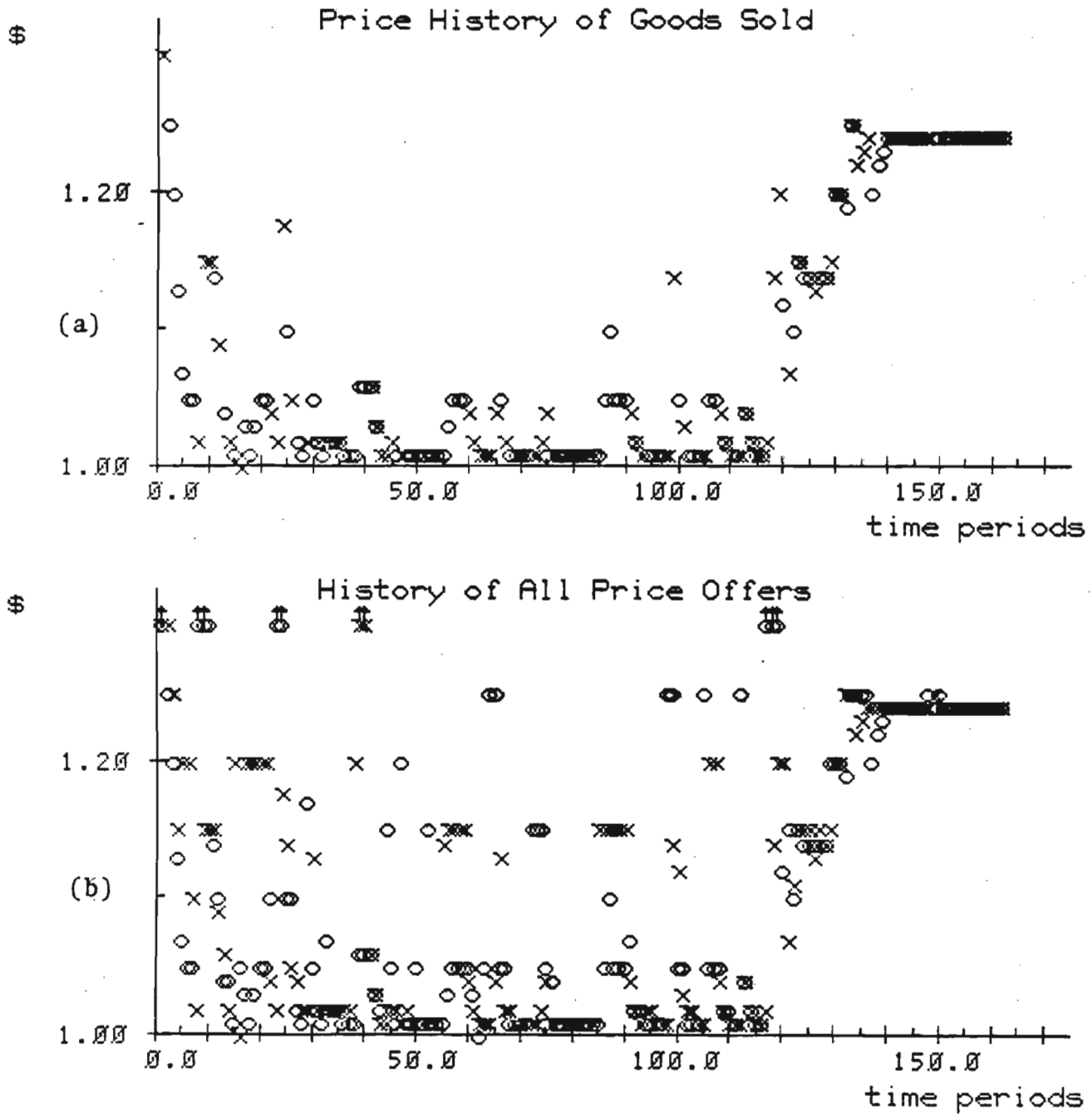

Price History of Goods Sold
$
1.20Ø
(a)
1.ØØ
Ø.Ø
5Ø.Ø
1ØØ.Ø
15Ø.Ø
time periods
History of All Price Offers
$
1.20Ø
(b)
1.ØØ
Ø.Ø
5Ø.Ø
1ØØ.Ø
15Ø.Ø
time periods

Figure 3

between 15, 20, and 25 time periods. The markets in FS with the most similar design are their duopoly and triopoly experiments with incomplete information, and the last period for each of these was period 15. All of the markets here continued at least for 16 time periods. Given this information, we will compare the equilibrium outcomes to the outcomes observed if the data sequence for each market is truncated after its 15th period.

The data we will use is summarized in Table 1 and in the graphs of Figure 4.[6] As can be seen, some of these markets failed to reach an equilibrium. Those markets that did fail to reach an equilibrium did so for a variety of reasons: copies 3 and 14 did not reach an equilibrium within the time constraints of one session, copy 6 had one subject walk out in the middle of the experiment due to low earnings, and copies 7 through 10 ended with a system crash. All of these markets did get through 15 periods.

The data reported for the previous studies are only from their design that is the closest to the one used here, and each is normalized to make comparisons between the different studies.[7] In spite of these choices, there remain several potentially important differences in the designs, in addition to the difference in the number of time periods run. KSW had three sellers and four buyers, FS had two or three sellers and "many" buyers, and I had two sellers and "many" buyers. KSW had relatively low physical capacities where no seller had the capacity to serve the whole market, but FS and I had large capacities. KSW and I allowed the sellers to make quantity choices, while FS did not.

copy	equilibrium outcome	periods	outcome after 15 periods
1	$ 1.24 , 6 units	162	$ 1.01 , 11 units
2	1.18 7	123	1.15 8
3	none	113	1.03 11
4	1.20 7	106	1.11 8
5	1.03 11	107	1.01 11
6	none	23	1.02 10
7	none	23	1.01 11
8	none	29	1.08 10
9	none	30	1.08 10
10	none	16	1.009 11
11	1.20 7	84	1.034 11
12	1.05 10	68	1.10 8
13	1.20 7	59	1.20 7
14	none	59	1.055 10
15	1.24 6	83	1.18 7

Table 1: Data Summary

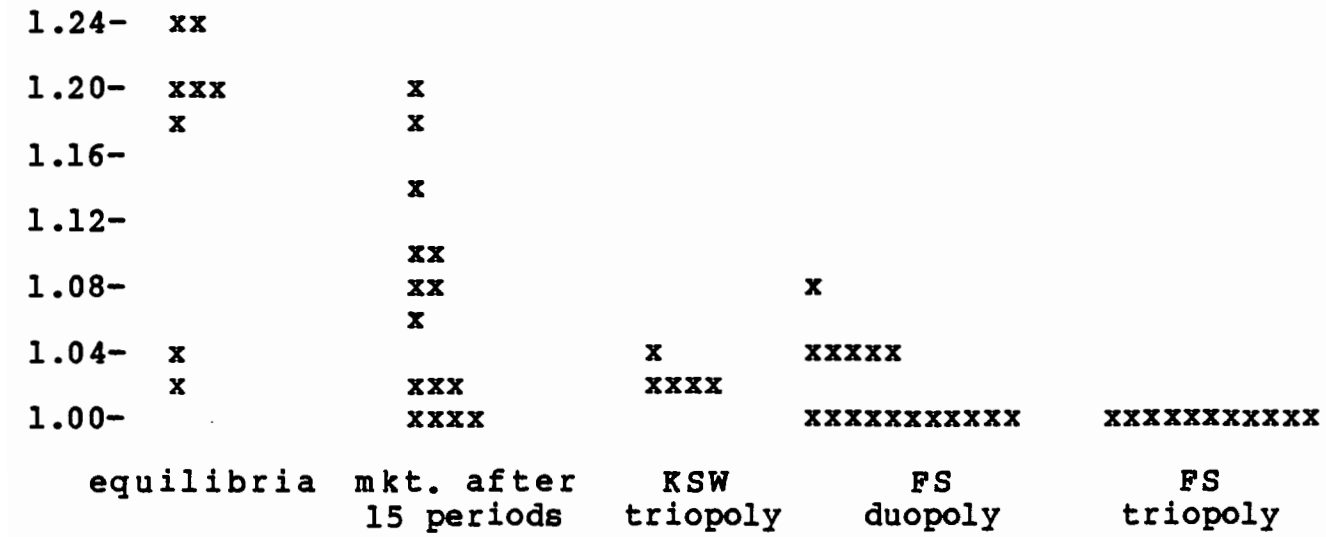


Figure 4: Sample Price Distributions

FS implicitly gave each seller information that fully described market demand, while KSW and I gave the sellers no market demand information. FS ran all experiments orally and kept records of transactions by hand, while for KSW and I all experiments were computerized. Instructions were presented differently in each set of experiments. Thus, because of these design differences, many possible explanations will remain as possible causes of any observed differences in behavior.

All of the hypotheses to be tested are concerned with whether the data from some pair of samples could have come from the same underlying probability distribution. As these outcomes typically lie along the market demand curve, we will only consider a one-dimensional description of each outcome -- the price (or an average if there is more than one) associated with some positive level of sales. For these tests we will use a one-tailed Kolmogorov-Smirnov test, a standard non-parametric test of this type of hypothesis. The major advantage of using a non-parametric test such as this is that the statistical assumptions made in any null hypothesis are very weak. With this test the only statistical assumption is one of continuous distribution functions.

The first hypothesis to be tested concerns whether the pattern noted previously appears to be important for these markets. The null hypothesis is that the data from the truncated sequences come from the same probability distribution as the equilibrium data. Using a one-tailed Kolmogorov-Smirnov test, we can reject this hypothesis at a 2% significance level.[8] For the markets here, this suggests that equilibrium prices are signifi-

cantly greater than the prices in the initial period. The only difference between these markets is the number of time periods, so this is presumably the cause of this result.

The next hypotheses to be tested concern the comparisons with the data from KSW and FS. The null hypotheses for these tests must be joint hypotheses that include an assumption that the normalization process is appropriate. The first of these has the null hypothesis that the KSW data and the equilibrium data here come from the same probability distribution (plus the probability distributions are continuous and the normalization is appropriate). Using the one-tailed test, this hypothesis can be rejected at a 0.1% significance level. This test rejects this same hypothesis for the FS duopoly markets at a 0.5% significance level and for the FS triopoly markets at a 0.05% level. These results indicate that the equilibrium prices observed in the markets here are significantly higher than those prices observed in the previous two studies. The data from these previous studies cannot provide an estimate of an equilibrium that is consistent with the equilibrium data here (this may be different from the equilibrium data that would be generated in their markets).

The second of these comparisons with the previous studies' data has the null hypothesis that the their data and the data from the truncated sequences come from the same probability distribution. The hypothesis for KSW cannot be rejected at any standard significance levels (>10%), but the hypothesis for the FS duopoly can be rejected at a 3.6% level and for the FS tri-

opoly at a 0.1% level. The cause of these results is presumably some of the design differences between the different sets of laboratory markets other than the difference in the number of time periods. While the evidence does not yet make a clear case, the number of sellers is certainly suspected as being one of the contributory causes.

To many researchers conducting economic experiments, the large number of time periods run before an equilibrium was reached might be surprising. Copies 1 through 6 and 14 had used subjects that were inexperienced with this design, and for these subjects it took a minimum of 106 time periods before the equilibrium criteria were satisfied. Copies 11, 12, 13, and 15 used subjects familiar with the design but unfamiliar with the particular market parameters used and they took a minimum of 83 periods to reach an equilibrium. This is from four to eight times the number of time periods run in most experimental designs.

Another interesting observation is that the three outcomes with a price of $1.20 all had sellers who found ways to share the market equally over time, even though the amount demanded was seven units, and the distribution between sellers was deterministic and limited to whole numbers. All of them developed a cyclic pattern where the profits were split evenly over time, giving them average profits equal to what would have been earned if they could have sold 3.5 units. In two of these markets the sellers chose (1.20,12) and (1.25,12), and in each successive period they switched their choices. In this way, a seller would receive all of the profits from selling seven units ($1.40) in one period and nothing the next. In the third market one seller

alternated (in this case every eight periods) choosing (1.20,12) and (1.20,5) while the other always chose (1.20,12). In this way, because sales were proportional to the quantities offered, each seller's profits alternated between $0.60 and $0.80. Note that this was accomplished without any direct communication or any knowledge of the rival's costs.

It also appears that the sellers thought they had reached the monopolistic outcome for at least some of these outcomes that are near but below the monopolistic outcome. Some had apparently searched and never found that $1.24 yielded the highest profit ($1.44) for a monopolist. This can still be rational behavior, since it is costly to choose a different price to get a sample of the market demand once a "collusive agreement" has been reached. It is also more difficult to get a sample for a price higher than the agreement price because both sellers would need to raise their price, while sampling market demand for a lower price only requires one seller to change his price. It appears that because of this opportunity cost for sampling the market demand, only a relatively few points were sampled, and some markets never found the true profit maximizing point for a monopolist. ($1.20 was the profit maximizing point if the sellers were restricted to increments of $0.05.) Also, if a market reached an equilibrium without sampling the market demand at points yielding higher profits, no information is ever revealed to contradict such beliefs by indicating that higher profits are available.[9]

4. Conclusion

The data for the experimental markets reported on here demonstrate that restricting tests only to data from markets in equilibrium can affect the results of some tests of equilibrium models. Because of this, for all experimental studies it seems desirable to limit tests of equilibrium models to data from markets in equilibrium, or to run a few markets for a relatively large number of time periods to demonstrate that the estimates of equilibria from the short experiments are similar to the equilibria actually achieved in the long ones. Such procedures offer an extra measure of control over the data that is to be tested against any equilibrium predictions.

For oligopoly markets where sellers post prices and goods are made-to-order, this data suggests that "collusive" behavior may be more likely in equilibrium than previous experimental tests have indicated. It also appears that more time may be needed to generate this behavior than had been expected. Nevertheless, many more experiments designed to test which market conditions make collusive behavior more or less likely are needed before these statements should be given much weight.

These observations lead to several interesting questions that should be addressed in some future research. Is this collusive behavior caused by market elements within a particular time period, or are the dynamic links across time periods essential for it to occur? Is there a big difference in behavior in markets with two sellers as opposed to three? How does equilibrium behavior change when the opportunity cost of obtaining market

demand information is varied? Because of the seemingly long time needed for collusive behavior to emerge, will relatively small shocks to the system, such as a relatively small amount of uncertainty in market demand or rivals' costs, eliminate collusive behavior?

Footnotes

*. The views expressed here are not to be considered those of the Commission or of any individual Commissioner. I wish to thank the Bureau of Economics for its support of this experimental work; Vernon Smith, Mark Isaac, Charles Holt, Jeffrey Eisenach and many colleagues here, especially Michael Lynch, Gerard Butters, and Pauline Ippolito, for many useful conversations; and Peter Knez for his assistance in running the experimental markets reported on here.

1. An approach that discards outcomes that have not reached equilibrium before some fixed number of time periods or before some real time limit is met may cause some problems. A bias might be introduced as the discarded markets might lead to different equilibrium than those markets whose data are kept. This might happen, for example, if "noncooperative" subjects typically have an extremely long disequilibrium period and then reach an equilibrium with relatively low profits. In this case, this approach would be systematically eliminating low profit equilibria.

2. Quitting only after an equilibrium has been reached reduces the number, and thus the expenses, for those markets which do not reach equilibrium, and reduces the expenses for those market periods after the market has been shown to be in equilibrium and the equilibrium is just maintained.

3. This measure for indicating some constancy of behavior is not appropriate if either a cyclic pattern is followed or some

strategies are not fully revealed, such as with mixed strategies or actions that depend upon some previous actions. Using rolling averages of profits over, say, ten periods might be an improvement with cyclic patterns and mixed strategies, but for these markets, this was not necessary as their was no indication that mixed strategies were used and those markets that were not terminated when this measure was exactly zero for several periods developed an obvious cyclic pattern. I am not sure what criterion would be most useful for strategies where actions depend upon the previous history of the market.

4. See the data for the markets considered in Ketcham-Smith-Williams (1984) and Isaac-Ramey-Williams (1984). This data was not available when the designs of these previous studies were established, so that simulating buyer behavior would have had a higher cost for these earlier studies than here. These markets, of course, ran for many fewer periods than the laboratory markets here, but the buyers' behavior is likely to satisfy reasonable equilibrium criterion (see fn. 3 as the buyers' actions depend upon the sellers' previous actions), so that their behavior might be expected to continue indefinitely.

5. The program allows quite a variety of oligopoly markets to be run. There may be: from one to sixteen firms, a dynamic form or a static form of the market, a specified time lag before a firm learns of his rivals' choices, one of several posted-price institutions, one of several rationing rules, any market demand or cost functions, a specified shift of cost and demand curves to disguise them, and a specified lump-sum payment for entering the market.

6. All data is available from the author.

7. Prices are normalized so that the monopoly and competitive prices of all sets of experiments match. Each price datum in KSW is normalized to give a price equal to .65 times the KSW price plus 1.00. Each price datum in FS is normalized to give a price equal to .08 times the FS price plus .96.

8. It might be that those outcomes with low prices in period 15 also have low prices in equilibrium. These markets may typically require more time to reach equilibrium and be more likely to have frustrated subjects wanting to leave the experimental market. If so copies 3, 6, and 14 would be providing a bias to the observed distribution of the truncated series relative to the equilibrium distribution, where they were not included. If this data is discarded, the null hypothesis is rejected at a 5.5% significance level.

9. This is similar to the two-arm bandit problem considered in Rothschild (1974).

References

Alger, D. (1984), "Equilibria and Rationalizable Strategic Behavior," mimeo.

Ketcham, J., V. L. Smith, and A. Williams (1984), "A Comparison of Posted-Offer and Double-Auction Pricing Institutions", Review of Economic Studies, 51(4), 595614.

Fouraker, L. and S. Siegel (1963), Bargaining Behavior, Greenwood Press, 167-199.

Isaac, M., V. Ramey, and A. Williams (1984), "The Effects of Market Organization on Conspiracies in Restraint of Trade," Journal of Economic Behavior and Organization, 1-32.

Rothschild, M. (1974), "A Two-Armed Bandit Theory of Market Pricing," Journal of Economic Theory, 185-202.

Smith, V. L. (1976), "Induced Value Theory," American Economic Review, 66(2), 274-279.