Theory and Misbehavior in First-Price Auctions: Comment

By JOHN H. KAGEL AND ALVIN E. ROTH*

In his recent paper in this *Review*, Glenn Harrison (1989) argues that the conclusions of James Cox et al. (1982, 1983, 1985, 1988) in their studies of first-price private-value auctions are not well supported, because of shortcomings in the way their experimental investigations were designed, analyzed, and reported. Harrison argues that the expected cost of deviations from risk-neutral Nash equilibrium (RNNE) bidding in these auctions was quite small (less than $0.05 at the median), so that in terms of expected monetary payoffs ("payoff space") many subjects had little to lose from deviating from the RNNE strategy. Harrison suggests that the significance of the differences Cox, Vernon Smith, and James Walker (hereafter CSW) report between subjects' bids and the RNNE bids (deviations in the "message space") may therefore need to be reexamined. In discussing Harrison versus CSW we have three primary points to make.[1]

First, in arguing that "it is more natural to evaluate subject behavior in expected payoff space" (Harrison, 1989 p. 749), we think Harrison has overstated his case. However, we agree with his more important point that looking at the cost of deviations is a useful diagnostic tool for determining when experimenters are likely to have lost control over subjects' incentives. Further, as we will show in Section I, this part of Harrison's critique applies with special force to CSW's studies of bidding.

Second, a broader examination of the results of private-value auction experiments indicates that risk aversion cannot be the only factor and may well not be the most important factor behind bidding above the RNNE found so often in first-price private-value auctions. The most telling evidence here is bidding above the *dominant* bid price found in second-price auctions (Kagel et al., 1987; Kagel and Levin, 1990) and the risk-loving found under several treatment conditions in CSW's (1984) own multiple-unit discriminative auctions (auctions in which the high bidders pay their bid price). These and other data inconsistent with risk-averse bidding are largely ignored in CSW (1988) but are nevertheless relevant to the substantive issue of risk aversion in private-value auctions. They are discussed in Section II.

Third, there are data gathered in other investigations which provide strong support for the view that the deviations from RNNE bidding reported in first-price auctions are not the results of the low expected cost of such deviations. However, these data, unlike the higher-stakes payoff data that CSW offer in response to Harrison, are not consistent with CSW's subsidiary conclusions that the data can be well accounted for by a narrow class of risk-aversion parameters for the bidders, together with the assumption that all agents are playing a Nash equilibrium of the resulting game of incomplete information. A key difference between these experiments and CSW's is that if subjects do not respond to CSW's treatment condition (increasing the payoffs from experimental to U.S. dollars) their behavior will be consistent with CSW's theory. In contrast,

*Department of Economics, University of Pittsburgh, Pittsburgh, PA 15260. We thank Jack Ochs and Emilie Roth for thoughtful discussions on earlier drafts of the paper, Jim Cox and Glenn Harrison for helpful comments on the initial draft of the paper, Susan Garvin for research assistance, and Ray Battalio, Carl Kogut, and Don Meyer for providing us with access to their data. Research support was provided by the Information Science and Technology and Economics divisions of the National Science Foundation, the Alfred P. Sloan Foundation, and the Russell Sage Foundation. The usual caveat applies with special force.

[1]We do not respond to specific comments that CSW (1992) make in response to our comment as, in order to avoid indefinite regress, the ground rules for this debate required us to comment on CSW's criticism of Harrison, after which they would be given the opportunity to respond to our comment.

in the experiments reported here substantial adjustments in bidding patterns are required if behavior is to satisfy constant relative risk aversion, as CSW contend. The fact that behavior satisfies CSW's theory when no change is required, but does not satisfy it when extensive changes are required, offers an important methodological lesson regarding what constitutes a demanding test of a theory. These results are discussed in Section III.

Before proceeding, let us say that we think this debate should be viewed as part of the continuing evolution of experimental methods in economics. We salute Glenn Harrison for persevering in bringing these issues to wide attention. Indeed, one of the great strengths of experimental economics is the ease with which it permits investigators to test one another's conclusions and the consequent speed with which these can be modified or even reversed when necessary.

I. The Harrison Critique and How it Applies to CSW

The papers by CSW (1983, 1985, 1988) concern bidding behavior in independent private-value auctions. Their main conclusion is that in first-price auctions not all bidders are risk-neutral, and they go on to present a number of subsidiary conclusions to the effect that the data are explained well by a family of models in which the risk aversion of each agent is parameterized in a simple way (constant relative risk aversion). Harrison's (1989) argument is that, unless similar behavior can also be observed among more highly motivated subjects, the observations reported by CSW might be due not to risk aversion, but simply to the inconsequential monetary motivation on the margin (and the consequent importance of other, uncontrolled but systematic, motivating factors). To support his argument, Harrison (1989 tables 2 and 3) replicates the experimental environment of CSW and notes that while the median observed deviations in *bids* from the risk-neutral Nash equilibrium (RNNE) in first-price auctions are relatively large (averaging $0.61), the median *forgone expected income* associated with these bids is relatively small (averaging $0.035).

CSW (1989a, b) argue in their replies that the claim that a monetary motivation of pennies is less important than a monetary motivation of dollars somehow involves an illegitimate utility comparison. This fails to address Harrison's main point, that CSW's results may simply reflect a lack of experimental control due to poor subject motivation. As a diagnostic tool, evaluating outcomes in payoff space reflects the common-sense observation that when financial incentives are "small" (i.e., there are "low" expected costs to deviating from a particular outcome), other arguments in a subject's utility function may guide behavior, the experimenter may have lost control, and one should exercise caution in evaluating the significance of the results. While the evaluation of what is "small" is subjective, when the expected cost of deviating from the RNNE lies between $0.00 and $0.06, as it does at the median of the experimental data reported in Harrison (1989), we would presume, along with Harrison, that many subjects have little to lose from deviating from the RNNE. Hence, they may be more likely to do so than if it were more expensive.

Harrison's (1989) criticism is particularly relevant to CSW's development of the constant-relative-risk-aversion model (CRRAM) for bidders whose private values, v_i, are each randomly drawn from a uniform distribution $[V_L, V_H]$. CSW (1988) devote a good deal of attention to the fact that individual bid functions do not have a zero intercept, which is very sensitive to bidders' behavior when they have the lowest valuations. Further, CSW have not included in their analysis data from bidders whose values are greatest, as they have been unable to obtain an analytic solution for the CRRAM equilibrium bid function for cases in which the v_i would produce bids that lie above b^*, where b^* is the maximum bid of the most risk-loving subject in the auction (Cox et al., 1982). Since the expected cost of deviating from the RNNE or CRRAM equilibrium bid function increases with the bidder's private valuation (as the likelihood of

TABLE 1—DEVIATIONS FROM RNNE BIDDING AS A FUNCTION OF BIDDERS' PRIVATE VALUATIONS

Experiment	Correlation coefficient (P)	Average proportionate bid deviations (SE): Lowest 20 percent of v_i	Middle 20 percent of v_i	Highest 20 percent of v_i
UH_1 ($N=6$)	−0.316 (0.001)	0.125 (0.030)	0.093 (0.006)	0.053 (0.004)
UH_2 ($N=6$)	−0.391 (0.001)	0.142 (0.023)	0.068 (0.009)	0.045 (0.006)
UH_3 ($N=6$)	−0.417 (0.001)	0.277 (0.085)	0.044 (0.007)	0.030 (0.005)
$A\&M_1$ ($N=10$)	−0.280 (0.001)	0.218 (0.043)	0.111 (0.028)	0.069 (0.003)
$A\&M_2$ ($N=10$)	−0.386 (0.001)	0.171 (0.047)	0.072 (0.017)	0.037 (0.005)
$A\&M_3$ ($N=10$)	−0.274 (0.001)	0.123 (0.029)	0.057 (0.004)	0.053 (0.004)

Notes: The correlation coefficients (r) reported are those between private valuations (V_i) and proportionate bid deviations ($|Ab_i - RNb_i|/v_i$), where Ab_i = actual bid of subject i, and RNb_i = RNNE bid of subject i; $P = \Pr(|r| = 0)$; N = number of bidders; SE = standard error of the mean. Key to experiments: UH_i = University of Houston auction series i, $v_i \in [0, 30]$; $A\&M_i$ = Texas A&M University auction series i, $v_i \in [0.10, 58.77]$.

winning the auction is higher for those with higher private valuations), CSW's investigation of CRRAM is based on private valuations for which the expected cost of deviating from equilibrium is the lowest. In other words, CSW's analysis of CRRAM is based on the private valuations with the lowest probability of winning the auction, hence the lowest expected cost of deviating from either the RNNE or the CRRAM equilibrium. This is worth emphasizing in any arguments regarding the relevance of Harrison's critique to the work of CSW.

What percentage of private valuations are thrown out by excluding bids above b^*, the maximum bid of the most risk-prone agent in the population? Assuming that the most risk-tolerant subject is risk-neutral (a fairly conservative assumption), b^* is, for example, \$7.50 in Harrison's experiments with $v_i \in [\$0.01, \$10.00]$ and four bidders. Thus, as much as 25 percent of the private valuations drawn could be excluded in the analysis of CRRAM bidding functions. (We say as much as 25 percent since CSW [1988] do not report the number of observations they eliminate.)[2] Therefore, Harrison's critique applies with special force to the CRRAM model.

Do bidders with the lowest probability of winning the auction show the largest deviations from RNNE bidding as the Harrison critique suggests? Table 1 presents a new analysis of data from a series of private-value auction experiments (reported in Douglas Dyer et al. [1989] and Raymond Battalio et al. [1990]) in which we examine this question. In each case we computed a simple correlation coefficient between private valuations and the absolute value of the size of the deviation from RNNE bidding relative to the underlying private valuation ([actual bid − RNNE predicted bid]/[private valuation]). In addition, we computed the means

[2]Note that the percentage of private valuations omitted from the analysis decreases with increases in the number of bidders. However, the central point, that CRRAM is developed on the basis of the bids with the lowest expected cost of deviating from either the RNNE or CRRAM equilibrium continues to hold.

of these proportionate deviations for the lowest and highest 20 percent of the private valuations, and for the middle 20 percent. Table 1 shows statistically significant negative correlations in each case; with lower private valuations (and lower probability of winning the auction), the absolute size of the deviation from RNNE bidding is proportionately larger. Further, the average size of this proportionate deviation is approximately two times larger in cases when bids lie in the lower 20 percent of the resale-value distribution compared to the middle 20 percent of the distribution, and 2–3 times larger compared to the top 20 percent of the distribution. That is, the average size of the proportionate deviations from RNNE bidding decreases in a fairly continuous manner in going from the lowest to the highest private valuations.[3]

II. Experimental Evidence Concerning "Overbidding"

One response to Harrison's criticism of CSW is that the deviations from the RNNE are systematically biased, in that subjects typically bid *above* the RNNE, rather than randomly distributing their bids above and below the RNNE as one might expect in the absence of adequately motivated subjects. However, there is considerable evidence that risk aversion is not the only factor, and possibly not the most important factor, promoting bidding above the RNNE in first-price private-value auction experiments.

In second-price private-value auctions (the high bidder wins the item but pays the second-highest bid price), bidders have a *dominant* strategy to bid their value for the item, *irrespective of risk attitudes* (William Vickrey, 1961). Nevertheless, subjects consistently bid above their private values by small amounts (Kagel et al., 1987; Kagel and Levin, 1990). In Kagel et al. (1987), overbidding in second-price auctions is attributed to perceptual errors on the subjects' part which are resistent to extinction because of the low expected cost of these deviations from the dominant bidding strategy.[4] In light of this overbidding, however, it is natural to entertain the hypothesis, in the absence of data to the contrary, that related perceptual errors underlie systematic deviations from RNNE bidding reported in first-price private-value auctions, when the expected cost of these deviations is relatively small.

Interestingly, Cox, Roberson, and Smith (1982; henceforth CRS) had earlier reported contrary results, namely, that in second-price auctions bids converge to the dominant bidding strategy (v_i) from below. However, unlike Kagel et al. (1987), CRS *did not permit* subjects to bid above their private valuations.[5]

Some of CSW's own reported data are inconsistent with their conclusions that bidding above the RNNE can be accounted for entirely in terms of bidders' risk aversion. In CSW (1985) they report an experiment in which subjects participated first in auctions in which they were paid directly in money and then in auctions in which they were paid in lottery tickets, using the technique of binary lottery games (Roth and Michael Malouf, 1979).[6] Since utility maximizers must be risk-neutral in lottery tickets, the binary lottery condition allows the predictions of the CRRAM theory to be tested

[3]These systematic variations in the deviations from RNNE bidding suggest heteroscedastic errors which would bias ordinary least-squares estimators of the slopes of individual-subject bid functions. Thomas Rietz (1990) argues that this, in conjunction with bid censoring in first-price auctions (it is irrational for subjects to bid above their valuations in a first-price auction, and those who do so fail to survive), generates artificially high rejection rates relative to a maintained hypothesis of risk-neutrality.

[4]An alternative explanation, that bidding above private values is an effort to limit rivals' profits, can be discounted in view of the fact that in theoretically isomorphic English auctions subjects quickly converge to the dominant bidding strategy (Kagel et al., 1987).

[5]See Kagel (1992) for further discussion of these and related second-price auction experiments inquiring into the basis for overbidding.

[6]Since its introduction into experimental economics, the technique of binary lottery games has been used by many investigators, and its strengths and limitations have been widely discussed (see e.g., Roth, 1988).

directly, since it predicts that bidders will bid precisely the RNNE bids in this case.

CSW (1985) report that most of the subjects who bid above the RNNE in the money games continued to bid above the RNNE in the binary lottery games, contrary to the predictions of the CRRAM model. Given their strong priors favoring the CRRAM model and their implicit assumption that all of the experimental data represent equilibrium bidding (plus some white-noise error term), CSW (1985 p. 165) reject the empirical adequacy of the binary lottery technique, at least as applied to auction data:

> Given the generally supportive results of earlier direct tests of the [CRRAM] bidding model, the predictive failure of Model II [binary lottery games] can be interpreted as providing (indirect) evidence against the compound lottery axiom of E[xpected] U[tility] T[heory] that is essential in Model II.

However, for people who prefer theories that organize a broad range of experimental data, no matter how well they might fit a small subset of the data, and who do not hold to the implicit assumption that all experimental data correspond to some equilibrium point prediction, these data are inconsistent with the risk-aversion hypothesis. The reason is that *for expected-utility maximizers* the binary lottery technique *must* be capable of controlling risk preferences.[7] Of course, one explanation that preserves expected-utility theory and is consistent with overbidding and the binary lottery data is that subjects are expected-utility maximizers who overbid for reasons other than risk aversion, for example from bidding errors or out-of-equilibrium behavior similar to that observed in second-price auctions.[8]

Another example of their own evidence disconfirming the risk-aversion hypothesis is CSW's (1984) multiple-unit discriminative auctions in which, for four of ten treatment conditions investigated, bids are significantly less than the RNNE prediction, which is inconsistent with risk aversion. After investigating several artifactual explanations for this behavior CSW (1984 p. 1008) conclude as follows:

> But at this juncture in the research program we have no explanation for the Group I experiments [multiple unit auctions] in Figure 12 that are inconsistent with the CRRA [constant relative risk aversion] and the VHR [Vickrey-Harris and Raviv] models (models of risk averse Nash equilibrium bidding). 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 in CRS (1982) in single unit auctions where there are only three bidders, or to strictly convex (risk preferring) preference for the monetary outcome.[9]

Here too investigators with strong enough prior beliefs in the CRRAM hypothesis may legitimately dismiss the recalcitrant treatment conditions as anomalous, but for people who are a bit more skeptical, these data

[7]See Kagel (1992) for a brief analysis of adjustment processes in private-value auction experiments.

[8]Of course, it is also possible that this part of CSW's data also may not stand up well to reexamination. Rietz (1990) questions their conclusions. We quote from his abstract:

> [CSW] use Roth and Malouf's procedure for inducing risk neutral preferences by paying subjects in the probability of winning a high value lottery prize. Expected utility maximizing subjects display linear preferences over probabilities, so they should behave risk neutrally in probability space. However, CSW continue to reject risk neutrality even after implementing this procedure on subjects who appear risk averse without it.
>
> Here, I use variants of the lottery procedure in eight sealed bid auction experiments. I find that, carefully implemented and appropriately tested, the procedure can induce behavior consistent with the intended risk preferences.

[9]The bidding below the RNNE referred to in first-price auctions with three bidders, and attributed to cooperative behavior in the original CRS paper, has not been replicated by CSW (1988) or by other investigators (Dyer et al., 1989).

also contradict the null hypothesis that bidding above the RNNE in first-price auctions can be explained primarily on the basis of risk aversion, particularly when viewed in conjunction with the second-price and lottery auction results.

Other investigators report anomalous findings as well. Kagel and Levin (1990) investigate third-price private-value auctions in which the high bidder wins the auction and pays the third-highest bid price. In third-price auctions, risk-averse bidders must bid *below* the RNNE bidding line, not above it, as in first-price auctions.[10] For auctions with five bidders, close to 80 percent of all bids lie below the RNNE line, consistent with the risk-aversion hypothesis. However, in auctions with ten bidders, approximately 65 percent of all bids lie *above* the RNNE line, which is inconsistent with risk aversion in this auction institution. This suggests that in this auction too there is overbidding that cannot be attributed solely to risk aversion.[11] These results raise further questions about whether bidding above the RNNE reported in first-price auctions can be strictly, or primarily, accounted for in terms of risk aversion.

In contrast, Dyer et al. (1989) report results from a series of first-price private-value auctions in which there is uncertainty regarding the number of bidders, and these data support the risk-aversion hypothesis. Assuming that absolute risk aversion is nonincreasing, a belief that many economists hold as a working hypothesis (Kenneth J. Arrow, 1971; Mark J. Machina, 1983), consistent with recent theoretical findings (Matthews, 1987; McAfee and McMillan, 1987), concealing information about the number of bidders raises more revenue for the seller than revealing information. Further, examination of individual bids shows that, while less than 50 percent of all individual bids satisfy the strict inequality requirements of the theory, in a large number of cases the deviations involve marginal violations of the theory, so that the data provide reasonably strong support for bidders' risk aversion. This, in conjunction with the experimental results reported in the next section, which show continued bidding above the RNNE when the expected costs of such bids are substantially higher than those reported in Harrison (1989) or most of CSW's experiments, suggests that such bidding is, at least in part, indicative of risk aversion but that it also involves bidding errors (relative to any well-defined theory of Nash equilibrium bidding) of one sort or another.

III. The Effects of Increasing Payoffs: Reevaluating Harrison's Conclusions Regarding Forgone Expected Income from Overbidding and CSW's Tests of CRRAM

One response to the Harrison critique is to conduct first-price auction experiments in which expected profits conditional on winning the auction (assuming RNNE bidding) are substantially higher then the experiments reported in Harrision (1989) or those typically conducted by CSW. The Dyer et al. (1989) and Battalio et al. (1990) experiments have done this by increasing the interval of the underlying support, $[V_L, V_H]$, of the uniform distribution from which private valuations are drawn. Comparing any two auctions with different underlying supports, and any two valuations that have the same probability of winning in both auctions (given that others are bidding the RNNE), a player with a deviant bid function, such as might be captured with the CRRAM functional form, has more to lose in terms of Harrison's risk-neutral metric in the auction with the wider interval, $[V_L, V_H]$. In other words, increasing the interval $[V_L, V_H]$ is

[10] More precisely, Kagel and Levin (1990) were able to solve for the case of constant absolute risk aversion. They argue that the prediction of bidding below the RNNE line should, however, be relatively robust to the form of the utility function specified.

[11] In a study of private-value auctions with bundling, Thomas Palfrey (1985 footnote 10) notes that, consistent with risk aversion, bidders generally "overbid" relative to the risk-neutral model when the commodity bundle has a single characteristic, so that the v_i are distributed uniformly. However, with four bidders, when the commodity bundle has two or more characteristics, so that the v_i are drawn from a triangular distribution, buyers frequently underbid relative to the risk-neutral model, which is inconsistent with risk aversion.

TABLE 2—OBSERVED BID DEVIATIONS AND FORGONE EXPECTED INCOME (ASSUMING RISK NEUTRALITY): ALL BIDDERS

	Observed bid deviations (dollars)			Forgone expected income (in dollars)		
Experiment	Median	Mean	SE	Median	Mean	SE
UH_1 ($N=3$)	2.64	2.64	0.127	0.306	0.606	0.700
UH_2 ($N=3$)	1.81	2.09	0.165	0.144	0.594	0.856
UH_3 ($N=3$)	2.40	2.30	0.145	0.236	0.541	0.642
$A\&M_1$ ($N=5$)	4.39	4.05	0.279	0.284	1.19	1.87
$A\&M_2$ ($N=5$)	2.35	1.91	0.313	0.045	0.487	0.940
$A\&M_3$ ($N=5$)	3.04	3.23	0.223	0.099	0.771	1.56

Notes: N = number of bidders; SE = standard error of the mean. Key to experiments: UH_i = University of Houston auction series i; $v_i \in [0, 30]$; $A\&M_i$ = Texas A&M University auction series i; $v_i \in [0.10, 58.77]$.

responsive to the Harrison critique in the sense that, for bidders adhering to a deviant bid function, the cost of this deviant behavior, in terms of Harrison's risk-neutral metric, will have increased for both the median bidder and the expected winning bidder.

Table 2 reports bid deviations and forgone expected income (assuming RNNE bidding) for all bidders from the experiments reported in Dyer et al. (1989) and Battalio et al. (1990).[12] Median bid deviations average 2.77, with a high of 4.39 in the $A\&M_1$ auction series and a low of 1.81 in the UH_2 auction series. Bids consistently exceed the RNNE prediction in these auctions, consistent with the results reported in Harrison (1989) and the CSW studies. Median forgone expected income averages a little over \$0.18, with a high of \$0.306 in auction series UH_1 and a low of \$0.045 in auction series $A\&M_2$. Median forgone expected income is substantially higher in these auction series than those reported in Harrison (1989) or typically observed in CSW's experiments, as are expected profits conditional on winning the auction.[13] Of course it is not possible to determine whether the forgone expected income in these auctions, or any level of forgone expected income for that matter, is sufficiently large to be salient for subjects. However, it is some three times larger than the median forgone expected income levels reported in Harrison and the typical CSW study, so that we can conclude that scaling up the expected loss function, to this level at least, does not eliminate overbidding relative to the RNNE.[14]

[12]With the exception of $A\&M_2$ and most of the data from $A\&M_3$, these auction markets employed a dual-market bidding procedure in which subjects were simultaneously bidding in two or more markets at the same time, with the market to be paid off as determined on the basis of a coin flip or some other random device. Kagel et al. (1987) and Battalio et al. (1990) report no systematic behavioral differences under dual-market as compared to single-market procedures in private-value auctions. Hence, we do not adjust expected returns by the probability of playing in any given auction.

[13]Expected profit conditional on winning the auction (assuming RNNE bidding) was \$7.43 (\$2.48 per subject) in the UH series and \$9.78 (\$1.96 per subject) in the A&M series, as compared to \$2.00 (\$0.50 per subject) in Harrison's experiments, which is slightly more than average expected profit per subject in CSW (see CSW, 1988).

[14]The average cost of deviating from RNNE bidding of \$25 reported in CSW (1989a) is quite misleading as (i) it sums up forgone profits across 25 auction periods,

TABLE 3—OBSERVED BID DEVIATIONS AND FORGONE EXPECTED INCOME (ASSUMING RISK NEUTRALITY): HIGH VALUE HOLDERS

	Observed bid deviations (dollars)			Forgone expected income (dollars)		
Experiment	Median	Mean	SE	Median	Mean	SE
UH_1 ($N = 3$)	3.27	3.42	0.187	0.72	0.97	0.101
UH_2 ($N = 3$)	3.68	3.15	0.339	0.99	1.21	0.142
UH_3 ($N = 3$)	3.72	3.42	0.222	1.02	1.01	0.105
$A\&M_1$ ($N = 5$)	7.85	7.31	0.372	2.97	3.35	0.325
$A\&M_2$ ($N = 5$)	5.18	4.63	0.480	1.05	1.43	0.240
$A\&M_3$ ($N = 5$)	5.94	5.28	0.544	1.27	2.23	0.361

Notes: N = number of bidders; SE = standard error of the mean. Key to experiments: UH_i = University of Houston auction series i; $v_i \in [0, 30]$; $A\&M_i$ = Texas A&M University auction series i; $v_i \in [0.10, 58.77]$.

Table 3 reports bid deviations and forgone expected income (relative to RNNE bidding) for the high signal holders in each auction period. The values here are all substantially higher than those reported in Table 2. In particular, average median forgone expected income is around $1.34 per auction period for the high signal holders compared to $0.18 for all bidders. This difference in median forgone expected income between the two tables results from the fact that bidders with lower resale values have substantially smaller chances of winning the auction. The primary point of Harrison's (1989) paper is that one consequence of this is that bidders with lower resale values have sharply reduced financial incentives. The net effect is that bidders with unusually low resale values (say in the lower third of the distribution of possible resale values) often adopt one of three strategies: (i) a sizable percentage continue to bid "seriously," discounting their bids relative to their signal values in patterns similar to bids based on higher resale values; (ii) some bid zero or close to zero; and (iii) others bid their resale value, or very close to it, sometimes even bidding above the resale value. Further, proportionately larger numbers of bidders with low resale values fail to respond to rudimentary changes in economic incentives: in an experiment where subjects' bids were contingent on whether there were three or six bidders in the market, approximately half the violations of the rudimentary bidding inequality $b_3(v) < b_6(v)$, where $b_n(v)$ stands for the contingent bid in the case of n bidders, were attributed to the lowest one-third of all resale values [for most of these violations $b_3(v) = b_6(v)$] (Dyer et al., 1989).

While the data in Tables 2 and 3 are consistent with risk-averse equilibrium bidding, additional data from these same experiments contradict the CRRAM model proposed by CSW. In particular, it appears that subjects overbid proportionately more

whereas the relevant measure in terms of Harrison's critique is forgone profits in a single auction period, and (ii) it takes the simple difference between the actual bid price and the RNNE bid price, without accounting for the probability of winning the auction with such a bid, as Harrison does and as we do in computing our cost measures. (Further, it involves looking at market data. As CSW argue at length, such data may involve the nonlinear portion of the CRRAM bid function and are therefore irrelevant to any evaluation of CRRAM.)

TABLE 4—ACTUAL PROFITS AND RNNE PROFITS WITH VARYING NUMBERS OF BIDDERS (STANDARD ERROR OF MEAN IN PARENTHESES)

	Small N			Large N			
Experiment	Actual profits	RNNE profits	Actual/RNNE	Actual profits	RNNE profits	Actual/RNNE	Difference in actual/RNNE: small N − large N
UH_1 ($N=3, 6$)	4.03	7.66	0.501	3.40	4.39	0.749	−0.248
	(0.276)	(0.267)	(0.023)	(0.325)	(0.110)	(0.061)	
UH_2 ($N=3, 6$)	4.02	7.57	0.553	3.46	4.43	0.766	−0.213
	(0.221)	(0.263)	(0.033)	(0.347)	(0.087)	(0.069)	
UH_3 ($N=3, 6$)	3.90	7.56	0.509	3.64	4.26	0.833	−0.324
	(0.247)	(0.302)	(0.021)	(0.293)	(0.156	(0.050)	
$A\&M_1$ ($N=5, 10$)	2.04	9.92	0.204	1.64	5.43	0.298	−0.094
	(0.225)	(0.215)	(0.021)	(0.251)	(0.058)	(0.045)	
$A\&M_2$ ($N=5, 10$)	5.20	9.92	0.517	3.87	5.32	0.711	−0.194
	(0.535)	(0.301)	(0.048)	(0.775)	(0.113)	(0.128)	
$A\&M_3$ ($N=5, 10$)	3.89	9.95	0.393	2.53	5.43	0.456	−0.063
	(0.507)	(0.257)	(0.050)	(0.364)	(0.080)	(0.062)	
$A\&M_4$ ($N=5, 10$)	2.07	4.90	0.412	1.49	2.57	0.564	−0.152
	(0.247)	(0.119)	(0.047)	(0.231)	(0.058)	(0.081)	
$A\&M_5$ ($N=5, 10$)	1.56	4.83	0.307	1.19	2.65	0.442	−0.135
	(0.162)	(0.106)	(0.028)	(0.138)	(0.052)	(0.047)	

Notes: All profits are given in units of dollars. N = number of bidders. Key to experiments: UH_i = University of Houston auction series i, $v_i \in [0, 30]$; $A\&M_i$ = Texas A&M University auction series i, $v_i \in [0.10, 58.77]$ for $i = 1, 2, 3$ and $v_i \in [0.10, 29.44]$ for $i = 4, 5$.

(act as if they are relatively more risk-averse) when there are higher expected profits to be earned, either (i) as a consequence of increasing V_H, the upper bound of the support from which private values are drawn, holding N (the number of bidders) constant, or (ii) increasing N without changing the distribution from which individual resale values are drawn. Evidence on this first point is reported in Kagel and Levin (1985) and Kagel et al. (1987). Evidence on this second point is offered in Table 4, which shows profit earned as a result of increasing N, without changing the distribution from which individual resale values are drawn, using data reported in Dyer et al. (1989) and Battalio et al. (1990). According to CRRAM, profit earned as a percentage of predicted RNNE profit should *decrease* with increases in N.[15] However, as Table 4 indicates, actual profit as a percentage of RNNE profit *increases* with increases in N in all cases, contrary to the CRRAM hypothesis.

[15]This can be seen as follows. Normalizing so that individual resale values are drawn from a uniform distribution with support $[0, V_H]$, the RNNE bid function is $[(N-1)/N]v_i$, yielding expected profits for the high bidder of $v_i - [(N-1)/N]v_i$. Under CRRAM, the bid function is $[(N-1)/(N-r_i)]v_i$, where r_i is the coefficient of relative risk aversion, with $r_i = 0$ corresponding to a risk-neutral bidder. Consequently, profit for the high bidder here is $v_i - [(N-1)/(N-r_i)]v_i$. Assuming that bidders with the high resale value win the auction, which happens in over 80 percent of all auction periods, profits for CRRAM bidders as a percentage of RNNE profits are $[1-(N-1)/(N-r_i)]/[1-(N-1)/N]$, which, for positive r_i, is decreasing for increases in N.

A simple t test based on the mean differences reported in the last column of Table 4 shows these increases to be significantly different from zero at better than the 0.01 probability level ($t = 5.92$, $d.f. = 7$). While the results of the test reported in Table 4 are a bit informal, concentrating as they do on the winning bids and not merely on those below b^*, more formal tests based strictly on private valuations below b^* yield the same result (Kagel and Levin, 1985).

These results stand in marked contrast to CSW's findings that the tripling of payoff values resulted in no significant changes in market prices (CSW, 1983) and no significant differences in mean intercept values or slopes of individual-subject bid functions (CSW, 1988), which imply that, on average, profit earned had tripled as CRRAM requires. These differences in outcomes may reflect the facts that: (i) in CSW's experiments profits were tripled by increasing the conversion rate from experimental dollars to U.S. currency, so that the distribution of private values, $[V_L, V_H]$, and all other variables over which bidders were choosing remained unchanged, requiring no changes in behavior to remain faithful to CRRAM, while (ii) varying the number of bidders while holding the distribution $[V_L, V_H]$ constant, or changing the distribution $[V_L, V_H]$ while holding the number of bidders constant, requires substantial adjustments in bidding patterns if behavior is to satisfy CRRAM, thereby providing a much more demanding test of CRRAM.

That is, CSW implemented an experimental manipulation for which their hypothesis predicted no change in subjects' behavior, and they detected no change and concluded that this supports their hypothesis. However, if the observed behavior were due to factors *not* accounted for by their CRRAM model, CSW's experimental manipulation (which left unchanged virtually all of the experimental environment) was unlikely to change these factors either, and so their failure to detect a change in behavior does little to support their hypothesis against Harrison's critique, which postulates just such uncontrolled factors. In contrast, Table 4 reports the result of a manipulation such that CRRAM predicts a *change* in behavior, and the predicted pattern of behavior was not observed.[16]

IV. Summary and Conclusions

This comment may be summarized in terms of three major points. First, CSW's argument that Harrison's (1989) measures of monetary motivation somehow involve illegitimate utility comparisons fails to address Harrison's main point, that CSW's results may simply reflect a lack of experimental control due to poor subject motivation. Indeed, as the results in Table 1 show, observed deviations from predicted (risk-neutral) bidding are inversely related to subjects' monetary motivation and are most pronounced for bidders holding the lowest private valuations.

Second, qualitative tests for risk aversion across different types of private-value auc-

[16]The design issue can be made clearer by the old joke about a man who brings a gorilla to a golf course and offers to bet that it is a champion golfer. He will offer anyone willing to bet $1,000 the following deal: after the gorilla takes each stroke, anyone who wants to settle his bet may do so; but the stakes for any gamblers who maintain their side of the bet will triple on each succeeding stroke. Faced with these generous terms, several men bet.

The gorilla addresses the ball, takes a mighty swing, and hits the ball three hundred yards down the fairway, where it comes to rest on the green within a foot of the hole. The bettors are astounded, and all settle their bets immediately. The gorilla again addresses the ball —and again hits it three hundred yards down the fairway.

Central to this joke is the experimental design it employs. After the first stroke, everyone changes priors on the hypothesis that the gorilla is a good golfer, since he can hit the ball 300 yards. However, the golfer hypothesis requires that the gorilla *change* his behavior when he is only one foot from the hole, and when he does not we see that he is not a good golfer at all.

CSW would have us test the golfer hypothesis by having the gorilla tee up on a different hole, tripling the stakes, and seeing if he hits the ball 300 yards again (in which case, they argue, we would have confirmed the hypothesis). In both cases, we would see the ball go three hundred yards again, but the design incorporated in the joke lets us see this as strong evidence against the golfer hypothesis, not for it.

tion markets yield mixed results. Although there are aspects of behavior supporting the hypothesis that risk aversion plays a role, a number of results directly contradict the risk-aversion hypothesis. The most telling of these include (i) bidding above the dominant strategy price in second-price auctions (where risk aversion plays no role), (ii) the failure of the binary lottery technique to induce risk-neutral bidding in CSW's auctions (whereas, the lottery technique must, by definition, risk-neutralize bidding if subjects are indeed expected-utility maximizers making equilibrium bids), (iii) bidding consistent with risk-loving observed in four out of ten treatment conditions in CSW's multiple-unit discriminative auctions, and (iv) bidding consistent with risk-loving observed in one of two treatment conditions involving third-price auctions. For someone holding sufficiently strong prior beliefs in CRRAM, any one of these results might be dismissed as an uncomfortable anomaly, but taken together, they suggest a pattern of behavior that is inconsistent with CSW's core hypotheses of risk aversion and equilibrium bidding. Alternatively, it might be argued that results (i), (iii), and (iv) are irrelevant to the analysis of first-price private-value auctions, as they occur in different types of auction markets. The problem with this line of reasoning is that the superior fit of the CRRAM model in first-price auctions results from its larger number of free parameters than the RNNE or homogeneous risk-aversion alternatives and thus does not directly test the core hypotheses of risk aversion and equilibrium bidding. These hypotheses have implications outside the domain of first-price auctions, and testing these implications provides one of the few ways to determine whether CSW's core hypotheses are actually able to organize bidding data in private-value auctions.

Third, data from experiments in which the cost to the median bidder, or the high expected bidder, are substantially higher than those reported in Harrison (1989), or the typical CSW experiment, show substantially higher forgone expected income using Harrison's metric 1. This indicates that the bid deviations reported in first-price auctions do not result simply from low forgone expected income. However, data from these same experiments show that subjects act as if they are relatively more risk-averse when there are higher expected profits to be earned, which directly violates predictions based on the assumption of constant relative risk aversion underlying CRRAM.

This last result differs substantially from CSW's experiments, in which the expected costs of deviating from the RNNE are increased by increasing the conversion rate from experimental into U.S. dollars. However, there is an important methodological difference between these two procedures. In CSW, if subjects are largely unresponsive to their treatment condition then CRRAM is satisfied. In contrast, increasing the underlying support from which private valuations are drawn while holding the number of bidders constant, or holding the support constant while increasing the number of bidders, requires substantially greater adjustments on the part of bidders to remain faithful to CRRAM. In other words, CSW's theory fails the more demanding test, much as it fails the more demanding test of organizing bidding patterns, even qualitatively, across different types of private-value auctions.

We believe that experimentation has more to offer economists than parameter estimates for narrow theoretical specifications. We think that much of the potential contribution of experimental methods to economics lies in their ability to provide serious tests of the basic comparative-static implications of hypotheses of economic interest. Furthermore, we think the most important methodological problems facing *any* experimenter concern how to protect against too easily accepting one's prior beliefs.

That being the case, we would be remiss if we did not note in conclusion that we think contemporary experimental economics is quite healthy in this respect. The literature is full of investigations conducted with a keen appreciation of competing hypotheses, and much of the work of experimental design is addressed to narrowing the scope for competing interpretations of the data. Even when this is not the case, experimental methods allow investigators to reexamine one another's conclusions relatively easily, and this is an important factor in the vitality of the experimental enterprise.

REFERENCES

Arrow, Kenneth J., *Essays in the Theory of Risk Bearing*, Amsterdam: North-Holland, 1971.

Battalio, Raymond C., Kogut, Carl A. and Meyer, Jack, "The Effect of Varying Number of Bidders in First Price Private Value Auctions: An Application of a Dual Market Bidding Technique," in L. Green and J. H. Kagel, eds., *Advances in Behavioral Economics*, Vol. 2, Norwood, NJ: Ablex, 1990, pp. 95–105.

Cox, James C., Roberson, Bruce and Smith, Vernon L., "Theory and Behavior of Single Object Auctions," in Vernon L. Smith, ed., *Research in Experimental Economics*, Vol. 2, Greenwich, CT: JAI Press, 1982, pp. 1–44.

______**, Smith, Vernon L. and Walker, James M.,** "Tests of a Heterogeneous Bidder's Theory of First Price Auctions," *Economics Letters*, 1983, *12* (3–4), 207–12.

______, ______ **and** ______, "Theory and Behavior of Multiple Unit Discriminative Auctions," *Journal of Finance*, September 1984, *39*, 983–1010.

______, ______ **and** ______, "Experimental Development of Sealed-Bid Auction Theory: Calibrating Controls for Risk Aversion," *American Economic Review*, May 1985 (*Papers and Proceedings*), *75*, 160–5.

______, ______ **and** ______, "Theory and Individual Behavior of First-Price Auctions," *Journal of Risk and Uncertainty*, March 1988, *1*, 61–99.

______, ______ **and** ______, (1989a) "Theory and Misbehavior in First-Price Auctions: Comment," unpublished manuscript presented at Economic Science Association meetings, Tucson, October 1989.

______, ______ **and** ______, (1989b) "A Comment on Laboratory Experimentation in Economics: A Methodological Overview," unpublished manuscript, Univeristy of Arizona, 1989.

______, ______ **and** ______, "Theory and Misbehavior in First-Price Auctions: Comment," *American Economic Review*, December 1992, *82*, 1392–1412.

Dyer, Douglas, Kagel, John H. and Levin, Dan, "Resolving Uncertainty About the Number of Bidders in Independent Private-Value Auctions: An Experimental Analysis," *Rand Journal of Economics*, Summer 1989, *20*, 268–79.

Harrison, Glenn W., "Theory and Misbehavior of First-Price Auctions," *American Economic Review*, September 1989, *79*, 749–62.

Kagel, John H., "Auctions: A Survey of Experimental Research," in John H. Kagel and Alvin E. Roth, eds., *The Handbook of Experimental Economics*, Princeton, NJ: Princeton University Press, 1992 (forthcoming).

______**, Harstad, Ronald M. and Levin, Dan,** "Information Impact and Allocation Rules in Auctions with Affiliated Private Values: A Laboratory Study," *Econometrica*, November 1987, *55*, 1275–1304.

______ **and Levin, Dan,** "Individual Bidder Behavior in First-Price Private Value Auctions," *Economics Letters*, 1985, *19* (2), 125–8.

______ **and** ______, "Independent Private Value Auctions: Bidder Behavior in First-, Second- and Third-Price Auctions with Varying Numbers of Bidders," mimeo, University of Houston, 1990.

Machina, Mark J., "The Economic Theory of Individual Behavior Toward Risk: Theory, Evidence and New Directions," Center for Research on Organization Efficiency, Stanford University Technical Report No. 433, 1983.

Matthews, Stephen A., "Comparing Auctions for Risk-Averse Buyers: A Buyer's Point of View," *Econometrica*, May 1987, *55*, 633–46.

McAfee, R. Preston and McMillan, John, "Auctions with a Stochastic Number of Bidders," *Journal of Economic Theory*, October 1987, *43*, 1–19.

Palfrey, Thomas R., "Buyer Behavior and the Welfare Effects of Bundling by a Multiproduct Monopolist: A Laboratory Investigation," in Vernon L. Smith, ed., *Research in Experimental Economics*, Vol. 3, Greenwich, CT: JAI Press, 1985, pp. 73–104.

Rietz, Thomas A., "Controlling Risk Preferences in Sealed Bid Auctions: Some Experimental Results," unpublished manuscript, Northwestern University, 1990.

Roth, Alvin E., "Laboratory Experimentation in Economics: A Methodological Overview," *Economic Journal*, December 1988, *98*, 974–1031.

______ **and Malouf, Michael W. K.**, "Game Theoretic Models and the Role of Information in Bargaining: An Experimental Study," *Psychological Review*, November 1979, *86*, 574–94.

Vickrey, William, "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, March 1961, *16*, 8–37.