

WP 96-16

November 1996



Working Paper

Department of Agricultural, Resource, and Managerial Economics
Cornell University, Ithaca, New York 14853-7801 USA

Developing a Demand Revealing Market Criterion for Contingent Valuation Validity Tests

by

Daniel Rondeau, Gregory L. Poe, and William D. Schulze

It is the policy of Cornell University actively to support equality of educational and employment opportunity. No person shall be denied admission to any educational program or activity or be denied employment on the basis of any legally prohibited discrimination involving, but not limited to, such factors as race, color, creed, religion, national or ethnic origin, sex, age or handicap. The University is committed to the maintenance of affirmative action programs which will assure the continuation of such equality of opportunity.

**Developing a Demand Revealing Market Criterion
for Contingent Valuation Validity Tests**

Daniel RONDEAU, Gregory L. POE and William D. SCHULZE

Department of Agricultural, Resource, and Managerial Economics
Cornell University

Correspondence should be addressed to:

Daniel Rondeau
Dept. Of Agricultural, Resource and Managerial Economics
449 Warren Hall
Cornell University
Ithaca, NY 14853
drr7@cornell.edu

Financial Support from the National Science Foundation, the US Environmental Protection Agency, the Social Sciences and Humanities Research Council of Canada, and the College of Agriculture and Life Sciences -- Cornell University, is gratefully acknowledged.

This is Cornell ARME Working Paper WP 96-16

Developing a Demand Revealing Market Criterion for Contingent Valuation Validity Tests

Abstract: Past research suggests that contingent valuation overstates demand for public goods. These estimates of hypothetical bias are probably invalid since they rely on voluntary contributions mechanisms which fail to reveal demand. An improved mechanism is shown to reveal aggregate demand in controlled experiments. However, individual contributions deviate from induced value.

Key Words: contingent valuation, voluntary contributions, provision point, experiments.

JEL classification H41, C92, Q20

Developing a Demand Revealing Market Criterion for Contingent Valuation Validity Tests

A critical issue in environmental economics and public policy is the ability of *contingent valuation* (CV) to estimate "actual" willingness to pay (WTP) for environmental commodities. Early validity field research compared hypothetical CV responses with values obtained from "simulated market" transactions for private (e.g. strawberries: Dickie *et al.*) and quasi-public goods (e.g. hunting permits: Bishop and Heberlein, 1979), concluding that "the overwhelming weight from simulated market experiments favors the use of contingent valuation for estimating willingness to pay" [Bishop and Heberlein, 1990, p. 101]. More recent efforts have sought to extend CV/simulated market comparisons to less familiar public goods with large nonuse components: Duffield and Patterson conducted such comparisons for leasing water rights for threatened trout streams, Siep and Strand evaluated hypothetical and actual sign-ups for an environmental organization, and Brown *et al.* compared hypothetical and real donations for the removal of roads at the Grand Canyon.¹ Together, these studies have demonstrated considerable differences between hypothetical and actual contributions, which have largely been attributed to biases associated with the hypothetical nature of CV. For example, Brown *et al.* [p. 164] argue that "Hypothetical questions, especially about donations to generally desirable environmental goods, seem to engender overestimates of actual WTP." Consistent with such analysis, efforts are underway to "calibrate" CV responses to correct for hypothetical bias [e.g. Champ *et al.*].

In this paper, we argue that past CV/simulated market comparisons may provide

¹ To the authors' credit, concern that such mechanisms allow for free-riding was noted in each article. It, of course, remains an open question whether there are incentives to free ride or behave strategically in the hypothetical CV questionnaire.

biased estimates of the hypothetical bias associated with valuing public goods. At issue is the fact that the voluntary contributions mechanisms used to solicit real monetary payments in these studies are likely to underestimate true preferences, and thus provide a false reference point for evaluating hypothetical bias. If this is true, the degree of hypothetical bias will be overstated and efforts to calibrate CV are premature. The immediate focus should instead be on developing mechanisms to better reveal true values as a reference point for validity tests. Building on an extensive research base in experimental economics, progress is reported on efforts to develop such a mechanism.

Experimental Economics Lessons: The Need for a One-Shot Demand Revealing Mechanism

An important finding from decades of economic experiments is that no public goods elicitation mechanism, even if it is theoretically incentive compatible, is perfectly demand revealing in practice. Here, demand revelation is defined in the purely empirical sense that individuals provide their true values through payments to a public fund. Public goods experiments are able to assess the degree of demand revelation because the true values of individual participants for a public good are induced by, and thus known to, the researcher.

Some public goods mechanisms that increase incentives for honest revelation have been developed over the years, and have been shown to approach demand revelation in laboratory experiments. Extending these mechanisms to CV field research is, however, problematic. Mechanisms such as the Groves-Ledyard [Groves and Ledyard] often involve complex incentive structures--greatly limiting their applicability and usefulness in situations outside of confined laboratory settings. Other mechanisms such as the Smith Auction [Smith 1979, 1980; Coursey and Smith; Harstead and Marrese] require unanimity, which

necessitates an interactive small-group situation. Moreover, these techniques require multiple rounds before they approximate individual (Groves-Ledyard) or group (Smith) demand revelation.

Given these difficulties, it is no surprise that CV researchers have turned to *voluntary contribution mechanisms* (VCM), which simply ask individuals to choose between submitting to a group fund with known returns and keeping their money in a private account. Such a technique is individually based, amenable to the single decision situations found in CV validity tests, and is readily viewed as a parallel to the standard hypothetical CV question. Yet, in theory, a standard VCM provides little or no economic incentive for one to contribute to a public good, let alone to truthfully reveal demand. Further, the large body of empirical research on VCM indicates that individuals contribute about 40 to 60 percent of the optimum in actual VCM public goods experiments involving real money (see Ledyard for a comprehensive review). From the standpoint of testing hypothetical bias, it is, thus, of serious concern that such VCM techniques have been used as the reference criterion in public goods CV/simulated markets experiments.

Three important additions to the VCM have been suggested and shown to improve contributions in public goods experiments. 1) A *provision point* (PP) is a minimum amount of aggregate contributions below which the public good is not provided. Isaac, Schmidtz and Walker, Suleiman and Rapoport, and Dawes, Orbell, Simmons and van de Kragt report that adding a PP significantly increases contributions. 2) A *money-back guarantee* (MBG) is a feature that can be added to the *provision point mechanism* (PPM) whereby individual contributions are refunded if the PP is not reached by the group. Isaac, Schmidtz and

Walker report contribution level on average four times higher in treatments with the MBG compared to baseline PP experiments. 3) A rebate rule for refunding of contributions in excess of the provision point also offers assurance against the potential loss of contributions. In the *proportional rebate* (PR) rule, all excess contributions are returned to individuals in proportion to the weight of one's contributions in the public good fund (see Marks for a study of alternative rebate rules).

The PPM with MBG and PR has been successful at increasing contributions to public goods in experimental settings. Unfortunately, this mechanism has thus far failed to produce demand revelation. However, all existing PPM experiments to date have been conducted in settings that greatly depart from field conditions. In the next sections, we report the results of a series of laboratory experiments in which we explore the performance of the PPM with MBG and PR in experimental conditions that more closely mimic key features and constraints encountered in CV field applications.

All experiments reported in this paper were conducted with subjects drawn from undergraduate Cornell University classes. At the beginning of each experiment, subjects read instructions that can be summarized as follows (instructions are available upon request). Participants are part of a group of students participating in a number of decision rounds. At the beginning of a round, each person in the group is given an initial balance of experimental dollars and must decide how much of this balance to keep and how much to allocate to a group fund. The group fund yields a return only if a predetermined investment cost (the PP) is met or exceeded. If the sum of contributions is below the PP, contributions are fully refunded (the MBG) and individual earnings for the round are equal to the initial

balance. Alternatively, if the group sum of bids equals or exceeds the PP, individual earnings for the round are the total of one's initial balance minus her bid, plus her personal return from the group fund, plus a rebate equal to her share of the contributions in excess of the PP. The PR rule was explained and illustrated by one or more examples. Subjects were aware of the number of participants in their group and that everyone had the same endowment, but were not told the level of the PP². Students knew their private payoff from the public good and that other students may not have the same payoff. Group size and subject types varied across experiments. These differences will be pointed out as required.

Experiment 1. Small Groups of Business Students

The first series of experiments that we report (which we will refer to as the "small group experiments") are the last three groups of a series of pretests conducted with students from an introductory economics class. The intent of these sessions was to test the PP/MBG/PR mechanism in small group situations akin to those conventionally used in economics experiments. All sessions were identical and serve as a baseline treatment for comparison with experiments presented in other studies as well as with our later experiments.

Individual payoffs (i.e. the induced values) from the group fund, if the provision point was met or exceeded, were the random numbers \$2.12, \$2.42, \$3.69, \$3.72, \$3.76 and

² In these experiments the total number of participants is known, but the provision point is unknown. This was done to introduce uncertainty in a manner such that participants would not gravitate towards a fair share value in which contributions approach the provision point value/number of participants. In contrast CV exercises and real public goods contributions such as land trusts, tend to have known provision points but unknown numbers of participants. Rondeau *et al.*, report large group experimental results for such conditions, which are quite similar to those reported for experiments 2 and 3 below.

\$3.90 experimental dollars, for aggregate benefits from (and aggregate demand for) the public good of \$19.63. The randomly chosen provision point was \$7.53, resulting in a benefit-cost ratio of 2.6. Experimental earnings were exchanged at the rate of one experimental dollar = \$0.25. The interested reader may consult Rapoport *et al.*; Asch *et al.*; Isaac, Schmitz and Walker; Bagnoli *et al.*; Cadsby *et al.* and Marks for PP experiments that share some of our design features. As is common in this type of experiments, the game was repeated several times. However, the context in which CV studies are conducted most closely resemble the initial round. Hence, we limit our attention to the first period data.

Let V_i and B_i respectively denote individual i 's induced value and contributions to the group fund. Then V_i/B_i is the proportion of individual i 's induced value contributed toward the public good. We report the mean and median of this ratio and we also report the ratio of the aggregate demand revealed to the aggregate demand induced ($\sum B_i / \sum V_i$). Pooling the data from all three small group experiments we find that individuals contributed an average of 64% of their induced value to the public good. The median for these 18 subjects is at 72% of value. In aggregate, the ratio of revealed to induced demands is 67%. The first column of Table 1 summarizes the results, which are consistent with aggregate demand results found in previous PPM research in small group settings (Marks, 64%; Cadsby and Maynes, 61%), but somewhat smaller than the 79% proportion computed from the first periods of Bagnoli and McKee's experiments with MBG and heterogenous valuations.

Figure 1 illustrates our results under a different light. Bars represent individual induced values. They are graphed in descending order to form the induced demand step function. Individual contributions are also ordered from high to low and plotted as a line to

represent the revealed demand curve. Contrary to the results of most hypothetical CV studies, we do not observe extreme outliers when actual cash is transacted. In fact, this dataset contains only two bids that could be considered individually irrational in the sense that $B_i > V_i$. At the other end, free riding makes for a steeper than desirable revealed demand curve. It is clear that neither our small group experiments using a PP/MBG/PR mechanism nor those PPM experiments cited above were successful at eliciting induced values from subjects. Demand revelation ratios in the 60-80% range appear to be relatively robust in small groups of students trained in economics. Thus, although these results represent sizeable improvements over the VCM, it would be premature to use small group PPMs to assess the validity of hypothetical CV surveys. Indeed, based solely on these results, one could conclude that the PP/MBG/PR is an inappropriate market public goods mechanism to use as a reference point for testing hypothetical bias in CV.

Table 1
Summary Statistics in Cents and Percentages

Experiment/ Subjects	1) Small Groups Intro Economics	2) Environmental Economics	3) Natural Resources
n	18	50	45
Mean V_i	327	325.8	300
Mean B_i	218.1	349.0	286.0
Median B_i	200	380	300
Mean B_i/V_i	64.44%	110.01%	103.17%
SD (B_i/V_i)	32.10%	48.53%	83.53%
Median B_i/V_i	72.0%	104.0%	93.3%
Dem. Revealed	66.65%	107.13%	95.19%

Experiment 2. Large Group of Environmental Economics Students

For the second experiment, three modifications were made to the original design. The group size was significantly increased, subjects were recruited from an environmental economics class, and the experiment consisted of a single-period of decision making. The decision to increase group size was motivated by the fact that CV studies often sample large groups, rather than the typical small group size used in most public goods experiments. In addition, Isaac, Walker and Williams found that individuals in groups of 40 and 100 individuals contributed significantly more to a VCM public good than did subjects in small groups. The only large group PPM experiment we are aware of was recently conducted by Rose *et al.* (n=100), and produced aggregate results consistent with demand revelation. In Rose *et al.*, subjects were given a dichotomous choice of contributing a fixed amount or nothing to a group fund within a PPM with a slightly different rebate rule consistent with a "Green Pricing" program offered by Niagara Mohawk Power Corporation. In the hope that group size effects carry over to the PPM with continuous contributions, we conducted experiment 2 with a group of 50 students.

We also modified the design to a single decision period. The primary motivations behind the adoption of the one-shot game are to conform to field conditions and to increase individual stakes. Experimental dollars were exchanged for real dollars one for one. The same six random induced values drawn for the small group experiments were used. The provision point was scaled up to \$62.75 to maintain the same benefit-cost ratio (this change does not affect results since the PP is unknown to participants).

Analysis of this experiment suggests that participants revealed their demand for the

public good in the aggregate. The second column of Table 1 provides these striking results. The mean and median individual proportion of value contributed to the group fund were respectively 110% and 104%. The mean of 110% is statistically different, at the 1% confidence level ($t=3.788 > 2.685 = t_{0.005, n=47}(\sigma_1 \neq \sigma_2)$) from the mean of 64% found in the small group experiments. However, this test is biased by the fact that roughly a quarter of all individual contributions in the large group experiment appear to have been constrained by the initial endowment of \$5. As such, the distribution of contributions is truncated, restraining both the mean and variance of the individual bid to value ratios. Yet, since the bias tends to lower the difference between the means, we are confident that the hypothesis of equality would still be rejected in the absence of the endowment constraint. Further, the median of 104% is a strong indicator of the central tendency of subjects to reveal demand and is unaffected by the truncation.

Figure 2 illustrates the aggregate results. The ratio of the area under the revealed and induced demand curves is 107%. In marked contrast to undervaluation in small group PPM experiments, the group as a whole slightly over-revealed demand. This ratio is also affected by the capping of contributions at \$5, but could only increase if the constraint was removed. In all, the results of this experiment are encouraging and suggest that it may be possible to construct a simulated market capable of eliciting aggregate demand and simple enough for field applications.

Experiment 3: Large group of natural resources/non-economics students

There is still an open debate over whether self-selection or training makes economics students more likely to act in self-interested ways (see Marwell and Ames; Isaac, McCue

and Plot; and Cadsby and Maynes). Because of the possible sensitivity of these experiments to subject type, we set out to replicate the results of experiment 2 in a slightly different environment. Subjects for the third experiment were a group of 45 students enrolled in an introductory natural resources course. Fewer than 5% of the students were currently taking or had previously taken an economics course. Also, in order to generate a better controlled and steeper demand function than previously, each subject was assigned to one of five induced values ranging from \$1.50 to \$4.50 (in real dollars), in increments of \$0.75, for a mean benefit of \$3 per person. The unannounced provision point was set at \$45 for a benefit-cost ratio of 3. In an attempt to reduce the number of bids constrained by the endowment, the initial balance was raised from \$5 to \$6.

In aggregate, participants in this experiment revealed 95% of induced demand. Individually, the median bid to value ratio is 93%, and the mean of 103% is not significantly different from 100% ($t=0.214$). The mean raw bid of \$2.86 is not statistically different from the mean induced value of \$3.00 ($t=0.49$). The median bid made by subjects is exactly \$3.00, equal to the mean and median individual induced benefit from the public good. Figure 3 illustrates, as noted for the previous experiment as well, that overbidding by some subjects approximately offsets the free-riding of others. Also, the revealed demand curve is steeper than the induced value function. Therefore, while all measures of central tendency calculated from the revealed data are accurate, the slope of the revealed demand curve does not appear to be reliable and should only be used with extreme care in policy relevant work.

Summary and Implications

Using large groups in an induced value experimental setting, this paper suggests that a proportional rebate, money back guarantee, provision point mechanism closely approximates demand revelation in the aggregate. This mechanism should be simple to implement and understand in field studies and is specifically designed for public goods with a predetermined magnitude, much like the public goods that CV is frequently applied to.

While the slopes of the resulting revealed demand curves are relatively poor approximations of the induced demand functions, the resulting measures of central tendency provided very strong predictors of the mean induced value. In the aggregate, free-riding is approximately compensated by the over-contributions of others -- a result that is in marked contrast to the 40 to 60 percent underrevelation associated with the VCM. As such, we argue that the PP/MBG/PR mechanism represents a substantial improvement over VCM measures used as a reference criterion in recent CV/simulated market validity tests.

From the perspective of conducting field validity tests that accurately measure the hypothetical bias of the CV method, these results are encouraging. However, the extent to which such laboratory experiments reflect contribution situations in the real world remains to be determined. While our experimental environments resembled field conditions more closely than previous studies, the behavior of college students in laboratory settings cannot readily be taken to reflect that of ordinary citizens in every day life without additional research, both in the lab and in the field.

Figure 1

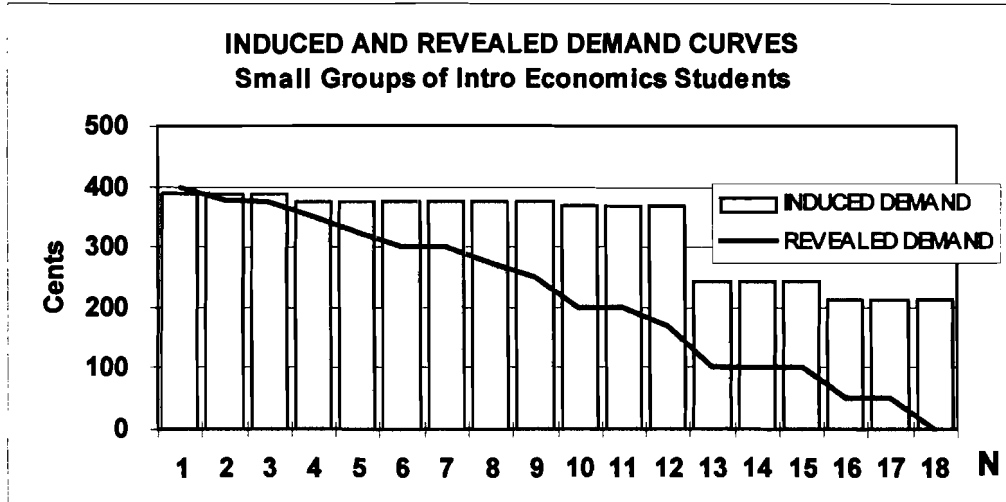


Figure 2

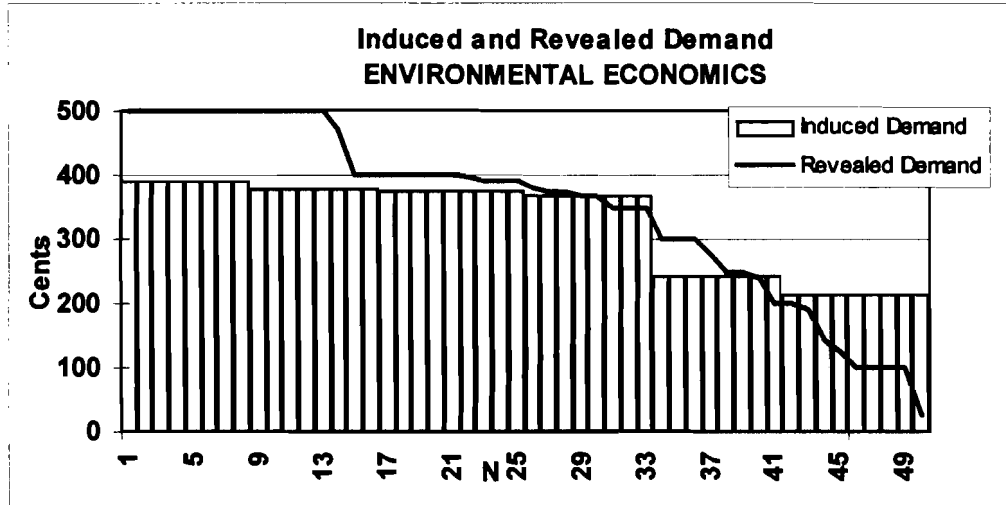
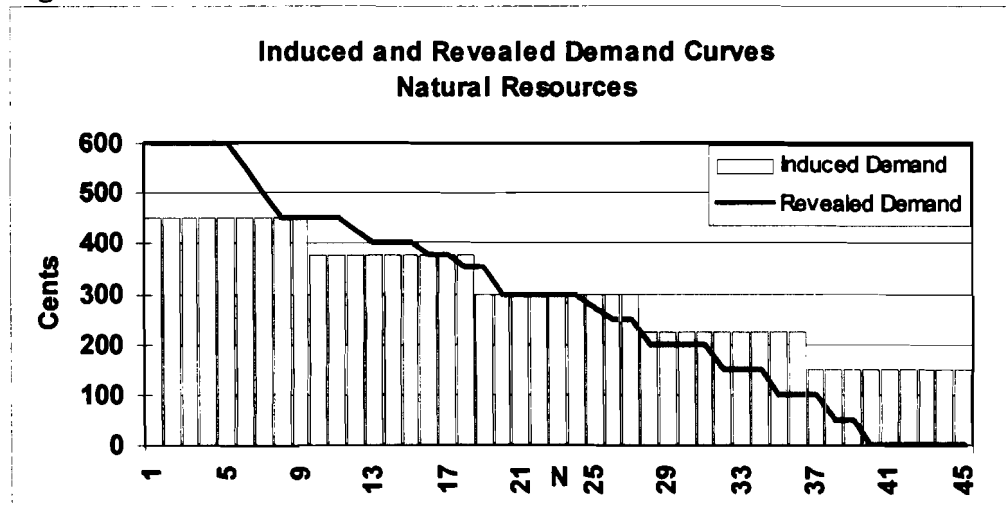


Figure 3



References

- Asch, P., G.A. Gigliotti and J.A. Polito. (1993). "Free Riding with Discrete and Continuous Public Goods: Some Experimental Evidence." *Public Choice* 77:293-305.
- Bagnoli, M. and M. McKee. (1991). "Voluntary Contributions Games: Efficient Private Provision of Public Goods." *Economic Inquiry* 29:351-366.
- Bishop, R. C., and T. A. Heberlein, 1979. "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?", *American Journal of Agricultural Economics*, 61:926-30.
- Bishop, R. C., T. A. Heberlein, and M.J. Kealy, 1983. "Contingent Valuation of Environmental Assets: Comparisons with a Simulated Market, *Natural Resources Journal*, 23:619-33.
- Bishop, R. C., and T. A. Heberlein, 1990. "Chapter 6: The Contingent Valuation Method", in D. V. Johnson, ed. *Economic Valuation of Natural Resources: Issues, Theory, and Applications*, 181-204.
- Bohm, P. (1972). "Estimating Demand for Public Goods: An Experiment." *European Economic Review* 3:111-130.
- Brookshire D.S., D.L. Coursey and W.D. Schulze. (1990). "Experiments in the Solicitation of Private and Public Values: An Overview." in L. Green and J.H. Kagel eds. *Advances in Behavioral Economics*, Vol. II. Ablex Publishing Corporation.
- Cadsby, C.B., and E. Maynes. (1996). "Choosing Between a Cooperative and Non-Cooperative Equilibrium: Nurses versus Economics and Business Students." Unpublished Manuscript.
- Coursey D. and V.L. Smith. (1984). "Experimental Tests of an Allocation Mechanism for Private, Public or Externality Goods." *Scandinavian Journal of Economics* 86:468-484.
- Davis, D.D. and C.A. Holt. (1993). *Experimental Economics*. Princeton, Princeton University Press. 572 pages.
- Dawes, R., J. Orbell, R. Simmons and A. van de Kragt. (1986). "Organizing Groups for Collective Action." *American Political Science Review* 8:1171-85.
- Dickie, M., A. Fisher, and S. Gerking, 1987. Market Transactions and Hypothetical Demand Data: A Comparative Study, *Journal of the American Statistical Society*, 82:69-75.
- Duffield, J. W. and D. A. Patterson, 1992. "Field Testing of Existence Values: An Instream Flow Trust Fund for Montana Rivers", paper presented at the AERE/ASSA meetings, New Orleans.
- Groves, T. and J. Ledyard. (1977). "Optimal Allocation of Public Goods: A Solution to the 'Free Rider' Problem." *Econometrica* 45:783-809.
- Harstad, R. and M. Marrese. (1982). "Behavioral Explanations of Efficient Public Good Allocations." *Journal of Public Economics* 19:367-383.

- Isaac, R.M., K. McCue and C. Plott. (1985). "Public Goods Provision in an Experimental Environment." *Journal of Public Economics* 26:51-74.
- Isaac, R.M., D. Schmidt and J. Walker. (1989). "The Assurance Problem in Laboratory Markets." *Public Choice* 62:217-236.
- Isaac, R.M., J. Walker and A. Williams. (1994). "Group Size and The Voluntary Provision of Public Goods: Experimental Evidence Using Very Large Groups." *Journal of Public Economics* 54:1-36.
- Ledyard, J.O. (1995). "Public Goods: A Survey of Experimental Research." in J.H. Kagel and A.E. Roth eds. Handbook of Experimental Economics. Princeton, Princeton University Press. pp. 111-194.
- Marks, M.B. (1993). "The Effect of Alternative Rebate Rules in the Provision Point Mechanism of Voluntary Contributions: An Experimental Investigation." Unpublished Manuscript.
- Marwell, G., and R. Ames. (1981). "Economists Free Ride, Does Anyone Else? Experiments on the Provision of Public Goods IV." *Journal of Public Economics* 15:295-310.
- Marwell G. and R. Ames. (1979). "Experiments on the Provision of Public Goods I: Resources, Interest, Group Size and the Free Rider Problem." *American Journal of Sociology* 84:926-937.
- Rapoport, A. and R. Suleiman. (1993). "Incremental Contribution in Step-Level Public Goods Games with Asymmetric Players." *Organizational Behavior and human Decision Processes* 55:55:171-194.
- Rondeau, D., W.D. Schulze and G.L. Poe. (1996). "Demand Revelation in Single-Period Provision Point Mechanisms with Incomplete Information". Paper presented at the annual meetings of the Economic Science Association, Tucson, October, 20.
- Rose, S., J. Clark, G.L. Poe, D. Rondeau and W.D. Schulze. (1996). "Field and Laboratory Tests of A Provision Point Mechanism." Paper presented at the annual meetings of the Economic Science Association. Tucson, October 19.
- Seip, K, and J. Strand, 1992. "Willingness to Pay for Environmental Goods in Norway: A Contingent Valuation Study with Real Payment". *Environemantl and Resource Economics*, 2:91-106.
- Smith, V.L. (1980). "Experiments with a Decentralized Mechanism for Public Goods Decision." *American Economic Review* 70:584-599.
- Smith, V.L. (1979). "An Experimental Comparison of Three Public Good Decision Mechanisms. *Scandinavian Journal of Economics* 81:198-215.
- Suleiman, R. and A. Rapoport. (1992). Provision of Step-Level Public Goods with Continuous Contribution." *Journal of Behavioral Decision Making* 5:133-153.

OTHER A.R.M.E. WORKING PAPERS

- | | | |
|-----------|---|--|
| No. 96-07 | Commodity Futures Prices as Forecasts | William G. Tomek |
| No. 96-08 | Old-Growth Forest and Jobs | Jon M. Conrad |
| No. 96-09 | The Economic Threshold With a Stochastic Pest Population: An Application to the European Red Mite | Jean-Daniel Saphores
Jon M. Conrad |
| No. 96-10 | Competitive Electricity Markets in New York State: Empirical Impacts of Industry Restructuring | Robert Ethier
Timothy D. Mount |
| No. 96-11 | Do Participants in Well Water Testing Programs Update Their Exposure and Health Risk Perceptions? | Gregory Poe
Harold van Es
Maxine Duroe
Christopher Crossman
Timothy Vandenberg
Richard Bishop |
| No. 96-12 | A Hedonic Approach to Estimating Operation and Maintenance Costs for New York Municipal Water Systems | Todd M. Schmit
Richard N. Boisvert |
| No. 96-13 | The Productivity of Dairy Farms Measured by Non-Parametric Malmquist Indices | Loren W. Tauer |
| No. 96-14 | A Unique Measure of the Welfare Effects of Price Support Programs for Corn on Family-Farm Households by Size Distribution | Tebogo B. Seleka
Harry de Gorter |
| No. 96-15 | Economies of Size in Water Treatment vs. Diseconomies of Dispersion for Small Public Water Systems | Richard N. Boisvert
Todd M. Schmit |