THE JOURNAL OF FINANCE • VOL. XXXIX, NO. 4 • SEPTEMBER 1984

Theory and Behavior of Multiple Unit Discriminative Auctions

JAMES C. COX, VERNON L. SMITH, and JAMES M. WALKER*

ABSTRACT

This paper reports the results of controlled experiments designed to test the Harris-Raviv generalization of the Vickrey theory of bidding in multiple unit discriminative auctions. The paper also discusses further development of the theory—in a way suggested by the experimental results—to include bidders with distinct risk preferences.

IN A WELL-KNOWN PAPER, Vickrey [20] formulated a Nash equilibrium model of bidding by risk neutral economic agents in single unit auctions. This analysis was subsequently extended in numerous papers. Vickrey [21] generalized his original model to include multiple unit auctions in which each of N risk neutral bidders can bid on one out of a total of Q homogeneous items up for auction, where $1 \leq Q < N$. In both of the Vickrey papers, individual values for the auctioned object(s) were assumed to be drawn from a uniform distribution. Holt [12] and Riley and Samuelson [19] for single unit auctions and Harris and Raviv [11] for multiple unit auctions have extended the Vickrey model to the case in which valuations are from a general distribution function and all agents have identical concave utility functions.[1]

This paper reports the empirical properties of individual bidding behavior and seller revenue for a group of 28 laboratory experiments designed to test the Harris-Raviv generalization of the Vickrey (hereafter, VHR) model of Nash equilibrium behavior in multiple unit discriminative auctions. In Section I, we summarize briefly the theoretical results of the VHR model which form the basis for the hypotheses that we test. Section II describes the experimental design that we use. Section III describes the experimental results and various tests based on the bidding behavior of individual subjects and on aggregate market (seller revenue) data. The results from six of the ten (N, Q) parameter designs reported

* All authors from the Department of Economics, University of Arizona. We are grateful to the National Science Foundation for research support and to an associate editor of this *Journal* for helpful comments and suggestions.

[1] Other important extensions of Vickrey's original work have been provided by Matthews [16], Maskin and Riley [15], Myerson [18], Milgrom and Weber [17], and others. Since these extensions have involved institutions with information or valuation conditions that have not yet been examined experimentally, they are not directly germane to the objectives of this paper. For example, Matthews [16] has modeled single unit auctions in which the seller selects a reservation price, whereas in all of the experimental studies to date the seller passively offers the item(s) for sale. Similarly, Milgrom and Weber [17] have extended auction theory to allow valuations to be positively dependent for risk neutral bidders.

in Section III are consistent with VHR predictions of seller revenue and levels of individual bids, but not with the implication that bidders have the same bid function. In another paper (see [7], which is further generalized in [5]), we offer a parametric theory of Nash equilibrium bidding behavior in which individuals are assumed to exhibit differing degrees of constant relative risk aversion. This theory "accounts" for the results from six of the ten (N, Q) parameter designs. However, neither our parametric model nor that of VHR is consistent with the observed bids (and resulting revenues) for four of the (N, Q) parameter designs. In the present paper, this evidence is remarkably stable, replicable, and internally consistent, as is indicated by the various procedural and internal validity tests that we describe in Section IV.

I. The Vickrey-Harris-Raviv Theory of Nash Bidding Behavior

Let $Q \geq 1$ units of a homogeneous good be offered in perfectly inelastic supply to $N > Q$ bidders. Each bidder submits a bid for a single unit with the understanding that each of the Q highest bidders will be awarded a unit of the good at a price equal to his bid; i.e., the institution is a discriminative sealed-bid auction. Let v_i be the monetary value of a unit of the good to bidder i, where $i = 1, 2, \cdots, N$. Assume that each v_i is drawn (with replacement) from a distribution with density $h(\cdot)$ and probability distribution function $H(\cdot)$ whose support is the interval $[0, \bar{v}]$. If bidder i bids b_i, and this bid is accepted, he gains the monetary surplus, $v_i - b_i$, with utility $u(v_i - b_i)$. If b_i is *not* accepted, bidder i's monetary gain is zero. Assume that $u(\cdot)$ is increasing, concave, and differentiable and that $u(0) = 0$. Suppose that bidder i expects each of his rivals to bid according to the differentiable bid function

$$b_j = b(v_j) \tag{1}$$

Let $b(\cdot)$ be increasing on $[0, \bar{v}]$ and denote by π the inverse of (1); i.e.,

$$\pi(b(v_j)) = v_j \tag{2}$$

The probability, $\psi(b)$, that any one of the $N - 1$ rivals of bidder i will bid less than some amount b in the range of (1) is the same as the probability, $H(\pi(b))$, that the rival will draw a value less than that v corresponding to b given by $\pi(b)$. Hence, the probability that a bid b by i will win is the probability, $G(\pi(b))$, that at least $N - Q$ of the values drawn by i's rivals are less than $\pi(b)$. This probability is given by the distribution function of the $(N - Q)$th order statistic for a sample of size $N - 1$ from the distribution H:

$$G(\pi(b)) = \frac{(N-1)!}{(N-Q-1)!(Q-1)!} \int_0^{\pi(b)} [H(v)]^{N-Q-1}[1 - H(v)]^{Q-1} h(v)\, dv \tag{3}$$

Now let bidder i choose $b_i = b_i^0$ to maximize $u(v_i - b_i)G(\pi(b_i))$. The function $b(\cdot)$ is a Nash equilibrium bid function only if $b_i^0 = b(v_i)$. Let $b_n(\cdot)$ be the special case of (1) where all bidders are known to be risk neutral and let $b_a(\cdot)$ be the special case of (1) where all bidders are known to have the same strictly concave utility function. Harris and Raviv [11] prove the following result for these Nash equilibrium risk neutral and risk averse bid functions:

$$b_n(v) = \frac{1}{G(v)} \int_0^v x\, dG(x) < b_a(v) < v \quad \text{for all} \quad v \in (0, \bar{v}] \tag{4}$$

If the density function $h(\cdot)$ is the constant density $(\bar{v})^{-1}$, the result (4) can be expressed in terms of the following Beta integrals, which for computational purposes can be written in terms of the indicated finite polynomials (see [6]):

$$b_n(v) = \frac{\bar{v} \int_0^{v/\bar{v}} Y^{N-Q}(1 - Y)^{Q-1}\, dY}{\int_0^{v/\bar{v}} Y^{N-Q-1}(1 - Y)^{Q-1}\, dY}$$

$$= \frac{\bar{v} \sum_{k=0}^{Q-1} \dfrac{(-1)^k (v/\bar{v})^{N-Q+k+1}(Q-1)!}{(N - Q + k + 1)k!(Q - 1 - k)!}}{\sum_{k=0}^{Q-1} \dfrac{(-1)^k (v/\bar{v})^{N-Q+k}(Q-1)!}{(N - Q + k)k!(Q - 1 - k)!}} < b_a(v) \tag{5}$$

In Figure 1, we plot $b_n(\cdot)$ for the parameter values, $\bar{v} = \$2.24$, $N = 10$, and $Q = 4$. Note that the risk neutral bid function defines the boundary below which any individual's observed bids are inconsistent with the VHR theory, while bids equal to or greater than $b_n(v_i)$ but less than v_i are consistent with the theory. Thus, in Section III, if b_i^* is agent i's observed bid corresponding to the observed value v_i^*, then we test the null hypothesis that $\Pr[b_i^* < b_n(v_i^*)] \geq 1/2$ for each i. The VHR theory will be called into question, if not falsified, if we are unable to reject this null hypothesis.[2]

However, rejection of this null hypothesis is insufficient to allow us to conclude that (with high probability) VHR stands as a nonfalsified theory of bidding behavior in discriminative auctions. This is because each subject's bids might be consistent with the inequality (5) but be inconsistent with the VHR theory's implication that all subjects bid according to the same equilibrium bid function. (This property of the theory implies that the allocations will be Pareto optimal.) Accordingly, in Section III, we report a direct test of the null hypothesis that individuals are homogeneous in their bidding behavior.

Typically in economics, when one tests the implications of a theory, the data are at best macromarket observations based on aggregation rather than observations on individual agents. Since a great variety of theories of individual behavior can yield qualitatively equivalent aggregate implications, such tests are extremely weak.[3] We have both individual and macromarket observations from

[2] Our test situation corresponds to the usual one in economics wherein there is only one formal theory. In this situation the classical falsification methodology does not articulate guidelines. The test procedure we use in Section III adopts a "naive" alternative theory to VHR, in which it is assumed that individual bids are at least as likely to be below the risk neutral bid function as above it. This naive theory becomes the null alternative to VHR, which states that bids are less likely to be below the risk neutral boundary.

[3] Macromarket tests are weak in the sense that if the observations fail to falsify the theory this

Footnote 3 continues on next page.

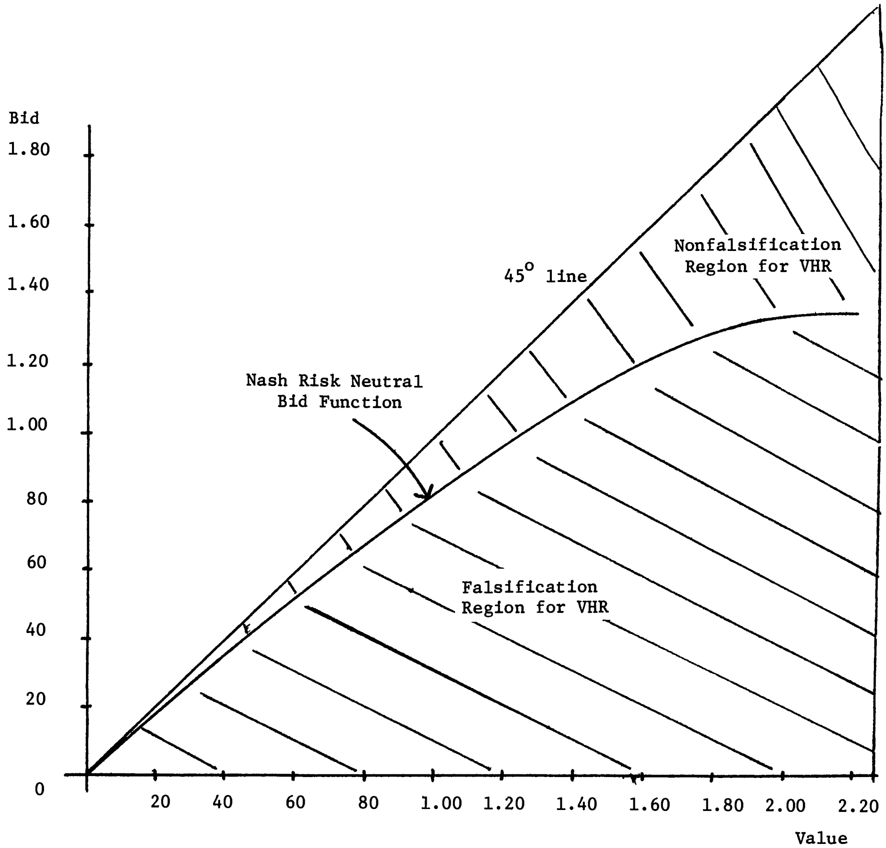

Figure 1. Critical Regions for VHR Bidding Theory for $(\bar{v}, N, Q) = (2.24, 10, 4)$

our laboratory experiments and, for purposes of summary comparisons, we also provide tests of the macromarket implications of the VHR theory.

From the ordering theorem in (4), it follows that seller revenue for risk neutral bidders (R_n) is less than seller revenue for risk averse bidders (R_a); i.e., if bidders are indexed so that $b_1 \leq b_2 \leq \cdots \leq b_N$, then

$$R_n = \textstyle\sum_{j=N-Q+1}^{N} b_n(v_j) < R_a = \sum_{j=N-Q+1}^{N} b_a(v_j) \tag{6}$$

Consequently, in Section III, we report tests of the macromarket null hypothesis that

$$R_n^* = \textstyle\sum_{j=N-Q+1}^{N} b_n(v_j^*) > R^* = \sum_{j=N-Q+1}^{N} b_j^* \tag{7}$$

where R^* is observed seller revenue and R_n^* is the theoretical risk neutral seller revenue conditional on the observed values, v_j^*, from a particular experimental realization. Reports of the seller revenue results will provide an overview of the

provides weaker support for the theory than if similar results had been obtained from tests of indivdual behavior. However, if the aggregate observations are inconsistent with the theory, this result is sufficient to falsify the theory.

experiments across all the parameter values studied, and allow us conveniently to integrate into the discussion the results of a large number of single unit auctions reported previously [4, 8].

II. Experimental Design

In a previous paper [6], we reported the results of 16 multiple unit auction experiments designed to make paired comparisons of the seller revenue properties of discriminative and uniform price sealed-bid auctions. Bidding data for the discriminative auction treatment in eight of these experiments (those for which the subjects were experienced) are included in this paper along with the results of 20 new discriminative auction experiments using experienced subjects. (By "experienced" we mean a subject who has been in a previous sealed-bid auction experiment of the type reported here.)

The complete experimental instructions given to all subjects are presented in [6]. Due to the length of the instructions, we will summarize here only the key points developed in the instructions and other significant experimental design features.

The experiments reported in this study were conducted using subjects drawn from a population of undergraduate students at the University of Arizona. The students were currently enrolled in lower level economics courses. All subjects were volunteers who had received a brief explanation that in our economics experiments the earnings resulting from their decisions would be paid to them in cash.

All experiments were conducted using the PLATO computer system. This approach to laboratory experiments enables us to ensure that all subjects are given identical market instructions with minimal experimenter/subject interaction. The computer system also facilitates the accounting process needed for each auction as well as curtailing subject transaction costs in bidding and obtaining information. No communication was allowed between subjects except that specified in the bidding process under 4 below.

Some important characteristics of the experiments can be summarized as follows.

1. Upon arriving at the PLATO computer site, each subject received $3.00 for keeping his/her appointment. The experimenter then assigned the subjects randomly to computer terminals at which time each subject reviewed, at his/her own pace, a set of instructions explaining the types of choices that subjects would encounter in the experiment.

2. In the instructions, subjects were informed (primarily with examples) of key properties of the discriminative auction institution in which they would be participating. Monetary values for the abstract objects being auctioned were induced on subjects using the concept of "resale values." That is, it was explained to each subject that for any winning bid the bidder would receive a profit equal to his/her resale value minus the amount bid. Each subject had one bid per auction and was informed in advance of each auction of his/her resale value for that auction.

3. Each bidder was informed (in nontechnical language) that all resale values assigned to each individual were drawn (with replacement) from a discrete uniform distribution contained in the interval $[0, \bar{v}]$. As required by the theory, subjects were informed of their own (certain) valuations for a specific auction, but were only informed of the distribution from which other subjects' values were drawn. Subjects also were informed that values were drawn independently for each auction.

4. At the end of each auction period, each subject was provided with a table of information on the outcome of that period's auction. In experiments $2'$, $3'$, and $6'$, the only information in the table was the subject's own bid, resale value, and profit in that auction. In all of the other experiments, the table displayed the highest accepted bid and highest rejected bid in that auction as well as the subject's own bid, resale value, and profit for that auction.[4] At the subject's discretion a table was provided which summarized the available information for all preceding auctions.

5. The only restriction on a subject's bids was that they had to be in the interval $[0, \bar{v}]$. Hence, a subject could bid in excess of his/her value, which would imply a loss if the bid was accepted. An occasional subject in the inexperienced groups bid in excess of value on more than a single occasion (see [6], Table 6). All such subjects were screened out of further experiments; i.e., they were excluded from the "experienced" subject pool. In this sense all "experienced" subjects were individually rational as well as having "practiced" familiarity with the discriminative auction institution.

6. Before making any bids in the experiment, subjects were reminded of pertinent summary market parameters such as number of bidders, number of units offered, and maximum and minimum possible resale values. Subjects were also endowed with $1.00 in "working capital." This was to compensate for any loss resulting if an individual were to bid in excess of his/her value and have the bid accepted. Subjects were informed in advance that there would be many auctions but were not informed of the specific number. After a predetermined number of auctions, subjects were informed that the experiment had ended, and were then payed privately the full amount of their earnings. Experimental sessions lasted approximately one and a half to two hours.

7. The expected gain to a risk neutral bidder can be calculated (see [6], Equation (3.6)) to yield $[(Q(Q + 1)\bar{v})/(2N(N + 1))]$. In order to provide some uniformity of subject motivation across experiments with different values of (N, Q), we chose $\bar{v}$ as a function of (N, Q) so that the expected salient reward for a VHR risk neutral bidder would be approximately the same ($10) in all experiments except the enhanced payoff experiments (subjects also receive $3 up front for keeping the session appointment).[5] The specific design parameters which were used are shown in Table I.

[4] The highest accepted bid and highest rejected bid are revealed (after the auction) to bidders in the primary auction for U.S. Treasury bills.

[5] Exercising this control over the general level of payoffs across treatments has both empirical and theoretical justifications that are particularly important to understand. Empirically, this control serves to immunize against the risk that, *a posteriori*, a critic may correctly argue that the interpretation of deviant results (if they occur) are confounded: The results may be accounted for by variation in payoff levels, and are not necessarily attributable to the scheduled "treatment." Also it should be

Table I

Design Parameters

Design Number	Maximum Resale Value, $\bar{v}$	Number of Bidders, N	Units Offered, Q	$E(R_n)$[d]	Number of Replications (Experiments)	Number of Discriminative Auctions per Experiment[e]
1	$0.80	10	7	$1.53	4	20
2	$2.24[a]	10	4	$4.89	6	20
2′	$2.24[a]	10	4	$4.89	2	30
2″	$6.72[a]	10	4	$14.66	1	30
3	$2.70[b]	5	2	$2.70	2	30
3′	$2.70[b]	5	2	$2.70	2	30
4	$3.78	6	2	$4.32	2	30
5	$5.04	7	2	$6.30	1	30
6	$6.48[c]	8	2	$8.64	2	30
6′	$6.48[c]	8	2	$8.64	2	30
7	$9.90	10	2	$14.40	1	30
8	$2.52	7	3	$3.78	1	30
9	$3.24	8	3	$5.40	1	30
10	$4.05	9	3	$7.29	1	30

[a] In one experiment (2DE2*) using this design, the payoffs were tripled. In two others (2′DE1BLK; 2′DE2BLK), the highest rejected and highest accepted bid information was blocked. In another (2″DE2), $\bar{v}$ was tripled.

[b] In two experiments (3′DE1BLK; 3′DE2BLK) using this design, the highest rejected and highest accepted bid information was blocked.

[c] In two experiments (6′DE1BLK; 6′DE2BLK) using this design, the highest rejected and highest accepted bid information was blocked.

[d] The calculation of theoretical risk neutral revenue is based on the expression derived in Cox, Smith, and Walker ([6], Statement 4.1)),

$$E(R_n) = \frac{Q(N-Q)\bar{v}}{N+1} < E(R_a)$$

[e] Designs 1 and 2 incorporated a sequencing of auctions that included a uniform price auction. (See [6]). However, in one experiment (2DE1) we replicated the earlier design 2 in a 30 discriminative auction sequence. (Actually designs 1 and 2 consisted of 23 auctions, and designs 3–10 of 33 auctions, but we omitted the first three auctions in all of our analyses to control for "start-up" variability).

III. Experimental Results

A. Levels of Individuals' Bids

From the theory of individual bidding behavior presented in Section I, the following null hypothesis can be proposed.

$$H_0^B\colon b_i^* < b_n(v_i^*) \quad \text{for all} \quad i \tag{8}$$

---

noted that the theory itself predicts that the expected gain becomes negligible for given Q and large enough N. Hence, for given Q, as N increases the "demand" parameter $\bar{v}$ must be expanded so as not to alter the relative importance of each bidder. This phenomenon is of course recognized in general equilibrium theory when, in examining large economies, the size of the economy is increased in the same proportion as the number of agents (in each characteristic class), so as not to bias the effect of increases in the number of agents. In Section IV we address the empirical and theoretical implications of the possibility that $\bar{v}$ or the reward level may itself be a "treatment" in accounting for the experimental outcomes.

Table II

Individual Bidding Results

Design	Number of Subjects	Observations per Subject	A	B (Percent)	C (Percent)
1	40	20	2	11 (28)	29 (72)
2	60	20[a]	11	31 (52)	29 (48)
2′	20	30	2	8 (40)	12 (60)
2″	10	30	1	3 (30)	7 (70)
3	10	30	10	10 (100)	0 (0)
3′	10	30	10	10 (100)	0 (0)
4	12	30	11	11 (72)	1 (8)
5	7	30	4	7 (100)	0 (0)
6	16	30	14	16 (100)	0 (0)
6′	16	30	16	16 (100)	0 (0)
7	10	30	7	10 (100)	0 (0)
8	7	30	1	6 (86)	1 (14)
9	8	30	1	5 (63)	3 (37)
10	9	30	6	9 (100)	0 (0)

Note: A, Subjects for whom H_0^B is rejected ($p = 0.05$). B, Subjects for whom sign of test statistic is consistent with VHR. C, Subjects for whom sign of test statistic is inconsistent with VHR.

[a] In experiments 2DE1 and 2DE2* there were 30 observations per subject.

Our test procedure for H_0^B is to apply the nonparametric Wilcoxon signed-ranks test [10] to the hypothesis that $\Pr[b_i^* - b_n(v_i^*) < 0] \geq 1/2$. This is a paired sample test; thus, for each bid b_i^* corresponding to the value v_i^* for individual i, there is a paired risk neutral theoretical bid $b_n(v_i^*)$ computed from (5). The Wilcoxon procedure does not require us to impose any assumptions on the distribution of the difference, $b_i^* - b_n(v_i^*)$.

The results of the Wilcoxon test for individual subjects are summarized in Table II. At this level of analysis, we begin to see the previously mentioned dichotomy between those experimental designs which support the theory and those which do not. In design 1, 29 (72%) of 40 subjects submitted bids that are too low to be consistent with VHR theory, while in design 2, 29 (48%) of 60 subjects were inconsistent with the theoretical predictions (see Column 6, Table II). In design 2′, where the highest rejected and highest accepted bid information was blocked (i.e., not revealed to the subjects), 12 (60%) of 20 subjects submitted bids that are too low to be consistent with VHR theory. Also, in design 2″, where $\bar{v}$ had been tripled to increase expected subject profits, seven (70%) of 10 subjects bid too low to be consistent with the theory, thus suggesting no effect on the level of bids from tripling the range of payoffs. Of the eight subjects in design 9, three (37%) bid at a level too low to be consistent with VHR. In designs 4 and 8, we observe only one bidder in each design bidding too low relative to the theory. However, in design 4, we find 11 of 12 bidders bidding significantly above the risk neutral predicted level, while in design 8 we find only one of seven does so. In contrast to these results, we find 0 of the 78 subjects in designs 3, 3′, 5, 6, 6′, 7, and 10 submitting bids at a level inconsistent with the predictions of the VHR theory. In summary, at the individual bidding level, VHR is strongly supported in designs 3, 3′, 5, 6, 6′, 7, and 10; it is weakly supported in designs 4, 8, and 9; and it fails decisively to receive support in designs 1, 2, 2′, and 2″.

It is particularly instructive to examine the bidding patterns of one or two representative subjects from some of the experimental designs. Figures 2–10 show plots of individual bids at corresponding assigned values for nine subjects. The risk neutral bid function, computed from the polynomial expression in (5) for the defining experimental parameters, is shown on the same scale for each bidder. Figures 2–10 all exhibit a nonlinear bid-value relationship which resembles the graphed concave form of the risk neutral VHR bid function. In many cases, as in Figures 5 and 7–10, the individual bids form a very tight nonlinear relationship. However, some individuals differ markedly in their bid-value behavior for the same experimental design parameters. This is illustrated by comparing Figures 2 and 3 for design 1 and Figures 4 and 5 for design 2. Figures 2 and 4 are typical of the bidding pattern for the great majority of bidders in the design 1 and 2 experiments. Initially the bids increase with value approximately as does the risk neutral equilibrium bid function; then they tend to cross under and remain below this bid function. Consequently, at the lower values the bids are relatively consistent with VHR equilibrium bidding, but at the higher values the bids are inconsistent with any form of the VHR noncooperative model. In contrast, as seen in Figures 6–10, individual bidders are quite consistent with the risk averse form of the VHR model over the entire range of values sampled in designs 3–7.

B. Bidder Heterogeneity

Although we find the level of individual bids strongly consistent with the VHR theory for designs 3, 3′, 5, 6, 6′, 7, and 10, and weakly consistent for designs 4,

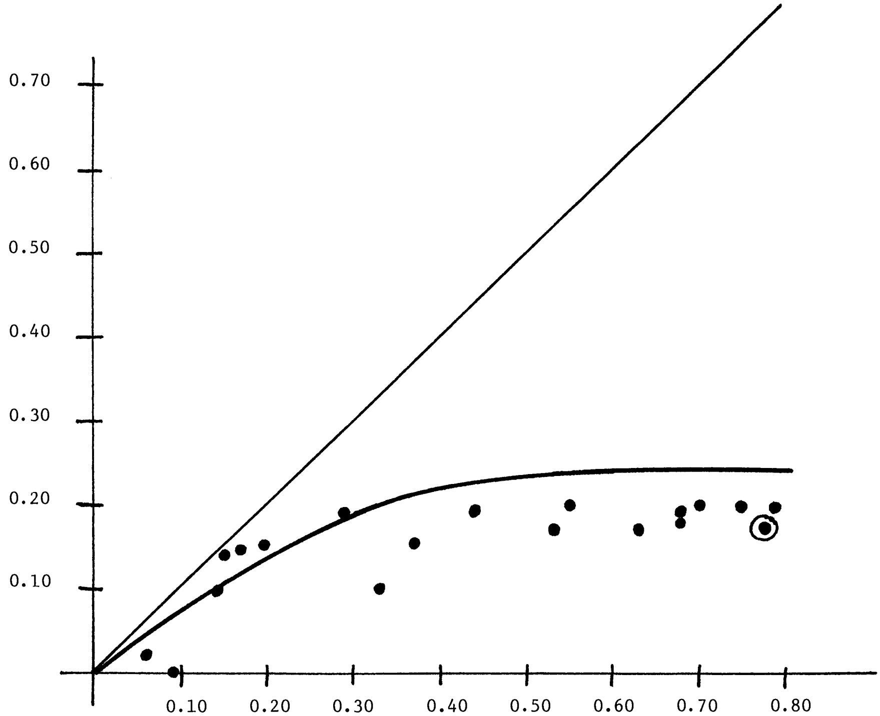

Figure 2. Bids, Subject 1, 1UDE2, $(\bar{v}, N, Q) = (0.80, 10, 7)$

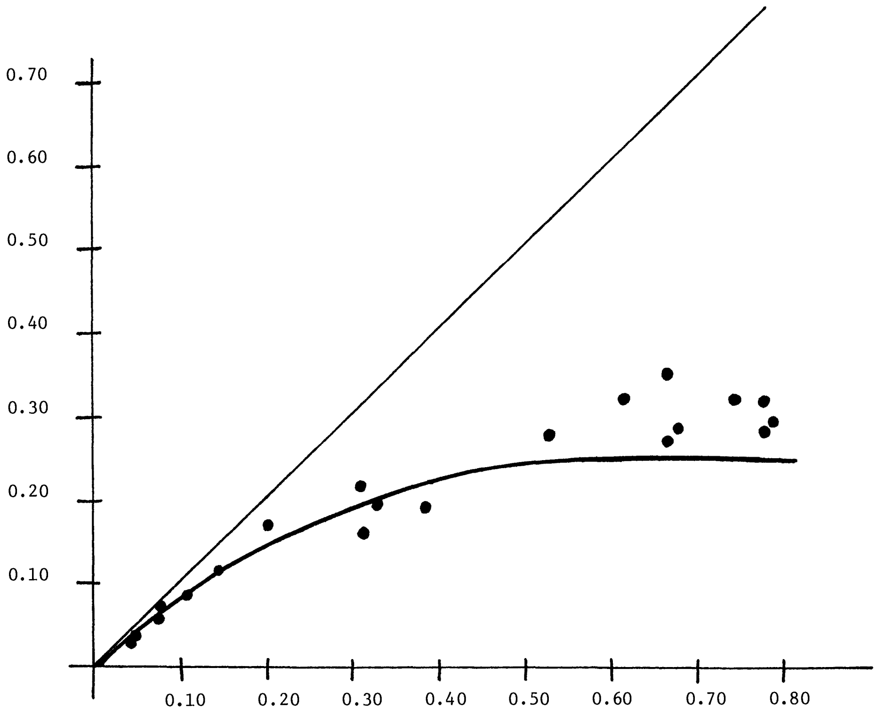

Figure 3. Bids, Subject 3, 1DUE1, $(\bar{v}, N, Q) = (0.80, 10, 7)$

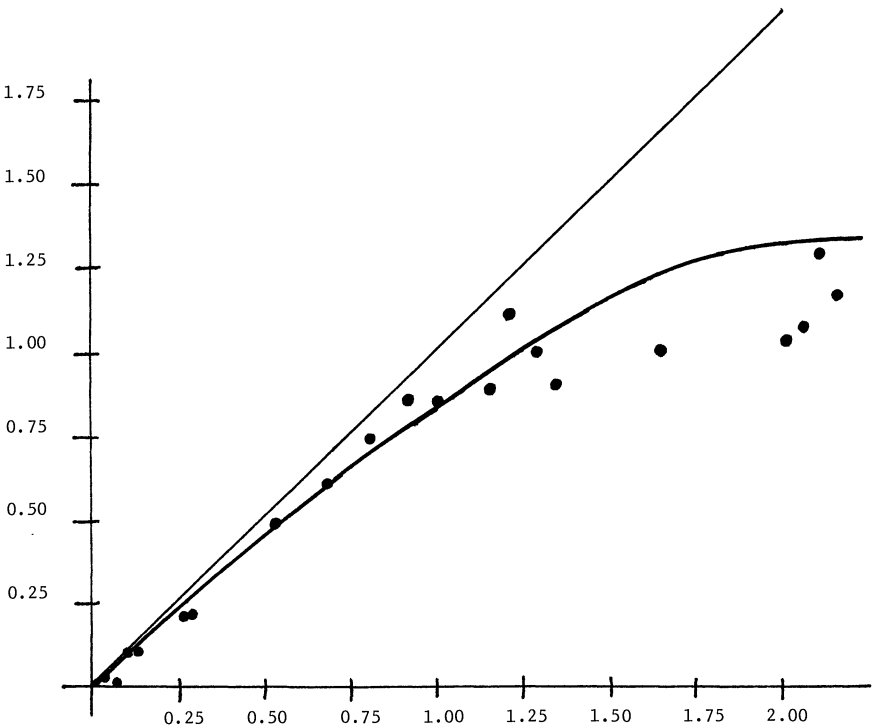

Figure 4. Bids, Subject 2, 2UDE1, $(\bar{v}, N, Q) = (2.24, 10, 4)$

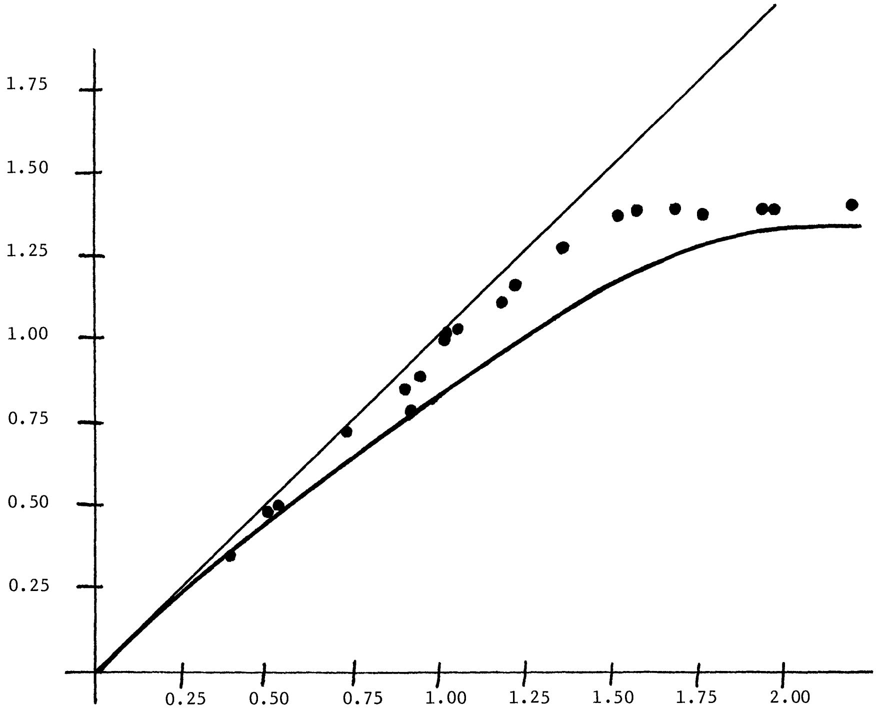

Figure 5. Bids, Subject 5, 2DUE1, ($\bar{v}$, N, Q) = (2.24, 10, 4)

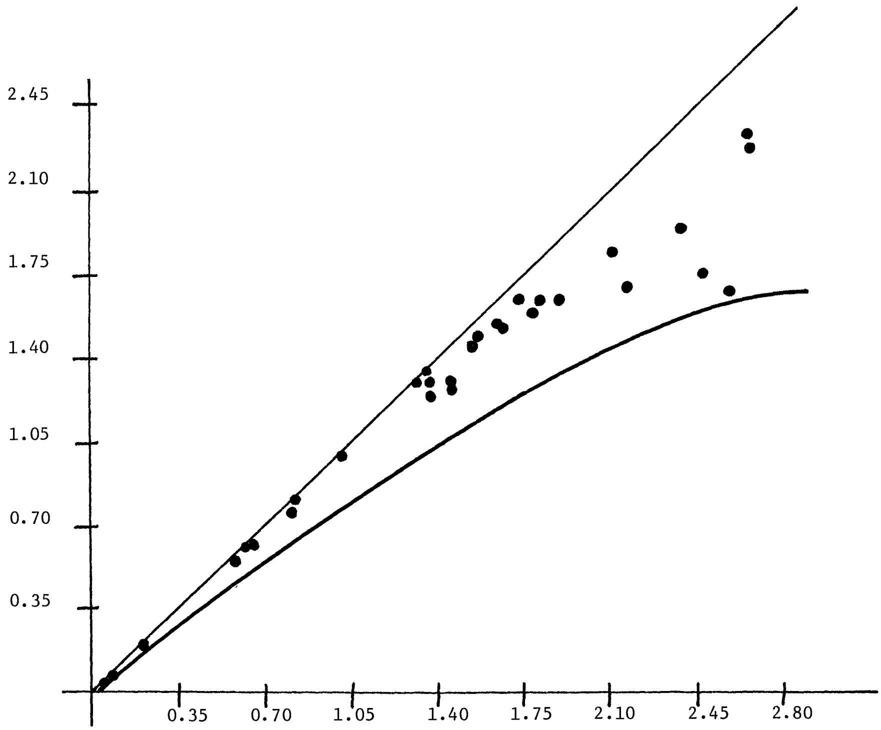

Figure 6. Bids, Subject 5, 3DE1, ($\bar{v}$, N, Q) = (2.70, 5, 2)

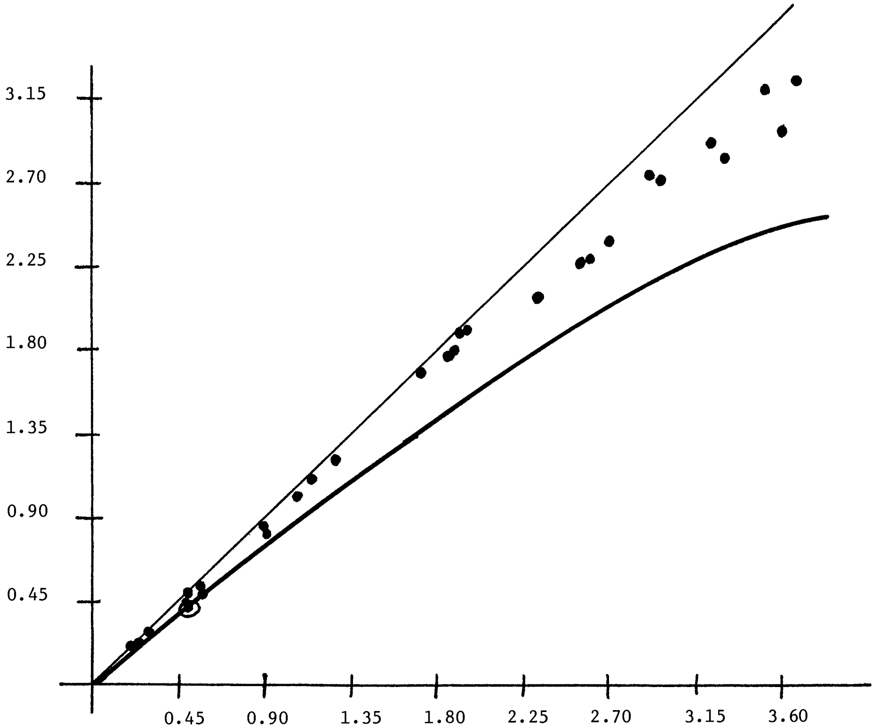

Figure 7. Bids, Subject 1, 4DE1, $(\bar{v}, N, Q) = (3.78, 6, 2)$

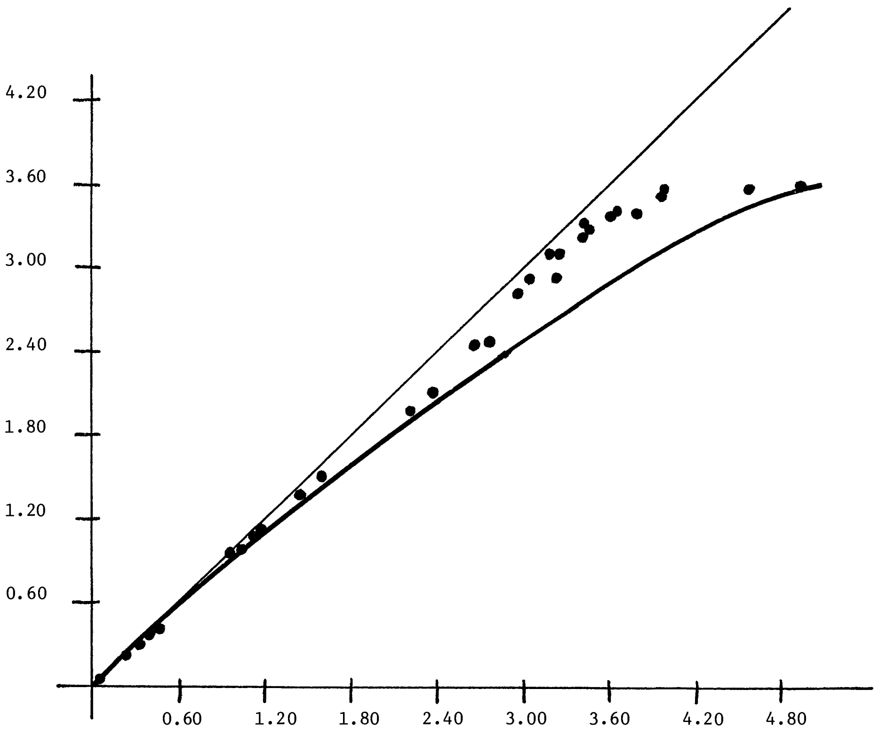

Figure 8. Bids, Subject 4, 5DE1, $(\bar{v}, N, Q) = (5.04, 7, 2)$

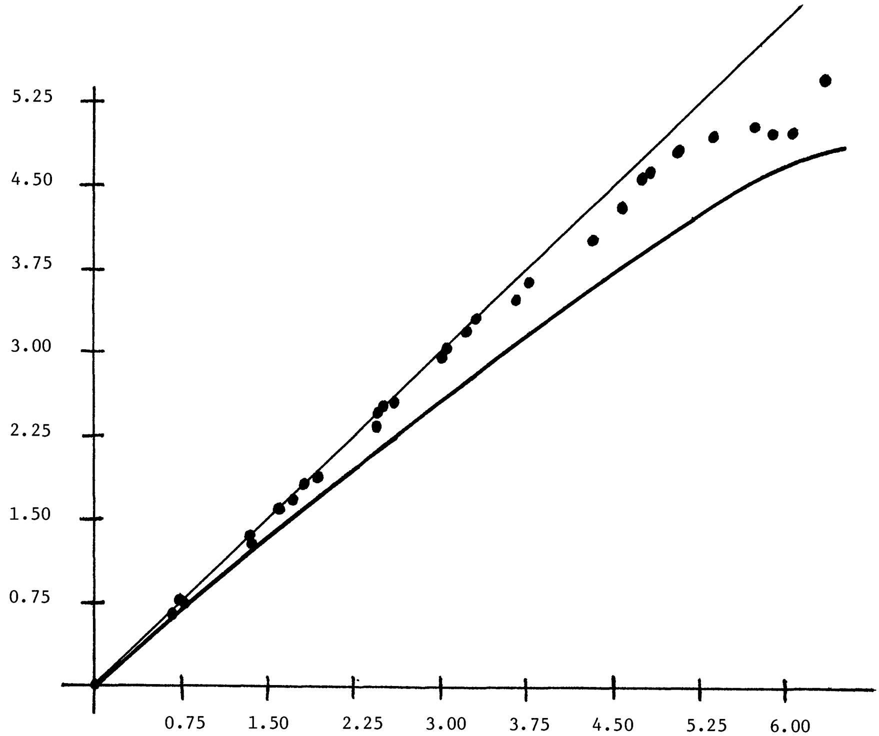

Figure 9. Bids, Subject 2, 6DE1, $(\bar{v}, N, Q) = (6.48, 8, 2)$

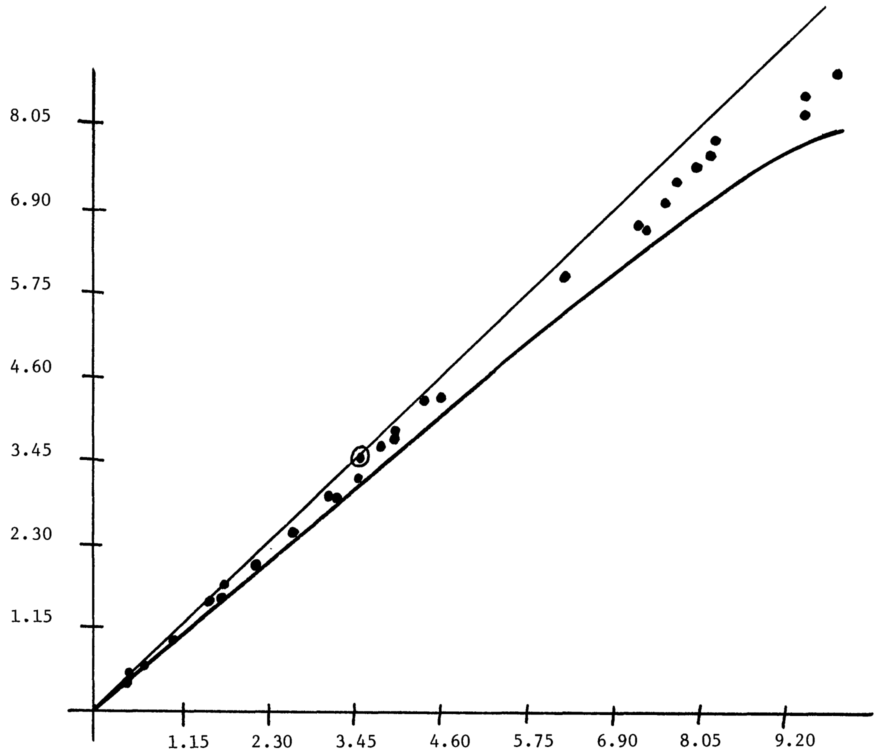

Figure 10. Bids, Subject 1, 7DE1, $(\bar{v}, N, Q) = (9.90, 10, 2)$

Table III

H-Test of Heterogeneity of Individual Bids

Session	Q	N	H-Statistic for $b_i^* - b_n(v_i^*)$
1UDE1	7	10	17.3[a]
1UDE2	7	10	30.9[a]
1DUE1	7	10	20.6[a]
1DUE2	7	10	39.2[a]
2UDE1	4	10	14.5
2UDE2	4	10	44.8[a]
2DUE1	4	10	31.2[a]
2DUE2	4	10	33.9[a]
2DE1	4	10	11.3
2DE2*	4	10	39.4[a]
2′DE1BLK	4	10	22.7[a]
2′DE2BLK	4	10	27.4[a]
2″DE2	4	10	17.5[a]
3DE1	2	5	34.5[a]
3DE2	2	5	2.7
3′DE1BLK	2	5	29.3[a]
3′DE2BLK	2	5	27.0[a]
4DE1	2	6	23.3[a]
4DE2	2	6	28.7[a]
5DE1	2	7	42.9[a]
6DE1	2	8	16.8[a]
6DE2	2	8	51.2[a]
6′DE1BLK	2	8	15.1[a]
6′DE2BLK	2	8	19.8[a]
7DE1	2	10	51.3[a]
8DE1	3	7	41.3[a]
9DE1	3	8	13.9
10DE1	3	9	22.5[a]

[a] Significant, $p = 0.05$.

8, and 9, such results are insufficient to conclude that VHR theory is supported (nonfalsified) in these designs. This is because the bidding behavior of individuals may depart from the common equilibrium bid function implied by the VHR model.

A direct statistical test of the null hypothesis, that the bids of individuals in each experiment are from the same distribution as opposed to the alternative that they are from different distributions, is reported in Table III. The first three columns of Table III identify the treatment conditions for each experiment. Thus, the experiment designated 1UDE1 indicates the following. The first "1" denotes that design 1 parameters of Table I were used. "U" indicates that the first sequences of auctions were conducted under uniform price rules and the following "D" indicates the next sequence of auctions were under discriminative price rules. The "E" represents experienced subjects and the final "1" indicates this was the first experiment of this type. One should note that the experiments using designs 3–10 and experiments 2DE1, 2DE2*, 2′DE1BLK, 2′DE2BLK, and 2″DE2 consist only of discriminative price auctions. Reported in Table III is the Kruskal-Wallis [13] nonparametric H-test applied to the deviations of individual bids from the predicted risk neutral bids, $b_i^* - b_n(v_i^*)$, for each subject in each experiment. Across all designs, in 24 of 28 experiments the null hypothesis is rejected. We conclude from these tests that individual bids are inconsistent with the VHR theory's implication that all bidders use the same bid function.[6]

C. *Revenue Results*

Even though the emphasis of this paper is the analysis of individual bidding behavior, it is also instructive to consider the impact of the previously summarized individual results on macro level revenue data for these sealed-bid auctions. For example, suppose all unaccepted but only some accepted individual bids were consistent with VHR theory. Then the tests on the individual data might show consistency with VHR, while the revenue data, based on the accepted bids only, might be inconsistent with VHR. Alternatively, if all unaccepted but only some accepted bids were inconsistent with VHR, the tests on the individual data could show inconsistency with VHR, while the revenue data might be consistent with VHR. Hence, tests on the revenue data are capable of yielding results that differ from the tests on individual bids.

A summary of revenue results from the discriminative auctions of all experiments using experienced subjects is given in Table IV. The fourth column of Table IV reports the observed mean (variance) of discriminative auction revenues for each experiment. In the fifth column appears the difference between the observed mean and the VHR predicted revenue (and its variance) in the discriminative auctions under the assumption of risk neutrality. From the VHR theory of bidding, the following null hypothesis can be proposed:

$$H_0^R\colon R^* - R_n^* = \textstyle\sum_{j=N-Q+1}^{N} b_j^* - \sum_{j=N-Q+1}^{N} b_n(v_j^*) < 0 \tag{9}$$

If the distribution of $R^* - R_n^*$ in successive auctions is approximately normal, and we observe an experimental sample of size T with mean $\bar{R}^* - \bar{R}_n^*$ and standard deviation S_δ, then a one-tailed paired sample t-test based on

[6] Given the traditional emphasis of economics on postulated differences in individual preferences, it is tempting, *ex post*, to be "not surprised" upon learning that in most of our groups the bidders do not behave as if they all use the same bid function. But when a theorist postulates homogeneous bidders, we assume that this postulate and the resulting theory is to be taken seriously as an approximation that could well be consistent with observation. If our subjects had not been such consistent bidders (in particular see Figures 3, 5, 7, 8, and 9) and had exhibited more noise in their empirical bid functions, then the null hypothesis could easily have failed to be rejected in most experiments. The conclusion would then have been that in terms of bidding behavior based on random value assignments, bidders are indeed homogeneous. We think that we would have attached a nonnegligible prior positive probability to this prospect. Once one has seen the data, it is of course impossible to assign credible prior probabilities. The fact that in four of the 28 experimental groups (14%) reported in Table III the null hypothesis could not be rejected suggests that the homogeneous bidders assumption is not meritless. Some subject samples do have this property. (We would expect to reject the null hypothesis in 5% of the samples even if it were true.)

Table IV

Comparison of Observed and VHR Revenues for Risk Neutral Bidders

Session	Q	N	$\bar{R}^*$ (V_R^*)	$\bar{R}^* - \bar{R}_n^*$ (V_δ)	T	Paired Sample t-Statistic
1UDE1	7	10	1.282 (0.029)	−0.254 (0.032)	20	−6.311
1DUE1	7	10	1.444 (0.053)	−0.070 (0.037)	20	−1.630
1UDE2	7	10	1.375 (0.018)	−0.162 (0.010)	20	−7.103
1DUE2	7	10	1.484 (0.017)	−0.031 (0.008)	20	−1.575
2UDE1	4	10	4.695 (0.386)	+0.002 (0.182)	20	+0.021
2DUE1	4	10	4.467 (0.095)	−0.347 (0.150)	20	−4.000
2UDE2	4	10	4.216 (0.145)	−0.477 (0.124)	20	−6.060
2DUE2	4	10	4.592 (0.126)	−0.222 (0.149)	20	−2.572
2DE1	4	10	4.593 (0.132)	−0.142 (0.088)	30	−2.619
2DE2*	4	10	4.499 (0.138)	−0.226 (0.092)	30	−4.094
2′DE1BLK	4	10	4.310 (0.114)	−0.428 (0.116)	30	−6.092
2′DE2BLK	4	10	4.698 (0.134)	−0.048 (0.066)	30	−1.018
2″DE2	4	10	12.588 (2.734)	−1.651 (3.534)	30	−4.810
3DE1	2	5	3.204 (0.149)	+0.456 (0.038)	30	+12.808
3DE2	2	5	3.134 (0.093)	+0.386 (0.038)	30	+10.787
3′DE1BLK	2	5	3.138 (0.176)	+0.378 (0.105)	30	+6.390
3′DE2BLK	2	5	2.950 (0.068)	+0.189 (0.070)	30	+3.980
4DE1	2	6	4.932 (0.507)	+0.518 (0.210)	30	+6.195
4DE2	2	6	4.552 (0.282)	+0.137 (0.095)	30	+2.436
5DE1	2	7	6.893 (0.317)	+0.418 (0.075)	30	+8.356
6DE1	2	8	9.036 (0.669)	+0.447 (0.162)	30	+6.090
6DE2	2	8	9.314 (0.127)	+0.725 (0.206)	30	+8.747
6′DE1BLK	2	8	9.668 (1.474)	+1.082 (0.155)	30	+15.072
6′DE2BLK	2	8	9.108 (0.757)	+0.522 (0.216)	30	+6.155
7DE1	2	10	14.910 (0.759)	+0.512 (0.430)	30	+4.275
8DE1	3	7	3.680 (0.106)	−0.181 (0.120)	30	−2.860
9DE1	3	8	5.162 (0.129)	−0.234 (0.183)	30	−3.002
10DE1	3	9	7.384 (0.528)	+0.262 (0.145)	30	+3.766

$\bar{R}^*$ = Mean of observed revenue for T auctions $= \frac{1}{T} \sum_{t=1}^{T} R^*(t)$

$= \frac{1}{T} \sum_{t=1}^{T} \sum_{j=N-Q+1}^{N} b_j^*(t)$

V_R^* = Variance of observed revenue $= \frac{1}{T-1} \sum_{t=1}^{T} [R^*(t) - \bar{R}^*]^2$

$\bar{R}_n^*$ = Mean revenue predicted by VHR theory for risk neutral bidders, conditional on realized experimental values, v_j^*, $= \frac{1}{T} \sum_{t=1}^{T} R_n^*(t)$

$= \frac{1}{T} \sum_{t=1}^{T} \sum_{j=N-Q+1}^{N} b_n(v_j^*(t))$

V_δ = Variance of revenue difference

$= \frac{1}{T-1} \sum_{t=1}^{T} [(R^*(t) - R_n^*(t)) - (\bar{R}^* - \bar{R}_n^*)]^2$

$t = (T)^{1/2}(\bar{R}^* - \bar{R}_n^*)/S_\delta$ is appropriate. The resulting t-values for each experiment are shown in the last column of Table IV.

From these t-values the test results are unequivocal: We reject the null hypothesis in favor of the risk averse Nash equilibrium model for designs 3–7 and 10, but we cannot reject the null hypothesis for designs 1, 2, 8, and 9.

These revenue results are plotted in Figure 11, together with corresponding data for single unit auctions from experiments reported by Cox, Roberson, and Smith [4].[7] In the single unit auctions (based on "first price" rules, where the winning bidder pays the amount bid), the above null hypothesis is rejected for $N = 4$, 5, 6, and 9 but not for $N = 3$ (see [4]). In Figure 11 mean observed revenue less the mean predicted revenue for risk neutral bidders (conditional on the sample values actually drawn) from VHR theory is plotted against the proportion of rejected bids. The revenue data from all experiments are pooled for each (N, Q) treatment condition except for those replications in which payoffs were tripled, $\bar{v}$ was tripled, or between auction information was blocked.

From the information presented in Table IV and Figure 11, one can clearly see the dichotomy of the results reported earlier. It is important to note that the results reported do not suggest a sort of random behavior where the experiments support the theory part of the time and fail to do so at other times. In those designs in which the bid and revenue level implications of the theory are supported, they are consistently supported. Yet in those designs in which the experiments do not support the theory, the theory is consistently called into question by both the individual and revenue data. Furthermore, this dichotomy is not affected if we triple payoffs, triple $\bar{v}$, or block the feedback of information on the highest accepted and rejected bids.

IV. Experimental and Theoretical Issues in the Research Program for Testing VHR Theory

The multiple unit sealed-bid auction experiments reported in the previous section have been part of a continuing research program by the authors (and other associates) in which the investigation has been constantly influenced by an interplay between our development and execution of the experiments and our efforts to further develop the theory. The form and direction of this investigation has been conditioned by a methodological perspective which recognizes that whenever a new confrontation between theory and experimental evidence suggests the existence of one or more inconsistencies, this could be due either to inadequacies in the theory or to inadequacies in the generation of the observa-

[7] There is a procedural difference between the Cox, Roberson, and Smith [4] first price sealed-bid auctions and the multiple unit discriminative auctions reported here. In the former, after each auction, subjects were informed of the market price (the highest, and only, accepted bid), but were not informed of the highest rejected bid. This procedure was used in single object auctions as an information control for comparing the first price sealed-bid auction with the Dutch auction. In the Dutch auction, where price falls until an acceptance occurs, one can never know what the highest rejected bid would have been. These alternative information conditions for discriminative multiple unit auctions have been examined in [22]. Of course the VHR theory predicts no difference in bidding behavior between these alternative information conditions.

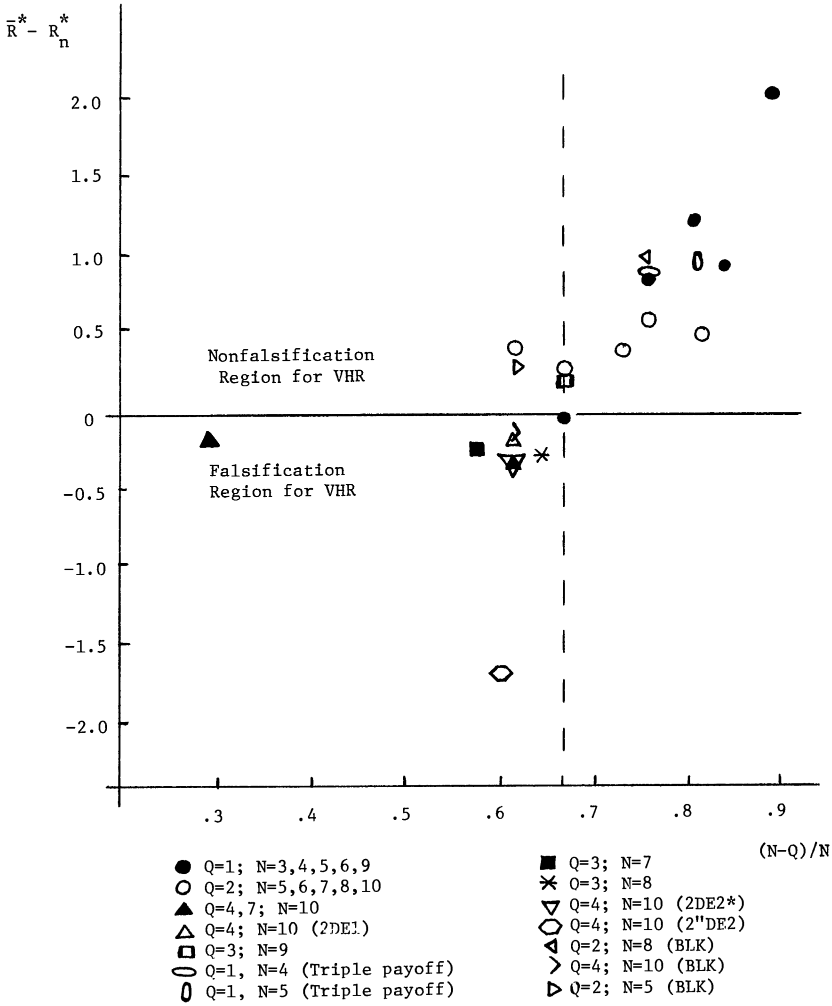

Figure 11. Critical Regions for VHR Revenues

tions—in this case the experimental test procedures. According to this perspective, any discrepancies between the observations and the implications (predictions) of the theory lead one, provisionally, to question both the theory and the observations for the purpose of taking research steps designed to resolve this inconsistency. Although no such thing as a final resolution can likely ever be asserted, our view is that before a theory can be considered to have been "falsified" by evidence, two conditions must be fulfilled. First, one must be satisfied, on the basis of an in-depth empirical examination of the sensitivity of the observations to variations in the experimental procedures, that the interpretation of the "falsifying" portion of the evidence is a defensibly correct interpretation. Second, one must be satisfied that there exists at least one other theory which does at least as well as the incumbent theory in accounting for the nonfalsifying observations, while at the same time accounting for some of the falsifying evidence.

This perspective is perhaps the dominant characteristic of the practice of science as it has been interpreted by Lakatos [14].

Our choice of the design parameters contained in Table I, the particular sequence in which the experiments based on these designs were conducted, the number of times each experimental design was replicated with different subjects, and our sensitivity tests of possible artifactual subject and level-of-reward effects on the results were all motivated and conditioned by the above considerations. In this section, we provide a detailed discussion of the questions and conjectures in our minds at the time we elected to run the particular experiments which constitute the series reported in Tables I–IV.

A. *Initial Choice of Design Parameters Influenced by First Price Auction Results*

Sometimes it is an arbitrary decision as to where in the parameter space one begins a program of new experimental research. Thus, the choice of some parameters may more likely be influenced by conjectures about where one might get the most informative observations for the money spent, than by considerations derivable from formal theory. In this case, multiple unit auction theory does not provide any guidelines as to where to begin a testing program in the parameter space defined by $N > Q \geq 2$. But in 1979, when we initiated our continuing series of multiple unit experiments, the first price (single object) auction experiments subsequently reported by Coppinger, Smith, and Titus [3] and by Cox, Roberson, and Smith [4] (hereafter CRS) had been completed, and these results strongly influenced our choice of the design parameters (N, Q) in the first several multiple unit experiments that we scheduled. These single object results also provided some guideline characteristics as to what to expect in the pattern of results from the multiple unit experiments. From the CRS experiments, we had the following conclusions: (1) The risk averse model of Nash equilibrium bidding behavior was strongly supported in all replications for $N = 4$, 5, 6, and 9. (2) For $N = 3$, CRS report some replications consistent with the risk averse model and some that are not. Pooling the revenue results across all replications for $N = 3$, CRS [4, p. 25] note that one cannot reject the risk neutral model. But since the subjects in the groups of size $N = 3$ were from the same subject pools as those in the groups of size $N > 3$, there was no basis for concluding that the deviant behavior when $N = 3$ was due to subject risk preferences (risk neutral, risk averse, or risk preferring) *independently of group size.*

These first price auction results suggested that the noncooperative model fails when group size is sufficiently small. This led us to conjecture that in the multiple unit auctions the noncooperative model might fail when group size was sufficiently small relative to the number of units offered. In the single unit auctions, the model failed only when the proportion of rejected bids was as low as two-thirds. This is seen in Figure 11 if one focuses on those mean revenue deviations that are plotted as solid circles. These five points, showing inconsistency with the model only when $(N, Q) = (3, 1)$, represented the state of our empirical knowledge when we undertook the first sequence of multiple unit auction experiments.

B. The First Series of Results for (N, Q) = (10, 7) and (10, 4)

Tentatively, we adopted the empirical hypothesis that when $\left(\frac{N-Q}{N}\right) \leq 0.67$ inconsistencies between the outcomes and VHR predictions might be expected. Consequently, we undertook a series of 16 experiments, each consisting of 23 discriminative auctions followed by 22 uniform price auctions, or vice versa, in which $\left(\frac{N-Q}{N}\right) < 0.67$. For eight of these experiments $(N, Q) = (10, 7)$, and for another eight $(N, Q) = (10, 4)$. One purpose of this series was to compare the revenue properties of the discriminative and uniform price auctions. These revenue results and comparisons are reported in Cox, Smith, and Walker [6] (hereafter CSW [6]). Since behavior may change with increased familiarity, practice, and the opportunity to reflect by experimental subjects, we also investigated "experience" as a treatment variable in this initial series of experiments.

In the first 14 of these experiments, consisting of four inexperienced and three experienced groups run under each of the two (N, Q) parameter conditions, two conclusions characterized the revenue outcomes in the discriminative auctions:

1. Subject experience had a pronounced, consistent effect on the revenue results in those sessions for which $(N, Q) = (10, 7)$. Three of the four inexperienced groups were consistent with VHR theory, while *none* of the three experienced groups were consistent with VHR theory. Consequently, we decided to run "practice" experiments and concentrate our multiple unit auction analysis on the use of experienced subjects in all subsequent experiments. This increased our research program costs, but allowed us to control for transient effects in subject behavior. In this design in which 70% of the bids are accepted, it may be that experienced subjects learn that there is minimal risk in bidding "cooperatively" since "low" bids are still likely to be accepted.
2. None of the seven experimental sessions for $(N, Q) = (10, 4)$—four using inexperienced subjects and three using experienced subjects—yielded revenue results consistent with VHR theory.

C. Effect of Screening for Noncooperative Behavior in the Uniform Price Auction

In the uniform price auction, it is a dominant strategy to bid one's value. Hence, noncooperative behavior in this auction can be identified independently of risk aversion (utility is required only to be increasing in monetary surplus). Suppose we screen for those subjects who adopt, or learn to adopt, the dominant strategy in a uniform price auction, and then use those subjects in the discriminative auction designs in which the VHR noncooperative model has been failing. If this result is reversed, it suggests that the observed tendency of VHR theory to fail in these designs is not exclusively a characteristic of the designs, but is also, in part, an identifiable characteristic of the subjects. If the results are not affected by this screening procedure, this outcome reinforces the conclusion that the failure of VHR theory in these designs is due to the parameter values. The revenue results for the two experiments using subjects screened for their tendencies to use the noncooperative strategy in uniform price auctions are reported in Table IV as experiments 1DUE2 and 2DUE2 (reported as "super experienced" experiments 1DU8′Sx and 2DU7′Sx, in CSW [6]). Clearly, these experiments yield results that do not differ significantly from those with the same parameter design, but using the experienced, unscreened subjects.

In Figure 11 the pooled revenue results for each of the parameter values $(N, Q) = (10, 7)$ and $(10, 4)$ are plotted as solid triangles. At this juncture, the combined results of the single unit auctions and the two sets of multiple unit auctions supported the hypothesis that the VHR model fails when the proportion of rejected bids was two-thirds or less.

D. Effect of Further Tests of the Proposed "Proportion-of-Rejected Bids" Hypothesis

If it is the proportion of rejected bids that is driving the separation of the parameter space into VHR falsifying and nonfalsifying sets of experiments, then since the model fails to be supported when $(N, Q) = (10, 4)$, it should also fail to be supported when $(N, Q) = (5, 2)$. Further, if this proportion is specifically two-thirds, the failure of the model when $(N, Q) = (3, 1)$ should extend to the first such multiple unit case, $(N, Q) = (6, 2)$. Finally, by this reasoning, the model should not fail when $Q = 2$ and $N > 6$. Based on these considerations, we scheduled the series of eight experiments listed under designs 3–7 in Tables I–IV. As we have seen in the previous section, both the revenue and the individual bid data for all eight of these experiments support the VHR model. The hypothesis that some simple "proportion-of-rejected-bids rule" would account for the falsifying cases appeared to be effectively refuted by these new experimental results. The pooled revenue deviations for designs 3–7 (excluding 3′ and 6′) are plotted in Figure 11 as open circles.

E. Testing Again for Artifactual Subject Effects

We wanted to be quite sure that these new experimental outcomes and their interpretation had internal validity. We therefore asked ourselves whether there might be any differences, other than the treatment parameters, between the experiments in designs 1 and 2 and those in designs 3–7. Since the experiments in designs 3–7 were conducted using a new pool of experienced subjects, recruited specifically for this new experimental series, it was conceivable that some unknown difference in the two subject pools might account for the results. This seemed unlikely, since we have long and repeated experience in finding that experimental market behavior is robust with respect to wide variations in subject groups with comparable experience. But this robustness might fail to extend to this new experimental market paradigm, and any such extension must proceed on a case-by-case empirical basis. Also, since we were getting a discrete bifurcation in the empirical support for the theory, we could not ignore the possibility that we were dealing with phenomena that were uncommonly sensitive to subject differences. This was easy to check; we replicated design 2 using a sample of 10 subjects from the pool of experienced subjects used in designs 3–7. This experiment is reported as 2DE1 in Tables III and IV. The outcome is clearly in the falsification region for the VHR model along with the other experiments in design 2. The revenue result for this experiment is plotted in Figure 11 as an open triangle. We reject, as unlikely, the hypothesis that differences in the behavior of the design 1 and 2 and design 3–7 experiments are attributable to subject differences.

F. Is $\bar{v}$ or the Level of Cash Rewards a Treatment Parameter Affecting Observed Outcomes?

Recall from Section II that each experimental design is defined by the 3-tuple, $(\bar{v}, N, Q)$, and that initially we decided to choose $\bar{v}$ as a function of (N, Q) so as to hold constant the per capita theoretical risk neutral expected earnings of subjects across experiments with differing values for (N, Q). Since actual earnings may deviate from risk neutral expected earnings if subjects are risk averse (or risk-preferring), or for other reasons, the procedure can only provide some approximate control over per capita realized earnings. Up to this point in the research program, we had not probed the internal empirical issues that relate to this experimental procedure. Two questions arise concerning our *a priori* justification for this procedure. First, since $\bar{v}$ varies with (N, Q) across designs as a by-product of holding risk neutral per capita earnings constant, could this parameter be an artifactual "treatment" that accounts for those designs that so consistently yield falsifying outcomes? Thus, it might be reasonable to conjecture that controlling for reward level across designs is of minor importance so long as rewards are "adequate" in the sense of being above some unknown threshold motivational level. Second, is the reward level itself a treatment variable in the sense that the observed behavior in any of the designs is changed if we effect a substantial increase in the reward level? If we let α be the factor that converts "experimental dollars" into U.S. currency payments to subjects, then the utility to bidder i of the experimental outcome, $v_i - b_i$, is the utility of the monetary amount, $\alpha(v_i - b_i)$, which is $u[\alpha(v_i - b_i)]$ in the VHR model. If $u(\cdot)$ is strictly concave, then the argument summarized in Section I now requires bidder i to choose $b_i = b_i^0$ so as to maximize $u[\alpha(v_i - b_i)]G(\pi(b_i))$, and this leads to a risk averse equilibrium bid function $b_a(v_i, \alpha)$ containing the conversion factor α. Hence, the VHR model generally predicts a shift in the risk averse bid function if α is changed, and in all the experiments discussed up to now we have $\alpha = 1$.

In [7] (also see [4, 5, 8]), we have extended the VHR model to allow bidders to have different constant relative risk averse (CRRA) utility functions for monetary outcome; i.e., $u_i = u[\alpha(v_i - b_i); r_i] = [\alpha(v_i - b_i)]^{r_i}$. Since in this model α affects only the scale of utility, it leads to an equilibrium bid function that is independent of the α reward parameter. Also, note that the risk neutral version of the VHR model is the special case of the CRRA model where $r_i = 1$ for all i.

These considerations led us to conduct two new experiments with design 2. In the first experiment, session 2DE2* in Tables III and IV, we tripled the experimental earnings of all subjects to obtain their payoffs in U.S. currency (i.e., $\alpha = 3$). From the revenue results in Table IV it is seen that this "treatment" has no effect on the outcomes (a result that is predicted by the CRRA model). In Figure 11 note that the revenue deviation for experiment 2DE2* (plotted as the inverted open triangle) coincides exactly with the pooled revenue deviation for the design 2 experiments which did not triple subject payoffs.

In the second new experiment, session 2″DE2, we tripled all subject value assignments; i.e., each sequence of values assigned to some subject in session 2DE2* was tripled and then assigned to a corresponding subject in 2″DE2. From Table IV it is seen that the effect is to nearly triple revenues, but the outcome is still significantly below the risk neutral predicted revenue. Also see Figure 11 where the revenue deviation for experiment 2″DE2 is plotted as a hexagon.

In Figure 11 we also plot the pooled revenue deviations from two sets of first price auctions reported in [8] in which payoffs were tripled in comparison with their corresponding control experiments. The parameter values for these experiments correspond to $(N, Q) = (4, 1)$ and $(5, 1)$. The new single and multiple unit experiments strongly reinforce the conclusion that in both the VHR-consistent and the VHR-inconsistent designs the results are due to the parameters (N, Q) and not to any effect attributable to the payoff level or to $\bar{v}$. The results are also consistent with the CRRA utility function model of Nash equilibrium bidding in [7]. Since CRRA utility is the only utility function for which individual bids are independent of the payoff scale, and for which a scaler change in $\bar{v}$ leads to the same scaler change in v_i, b_i, and payoffs, these results are particularly important in narrowing the range of models that are supported by the bidding data.[8]

G. *Effect of Between Auction Information on Highest Rejected and Highest Accepted Bids*

At this point in our experimental procedure, all experiments had involved sequences of auctions where the between auction information available to subjects included the highest rejected and highest accepted bids in the previous auctions in a sequence. (See the fourth paragraph in Section II above.) Our intention in providing this information was to create an information setting similar to those in field discriminative auctions such as the U.S. Treasury bill auction. However, from the theoretical development of Section I, it can be seen that this information is not required by bidders in the Nash equilibrium model. Hence, one might conjecture that the between auction information on highest rejected and highest accepted bids could influence subject bidding behavior and thereby affect our tests of the theory. Furthermore, if the effect of between auction bid information was different in the designs that support VHR theory than in the designs that do not support the theory then some insight might be gained into the reasons for the experimental bifurcation of the (N, Q) parameter space.

[8] Let $\mu_i = b_i/\bar{v}$ be i's normalized bid (measured as a proportion of $\bar{v}$) when i's normalized value is $\nu_i = v_i/\bar{v}$. Then the probability that a bid of μ_i or less will win is $G(\mu_i)$ and CRRA expected utility is $[\bar{v}(\nu_i - \mu_i)]^{r_i}G(\mu_i)$. If we apply the derivation procedures reported in [7] (also see the CRRA examples in [5]), the result is a normalized equilibrium bid function of the form $\mu_i = b(\nu_i, r_i)$. Therefore, $\bar{v}$ affects only the scale of utility, bids, values, and payoffs. With CRRA utility, tripling $\bar{v}$ is indistinguishable from tripling the payoff conversion factor, α, and the theory predicts that neither parameter has any effect on *normalized* bidding behavior.

To test for possible bid information effects, we conducted six new experiments in which the highest rejected and highest accepted bid information was not provided to the subjects. These experimental designs are listed as 2′, 3′, and 6′ in Tables I and II. The bidder heterogeneity and market revenue tests for these experiments are reported in Tables III and IV as the six experimental sessions labeled with the letters "BLK." The market revenue deviations for these experiments are also plotted in Figure 11. The results are clear. The levels of individual bids and market revenues observed for design 2′ are too low to be consistent with VHR theory, as they are in designs 2, 2*, and 2″. In contrast, for designs 3′ and 6′ the levels of individual bids and market revenues are consistent with VHR theory, as they are in designs 3 and 6.[9]

V. Summary and Discussion

The primary implications of VHR bidding theory that are tested using the experimental data reported in this paper are: (1) all bidders use the same increasing bid function (bidders are homogeneous in their bidding behavior); (2) no bidder will bid below the risk neutral theoretical bid function; and (3) observed auction market revenue will not be less than the corresponding risk neutral theoretical revenue. Statistical tests of hypotheses derived from these implications, using the experimental data reported in this paper, lead to the following conclusions. (1) Within our experimental designs there is a bifurcation of the results into two distinct sets defined by a corresponding partition of the (N, Q) parameter space. The levels of individual bids and market revenues are consistent with the VHR model for one group of (N, Q) parameter designs, but inconsistent with the VHR model for another group of parameter designs. This bifurcation into Group I (inconsistent) and Group C (consistent) is illustrated in Figure 12, which includes a large number of single-unit auction experiments reported earlier [4, 8]. These results are internally consistent, replicable, and remarkably robust under several experimental probes attempting to uncover artifactual and subject sample effects. (2) Across all designs, the hypothesis of bidder homogeneity (that the bids of the N individuals in a given experiment all come from the same distribution) is rejected in 24 of 28 experiments.

What are the implications of these results for the VHR model? We begin by noting that given the complexity and mathematical sophistication of the theory, the VHR model does remarkably well. In Figure 12 two-thirds (10 of 15) of the parameter designs studied to date are consistent with the revenue (and individual bid level) predictions of the model. Second, we note that the VHR model assumes the following: (1) maximization of expected utility; (2) general concave utility; (3) Nash strategic behavior; (4) identical utility functions; and (5) zero subjective

[9] When this research program was initiated, we did not consider it to be conceivable that subjects would exhibit Nash equilibrium behavior without some learning adjustment over time, perhaps in response to the feedback of this bid information. Therefore, it seemed natural to provide this information even though it was redundant for a *calculating* Nash bidder. But as noted in [6] we have not identified any characteristic, replicable trends over time in the revenue data. This suggests that the feedback of bid information may not be of any significance, which is reinforced with the results from these six "BLK" experiments.

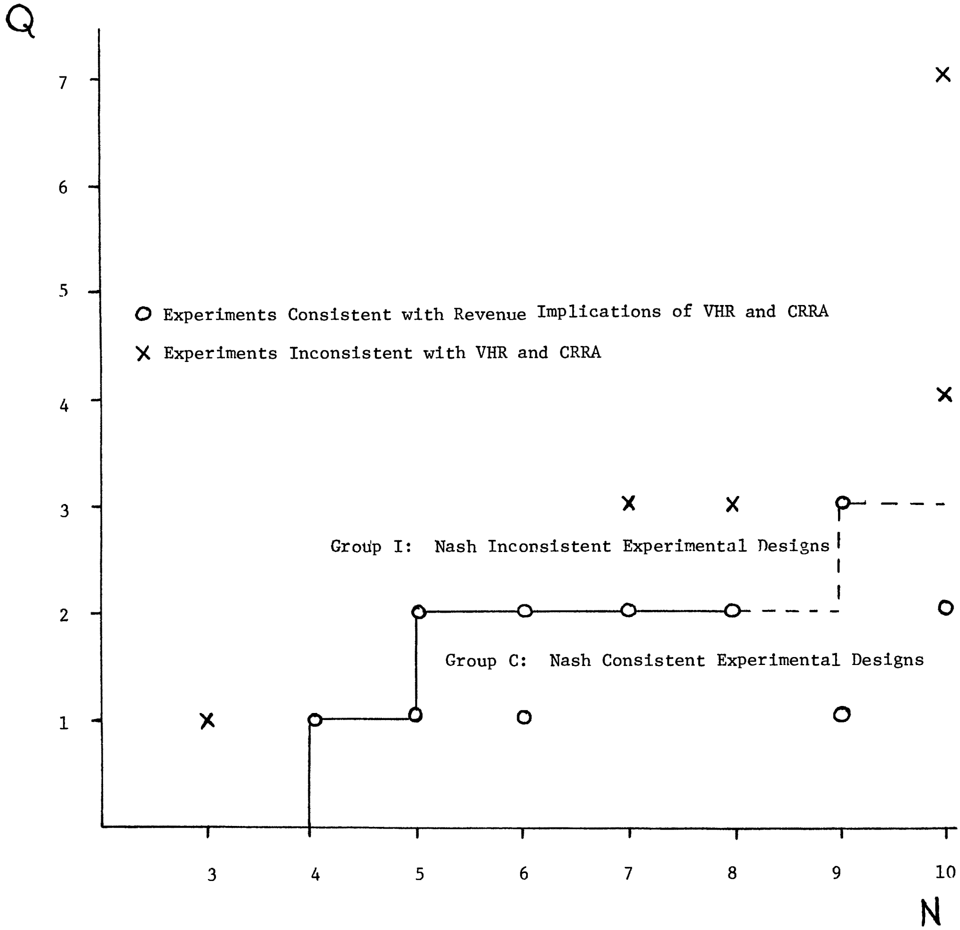

Figure 12. Bifurcation of the (N, Q) Parameter Space

cost of transacting (thinking, deciding, and acting; i.e., only $v - b$ is an argument of utility). When a theory is inconsistent with some observations, this suggests that the theory together with any or all of its assumptions is in question. But the pattern of failure of the theory may suggest "clues" as to how the theory might be modified to account for the falsifying observations, *given those ingredients of the theory that one chooses not to abandon.*[10] This criterion would seem to require us to retain assumptions (1) and (3); i.e., expected utility maximization and Nash

[10] Lakatos [14] refers to the "hard core" of a scientific research program as consisting of those assumptions which the scientist or the scientific community is unwilling to abandon in the face of falsifying evidence. Thus in the Newtonian research program, 17th and 18th century physicists did not abandon Newton's postulate of the inverse square law of mass attraction in the face of clear evidence that the planets did not move in perfect ellipses. Instead, what was modified, were the special forms (or models) of the "general" theory based on special simplifying assumptions, such as that the sun's mass is concentrated at a point, that the mass of each planet is concentrated at a point, that the mass interaction of the planets can be ignored, and so on. Thus the "hard core" was

Footnote 10 continues on next page.

strategic behavior appear to be essential if we are not going to abandon the basic concept of noncooperative equilibrium.[11] It is our position that the performance of the VHR model is sufficiently satisfactory to readily justify attempts to improve and extend it in directions disciplined by the above evidence, and that it would be quite premature at this stage to abandon its noncooperative, expected utility, foundations.

Accordingly, we have elsewhere developed heterogeneous bidder Nash models [4, 5, 7]. The multiple unit auction model developed in [7] (which generalizes the single unit auction model in [4]) replaces assumptions (2) and (4) with the specific assumption that bidders have distinct CRRA utility for monetary surplus. The CRRA model dominates the VHR model in that the former does as well as VHR in accounting for the Group C observations in Figure 12, while accounting for other evidence that is inconsistent with VHR. In particular the CRRA model is consistent with the rejection of the hypothesis that bidders bid as if they had a common bid function. It is also consistent with the experimental result that tripling payoffs has no significant effect on bidding behavior, and that tripling $\bar{v}$ effectively triples the level of individual bids.

But at this juncture in the research program we have no explanation of the Group I experiments in Figure 12 that are inconsistent with both the CRRA and VHR models. The Group I experiments are characterized by a pronounced tendency of individuals to bid below the risk neutral Vickrey bid function. It is natural to conjecture that this is due either to cooperative behavior, as suggested by CRS [4] in single unit auctions when there are only three bidders, or to strictly convex (risk-preferring) preferences for monetary outcome. However, it is one thing to conjecture that risk-preferring bidders might bid below the risk neutral bid function, but quite another matter to show that the presence of bidders with strictly convex utility can be consistent with the existence of Nash equilibrium bid functions. In [5] we have generalized the CRRA Nash equilibrium model to include a rich class of bidders in which some bidders can be risk-preferring.[12] Whether some such model can account for the Group I experiments remains an open question. Clearly, since both the Group I and Group C experiments used

surrounded by a "protective belt" of auxiliary hypotheses subject to change. Such marginal adjustments enabled Newtonian mechanics to better "account" for the observations; but anomalies remained, such as the perturbation in the orbit of the planet Mercury, which could not be accounted for by fiddling with the implementational postulates of Newton's laws. Hence, physicists simply lived with this and other anomalies until the 20th century when relativity theory accounted for some of these anomalies (such as the Mercury orbit), while at the same time also accounting for those observations which were consistent with Newton's system.

[11] An interesting and challenging alternative might be to retain Nash strategic behavior but replace expected utility maximization with Chew's weighted utility maximization [1, 2]. In effect, this approach abandons homogeneous subjective expectations. Although this alternative is worth exploring, it would appear that the more conservative line of defense should first include the more traditional expected utility calculus.

[12] It has been suggested that the failure of designs 1 and 2 to be consistent with the VHR or CRRA Nash theory could be due to the lower maximum payoff resulting from the lower value for $\bar{v}$ (see Table I) in these designs. This suggestion is most directly refuted by the results of designs 2′ and 2″, in which tripling payoffs or $\bar{v}$ did not affect the tendency of subjects to bid below the risk neutral level.

individuals from the same pool of experienced subjects, the difference in results cannot be due to differences in the risk characteristics of the subjects *independently* of the parameters (N, Q). This leaves open the possibility that there may be some interaction between the parameters (N, Q) and the presence of risk-preferring bidders that could account for the Group I outcomes without simultaneously losing consistency with the Group C outcomes. For example, the presence of only one risk-preferring bidder might lower all bids when there are only three bidders. We expect to continue our theoretical examination of these questions, and to employ direct experimental techniques for manipulating risk attitude as a means of constructing more robust tests of the effect of risk attitude on bidding behavior.

REFERENCES

1. Soo Hong Chew. "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox." *Econometrica*, Forthcoming.
2. ———. "Weighted Utility Theory, Certainty Equivalence, and General Monetary Lotteries." Discussion paper, University of Arizona, 1982.
3. Vicki M. Coppinger, Vernon L. Smith, and Jon A. Titus. "Incentives and Behavior in English, Dutch, and Sealed-Bid Auctions." *Economic Inquiry* 18 (January 1980), 1–22.
4. James C. Cox, Bruce Roberson, and Vernon L. Smith. "Theory and Behavior of Single Unit Auctions." In Vernon L. Smith (ed.), *Research in Experimental Economics*, Vol. 2, Greenwich: JAI Press, 1982.
5. James C. Cox and Vernon L. Smith. "Equilibrium Bidding Theory When Some Bidders May Be Risk-Preferring." Discussion paper, University of Arizona, 1983.
6. James C. Cox, Vernon L. Smith, and James M. Walker. "Expected Revenue in Discriminative and Uniform Price Sealed-Bid Auctions." In Vernon L. Smith (ed.), *Research in Experimental Economics*, Vol. 3, Greenwich: JAI Press, 1984 (in press).
7. ———. "Auction Market Theory of Heterogeneous Bidders." *Economics Letters* 9 (1982), 319–25.
8. ———. "Tests of a Heterogeneous Bidders Theory of First Price Auctions." *Economics Letters* 12 (1983), 207–12.
9. ———. "A Test That Discriminates Between Two Models of the Dutch-First Auction Nonisomorphism." *Journal of Economic Behavior and Organization* 4 (June–September 1983), 205–19.
10. Morris H. Degroot. *Probability and Statistics.* Reading: Addison-Wesley, 1975.
11. Milton Harris and Artur Raviv. "Allocation Mechanisms and the Design of Auctions." *Econometrica* 49 (November 1981), 1477–99.
12. Charles A. Holt, Jr. "Competitive Bidding for Contracts Under Alternative Auction Procedures." *Journal of Political Economy* 88 (June 1980), 433–45.
13. William H. Kruskal and W. Allen Wallis. "Use of Ranks in One-Criterion Analysis of Variance." *Journal of the American Statistical Association* 47 (December 1952), 583–621.
14. Imre Lakatos. *The Methodology of Scientific Research Programs.* Edited by John Worrall and Gregory Currie. Cambridge: Cambridge University Press, 1978.
15. Erik Maskin and John G. Riley. "Auctioning an Indivisible Object." Discussion paper no. 87D, JFK School of Government, Harvard University, 1980.
16. Steven Matthews. "Risk Aversion and the Efficiency of First- and Second-Price Auctions." CCBA Working paper no. 586, University of Illinois, 1979.
17. Paul R. Milgrom and Robert J. Weber. "A Theory of Auctions and Competitive Bidding." *Econometrica* 50 (September 1982), 1089–1122.
18. Roger B. Myerson. "Optimal Auction Design." *Mathematics of Operations Research* 6 (February 1981), 58–73.

19. John G. Riley and William F. Samuelson. "Optimal Auctions." *American Economic Review* 71 (June 1981), 381–92.
20. William Vickrey. "Counterspeculation, Auctions, and Competitive Sealed Tenders." *Journal of Finance* 16 (March 1961), 8–37.
21. ———. "Auction and Bidding Games." In *Recent Advances in Game Theory.* (Conference proceedings.) Princeton, Princeton University Press, 1962, pp. 15–27.
22. James M. Walker, Vernon L. Smith, and James C. Cox. "Bidding Behavior in Sealed Bid Discriminative Auctions: An Experimental Analysis of Multiple Sequence Auctions with Variations in Auction Information." Discussion paper: University of Arizona, 1983.