*Public Choice* **85**: 249–266, 1995.


# Heterogenous demand for public goods: Behavior in the voluntary contributions mechanism*

JOSEPH FISHER
*Department of Accounting, Indiana University, Bloomington, IN 47401, U.S.A.*

R. MARK ISAAC
*Department of Economics, University of Arizona, Tuscon, AZ 85721, U.S.A.*

JEFFREY W. SCHATZBERG
*Department of Accounting, University of Arizona, Tucson, AZ 85721, U.S.A.*

JAMES M. WALKER
*Department of Economics and Workshop in Political Theory and Policy Analysis, Indiana University, Bloomington, IN 47401, U.S.A.*

Accepted 16 December 1993

**Abstract.** Numerous laboratory experiments have investigated the performance of several processes for providing public goods through voluntary contributions. This research has been able to identify features of the institution or environment which are reliably likely to produce outcomes "close" to the free riding outcome or "substantially" greater than the pessimistic prediction of standard models. One such feature is the "marginal per-capita return" (MPCR) from the public good. Various authors have altered MPCR between groups or for an entire group at the same time. The experiments reported here address a different question, "What would happen if, within a group, some persons faced a 'high' MPCR while others faced a 'low' MPCR?"

## 1. Introduction

Over the past few years, numerous laboratory experiments have investigated the performance of several processes for providing public goods through voluntary contributions.[1] This evaluation has focused on both aggregate group behavior and on individual behavior. At the aggregate level, the effect of institutional and parametric variations has been compared to a social wel-

* The heterogenous group experiments reported here were conducted at the Economic Science Laboratory of the University of Arizona. All data are archived on the NovaNet computer system and are available from the authors upon request. The financial support of the National Science Foundation, grants SES-8820897 and SES-8821067, is gratefully acknowledged. Fisher and Walker are at Indiana University; Isaac and Schatzberg are at the University of Arizona.

fare measure (such as level of the public good provided, level of the public good as a percentage of the efficient level, overall system efficiency, etc.) At the individual level, behavior has been compared to theoretical (i.e., equilibrium) predictions.

One of the *a priori* conjectures which has received the most attention is the so called "free-riding" hypothesis. Informally, this proposition leads to the conjecture that public goods will be underprovided. In the experimental laboratory, precise predictions regarding individual behavior can be derived given a controlled decision making institution and specific parameterizations.[2] A broad, but misleading, summary of the experimental evidence to date is that, in those systems in which the free riding hypothesis should apply, it predicts only moderately well. This summary is misleading because it runs the risk of being (mis)interpreted as saying that intermediate levels of public good provision (much less than optimal but much more than the free riding prediction) is a uniform occurrence. This is not the case.

While it is true that intermediate levels of the public good are often provided, the research to date has been able to identify certain features of the institution or environment which are reliably likely to produce outcomes "close" to the free riding outcome or "substantially" greater than the pessimistic prediction of standard models. One such feature identified by Isaac, Walker, and Thomas (1984), Isaac and Walker (1988a), and Isaac, Walker, and Williams (1994) is the "marginal per-capita return" (MPCR) from the public good. Broadly defined, it is the marginal gain in moving an incremental unit of wealth to public goods provision relative to the costs of doing so. Even though the single period dominant strategy (and multi-period complete information Nash) equilibrium predictions were unchanged, these earlier studies reported replicable behavioral differences when groups faced an MPCR of 0.30 as opposed to 0.75.

In their work to date, Isaac, Walker, and colleagues have altered MPCR only between groups or, within an experimental sequence of decision periods, for an entire group at the same time. That is, comparisons have been between groups (or for the same group between decision periods) where *all* persons had an MPCR of 0.30 or *all* persons had an MPCR of 0.75. Of course, this is not the only way in which differential demands for a public good could occur. What would happen if, within a group, some persons faced a "high" MPCR while other persons faced a "low" MPCR? The experiments reported here address this question.

## 2. The voluntary contributions mechanism, prior research, and preliminary conjectures

### 2.1. *The voluntary contributions mechanism*

There have been several laboratory versions of a voluntary contributions mechanism. The one reported here is the NovaNet computerized version first reported by Isaac, Walker, and Thomas (1984). Each individual in a group of size n faces ten "investment" decision periods. In each period, the experimenters endow each participant (i) with $z_i$ tokens. Each token will be invested in an "individual exchange" (where it pays $p_i$ with certainty), or in a "group exchange". Note that an individual is free to divide his tokens between the two types of exchange but is not allowed to carry over tokens from one period to the next. Let $m_i$ represent individual i's contribution of tokens to the group exchange in a given period. The group exchange is a public good in that each individual receives a payment of:

$$(1/n)G(m_i + \Sigma m_j) \text{ cents}$$

where $G(\cdot)$ is an appropriately specified function and $\Sigma m_j$ represents the sum of the contributions of everyone else except person i. In fact, the $p_i$ amounts and the $G(\cdot)$ function were chosen so that the Pareto optimum (in this experiment defined simply as the outcome that provided the greatest total monetary payout from the experimenters to the subjects) was for every individual always to invest all tokens in the group exchange (i.e., to set $m_i = z_i$). In addition, these parameters were chosen so that the single period dominant strategy (and hence the unique, backward unravelling complete information Nash equilibrium) was for each person to invest zero tokens in the group exchange.

### 2.2. *Prior research*

Our starting point is the research reported by Isaac and Walker (1988a), hereafter I&W. They presented data from 12 experimental series (24 ten-period sessions) which suggested that MPCR was an important influence on group behavior, examining MPCRs of 0.30 and 0.75. The Pareto optimal decisions and the single period dominant strategy predictions were *identical* for either level of MPCR, and yet MPCR was a treatment found to be associated with different behavior. Specifically, groups facing an MPCR of 0.30 were more likely to have contributions which decayed to (or very close to) the "free riding" predictions. On the other hand, groups with an MPCR of 0.75 were more likely to exhibit the "intermediate" behavior of neither full free riding nor efficient

contributions. For example, in true end periods, groups with MPCR = .30 contributed on average only 3.65 percent of the optimal number of tokens to the public good, while groups with MPCR = .75 contributed 26.35 percent, on average.

There are several experimental papers involving other public goods institutions or other economic environments which examine the influence of group heterogeneity in demand for the public good. Isaac, McCue, and Plott (1985) using a different version of a voluntary contributions mechanism, had groups of 10, consisting of five "low" demanders and five "high" demanders.[3] No participant was informed of the demand conditions for anyone but himself. High demanders were found to contribute more in absolute terms but *not* in terms of percentage of the Lindahl optimal amount.

Marwell and Ames (1979) examine a treatment which mixes "high" and "low" interest persons. The concept of MPCR is difficult to recover from their design because of the numerous discontinuities of their group payoff function. However, from their discussions we infer that certainly the "low interest" persons are expected to undercontribute to the group good. They report that for small groups with unequal interest, the "low interest" subjects "behave very much like equal subjects in other conditions."

Fisher and Schatzberg (1988) (hereafter F&S) report on a number of binary choice prisoner's dilemma experiments comparing symmetric matrices having, alternatively, "high" and "low" rewards for cooperation (analogous, but not exactly comparable, to the concept of the MPCR), to asymmetric matrices in which one person faced high rewards for cooperation while the other faced low rewards. They found, similar to I&W, that persons in the symmetric conditions cooperated more with high incentive conditions, even though both matrices were classic prisoner's dilemmas. Furthermore, F&S found that in the asymmetric matrices the high incentive subjects did in fact cooperate more than the low incentive subjects. Interestingly, low incentive subjects were indistinguishable whether they came from symmetric or asymmetric matrix pairs, while high incentive persons differed in the two cases. High incentive subjects cooperated less frequently when in asymmetric matrices than in symmetric matrices, (i.e., there appeared to be a "poisoning-of-the-well" effect).

After completing a first draft of this paper we became aware of other research in process investigating heterogenous MPCRs. Brookshire, Coursey and Redington (1988) investigate a public good decision setting where subsets of groups obtain a disproportional benefit from the provision of the public good. They investigate a more extreme case in which one or two individuals have incentives to provide the public good unilaterally. On average, they did not find conclusive evidence that large interest individuals significantly increase total group contributions.[4]

### 2.3. *Informal conjectures*

As noted, all of the results reported in I&W investigate the effect of MPCR when the groups themselves are homogenous. That is, in that research any one experimental group was characterized in any one decision period by all of its members having an identical MPCR of either 0.30 or 0.75. An obvious extension is to investigate the effect of MPCR in a setting where individuals within a group are heterogenous (some facing a low MPCR and some facing a high MPCR). In other words, the goal is to investigate the F&S questions in the I&W public goods domain.

Virtually any attempt to model the influence of MPCR will be non-standard. Given the known end period, and the dominant strategy incentive to free-ride in any one period, any MPCR of less than one should yield the same pessimistic predictions in any standard treatment. Nevertheless, the research cited above demonstrates the existence of an MPCR effect in homogenous groups. There have been, however, several attempts at modelling contribution decisions in a way which captures the MPCR effect. See Isaac, Walker, and Williams (1994) for a discussion of these attempts.

In a heterogenous environment, the question of how to explain the effects of MPCR becomes even more complicated. It is not merely a question of asking whether earlier findings related to MPCR are replicated. One conjecture is that what really matters is one's own MPCR. In this case one would expect that persons of high or low MPCR type in heterogenous groups will exhibit behavior very similar to persons in homogenous groups of that type. The other possibility is that persons respond either directly or indirectly to the MPCRs of others. Suppose, as an example, the behavior of persons with an MPCR of 0.30 in the current experiments (where the possibility of asymmetric MPCRs is common knowledge) is different from that observed in the I&W environment. This suggests that a person's behavior is influenced not only by his own MPCR but somehow by the MPCRs of others (perhaps it is the knowledge or expectation of other person's MPCR; perhaps it is a less explicit reaction to the behavior of other persons).[5] But, if one persons's decisions are dependent on the MPCR of others, what is the nature of that interdependence? The F&S results suggest a "poisoning of the well" effect in which the contribution rates of high MPCR subjects fall when compared to baseline homogenous groups. Such a phenomenon would have important implications for the formation of "privileged" or "intermediate" groups (Olson, 1971: 49–50) in which some members have a greater incentive to provide the public good than others. Clearly, other forms of the interdependency among MPCRs may also exist. We will examine three broad conjectures (no effect, "seeding" and "poisoning of the well") in the next section after the experimental design and parameters are described in more detail.

### 3. Experimental design, parameters, and research hypotheses

We report the results of 20 experimental sessions from ten different groups. Each group consisted of four individuals experienced in the voluntary contributions process but not in the specific parameters of these experiments. Specifically, these individuals had been subjects in earlier voluntary contributions experiments, which we will refer to as "training sessions." These training sessions were designed for two reasons: 1) to familiarize the subjects with the mechanics of the institution; and 2) to keep the experience level of the research sessions consistent with that reported in I&W. No four-person group in the research sessions had participated as a group in the training sessions.[6]

Each of the ten groups participated in two ten-period experimental sessions. In all 20 of these research sessions, the $G(\cdot)$ function used to define the return to the public good was set equal to $2.4\Sigma m_i$, measured in cents. Different MPCRs were created by giving two subjects a $p_i$ equal to 2.0 cents and two others a $p_i$ equal to 0.8 cents. The two MPCRs thus defined were 0.30 and 0.75, as in the I&W experiments. In these research sessions as well as in the training sessions the same instructions were used. It was revealed to each subject that the return from the individual exchange might *not* be an experimental constant. The subjects were not told explicitly that different individuals necessarily had different $p_i$'s, but they were told that the $p_i$'s were not necessarily the same. In addition, each subject observed the following in both the research sessions and the training sessions: i) their own $p_i$ changed between the first ten-period session (what we will call here "Year 1") and the second ten-period session ("Year 2"); and ii) the $p_i$'s for Year 2 were distributed privately on separate pieces of paper. Because the same instructions and Year 1–Year 2 distinctions were used in both the training sessions and the research sessions, our operative assumption was that in returning to the same institution within a matter of days (the time between the training session and the research session), the subjects would not "home grow" incorrect expectations of *necessarily* homogenous MPCR.

At the completion of the first ten periods (Year 1), the subjects were informed of their earnings for those ten periods,[7] and then informed that a second ten periods would be conducted, but with possibly different $p_i$'s. Each subject was then privately informed of his/her new $p_i$. In fact, all of the subjects who had been low MPCR types in the first 10 periods were switched to having high MPCRs, and vice versa. It is important to note the difference between the changes in MPCR reported here and those in I&W. In I&W, changes in each subject's MPCR, across Year 1 and Year 2, always went in the same direction: as each subject's MPCR was raised (lowered), so was that of all other subjects. However, in this paper, MPCR increases across Year 1 and Year 2 for two subjects, while for the other two subjects, MPCR decreases.

Combining the informal conjectures of the previous section with the specifics of the experimental design, we propose two sets of hypotheses: the first for aggregate group decisions and the second for behavior disaggregated by MPCR type.

### 3.1. *Hypotheses: Aggregate group decisions*

**Hypothesis 1.** Group contributions to the group good as a percent of optimum will resemble that which would occur if the entire group were comprised of low MPCR types (MPCR = 0.30).

**Hypothesis 2.** Group contributions to the group good as a percent of optimum will resemble that which would occur if the entire group were comprised of high MPCR types (MPCR = 0.75).

**Hypothesis 3.** Group contributions to the group good as a percent of optimum will resemble an average of what would occur across either of the two homogenous MPCR conditions. That is, let H be the level of contributions from a homogenous group of high MPCR types and L be that of a homogenous group of low MPCR types. In the current research sessions, group contributions should resemble an average of H and L.

Notice that if subjects' behavior is influenced *only* by their own MPCR we should observe data in support of Hypothesis 3. High MPCR types will look like homogenous high MPCR individuals, low MPCR types will look like homogenous low MPCR individuals. The average across the group will be contributions as a percent of optimum between the homogenous pairs. There is no such clear-cut distinction if individual behavior depends on the vector of MPCRs (i.e., own and others' MPCRs); data consistent with any of the three hypotheses are possible. Just to give three examples, consider the following: i) if individuals in both the high and the low MPCR groups contribute an amount that is the average of what an individual would contribute *in the homogenous group*, then the data would support Hypothesis 3; ii) if individual behavior were to show a greater level of contribution the greater were the *maximum* MPCR of *any* member, then data consistent with Hypothesis 2 would occur (this would be an example of the presence of high MPCR types "seeding" contributions from others); iii) if individual contributions were pulled down by the *minimum* MPCR of any member, then the data would be consistent with Hypothesis 1 (this would be an example of the low MPCR types "poisoning the well" for others' contributions).

### 3.2. *Hypotheses: Data disaggregated by MPCR type*

**Hypothesis 4.** Contributions to the group good as a percent of optimum will be the same for low and high MPCR types.

The experimental results from other studies offer some mixed evidence related to Hypothesis 4. The Isaac-McCue-Plott data could suggest a difference (based on contributions in absolute terms) or no difference (based on percent of the Lindahl optimum). In the research reported here, persons with different MPCRs nevertheless have identical token endowments and identical optimal levels of contributions. Their token endowments differ in U.S. cash equivalence due to heterogenous returns from the individual exchange. The experimental results of F&S would suggest rejection of Hypothesis 4: low and high MPCR types will make different contributions to the group good as a percent of optimum.

Hypotheses 4 relates to whether low and high MPCR types behave like one another. In addition, it will be interesting to compare each type to prior experiments consisting exclusively of low or high MPCR types. Specifically, we will be looking for whether high (low) MPCR types in heterogenous groups behave similarly to their high (low) MPCR counterparts in homogenous groups.

Notice that a test of the above hypotheses and their related questions requires a baseline from experimental sessions in which groups were homogenous at MPCR of 0.30 or 0.75. We generated this baseline from the data of all four-person, non-communication, experienced groups reported in Isaac and Walker (1988a, 1988b).

## 4. Experimental results and interpretation

The focus of the data analysis reported here is on group behavior. The principal research question involves group and sub-group dynamics when there is heterogeneity and uncertainty among group members regarding that heterogeneity.[9]

We present the results using the same aggregate-disaggregate distinction as above. All new data are from the 20 experimental sessions using experienced subjects. In all of the tests that follow, the basic unit of observation is an individual's contribution to the group good as a percentage of optimum for a decision period. In the figures we present the data aggregated to the level of group or sub-group means for a decision period. We organize the results around three conclusions.

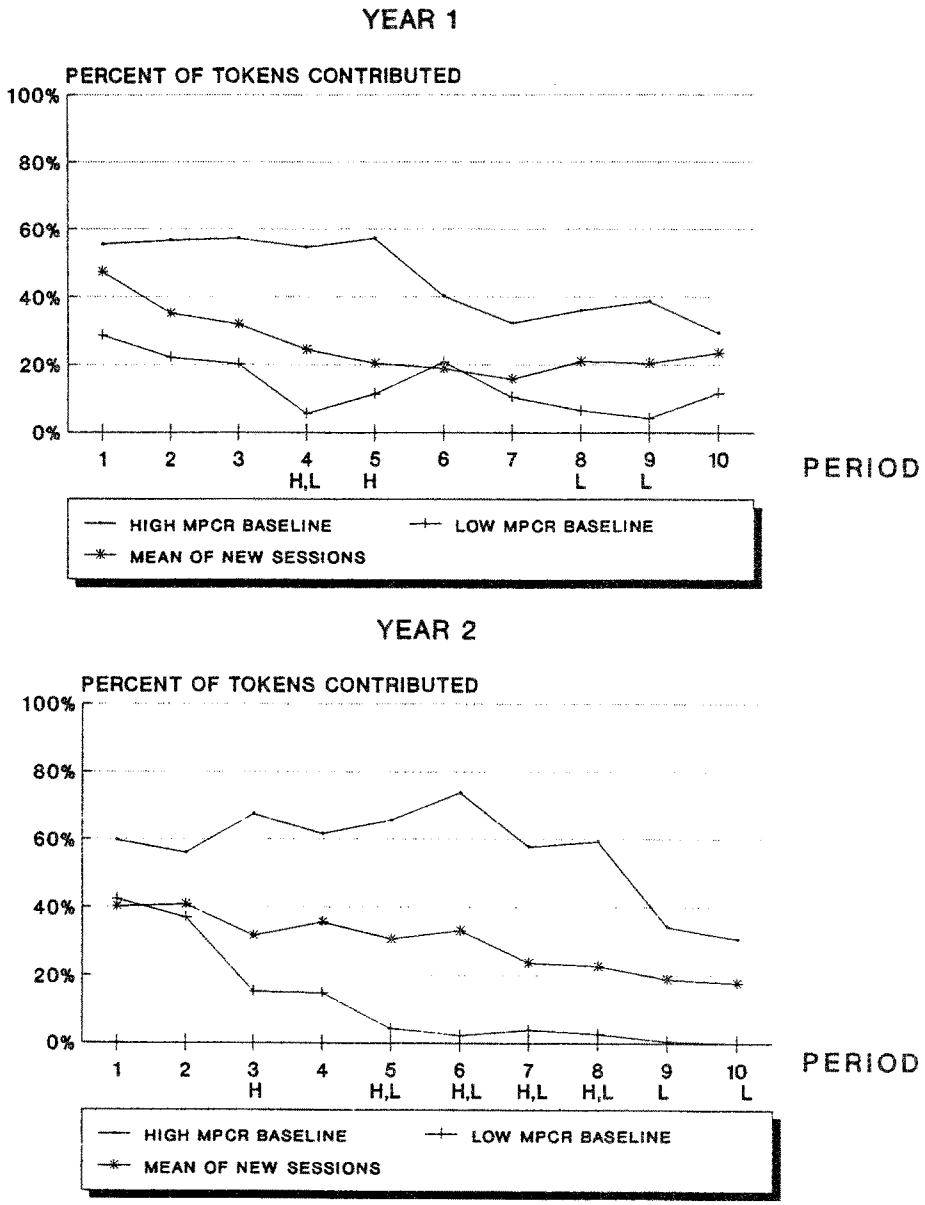


*Figure 1.* Contributions to the group good: Means of asymmetric groups compared to means of high and low MPCR baseline groups.

### 4.1. *Aggregate group decisions*

**Conclusion 1.** Hypothesis 3 best organizes the aggregate data. Specifically, the aggregate contributions to the group good from the heterogenous groups

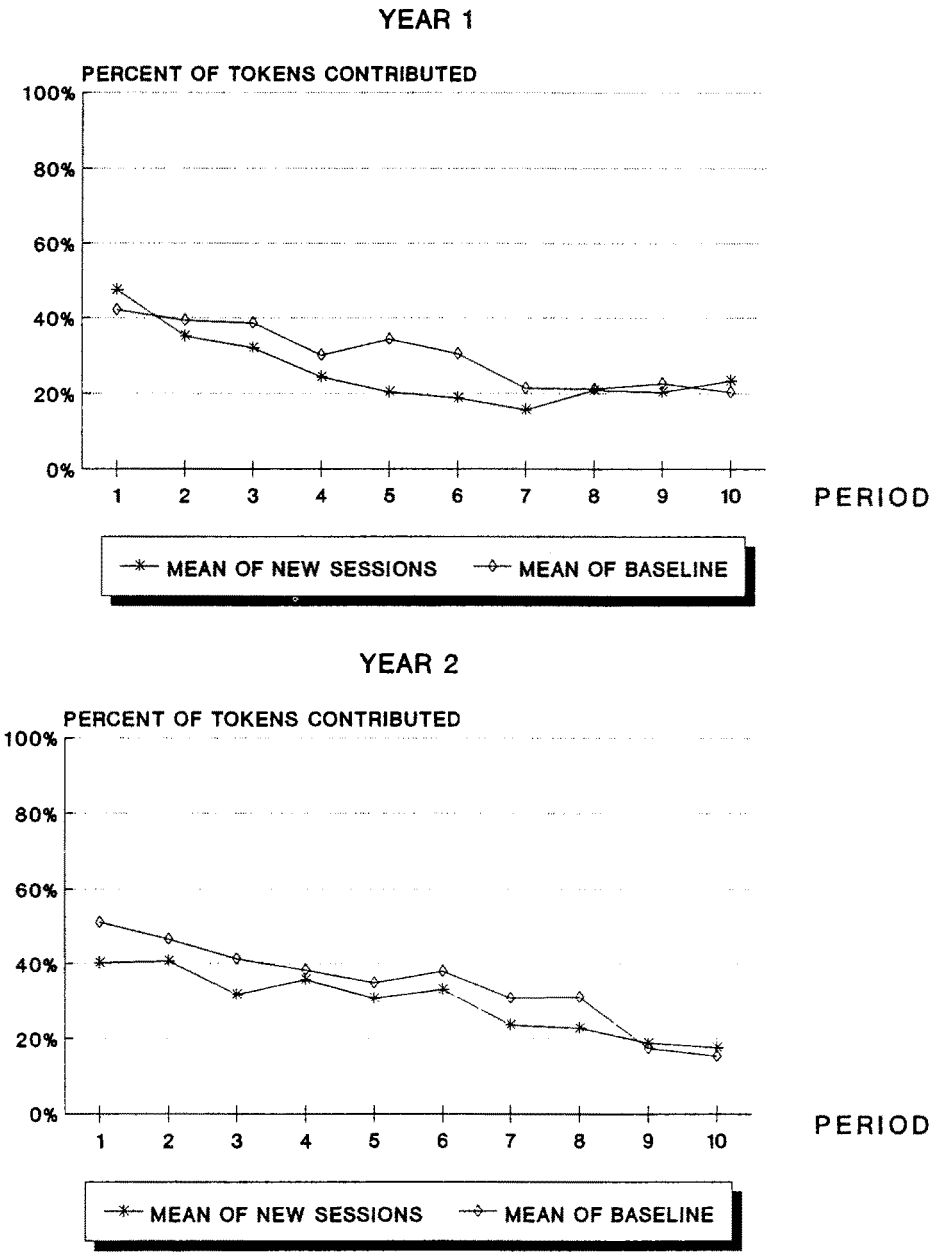


*Figure 2.* Contributions to the group good: Means of asymmetric groups compared to means of baseline groups.

resemble neither the extreme of the homogenous low or high MPCR groups, but rather are intermediate between the two extremes.

The essence of the group behavior hypotheses (Hypotheses 1–3) is whether a heterogenous group consisting of two low MPCR and two high MPCR types will provide the group good at a level that looks like the extremes of the homogenous groups or at an "intermediate" level. In Figure 1, we present group behavior related to these hypotheses. This figure displays the mean percentage of optimal group good provision obtained from the baseline symmetric groups of high MPCR type and low MPCR type, along with similar mean data from the heterogenous groups of this research. A letter "L" ["H"] indicates that the 90 percent confidence interval around the asymmetric group means is separated from the 90 percent confidence interval around the means from the homogenous low [high] baseline groups.

In general, the data take on an "intermediate" position. The mean of the heterogenous groups lies between the means of the homogenous groups in 9 of 10 periods in Year 1 and in 9 of 10 periods of Year 2. The basis for interpreting this intermediate position as representing distinct behavior is stronger in Year 2 where the mean from the heterogenous groups is statistically separate from at least one of the means from the homogenous groups in 7 of 10 periods. The same relationship can be seen in a different way in Figure 2, in which the mean data from the 20 new sessions is plotted against a pooled average of the data from both the high and low MPCR baseline sessions. In none of the periods are the means significantly different, at $\alpha = .10$.

### 4.2. *Data disaggregated by MPCR type*

**Conclusion 2.** Hypothesis 4 is generally not supported. Specifically, there appear to be differences in behavior between high and low MPCR types.

We begin an examination of the disaggregated data with Hypothesis 4, which relates to whether the two MPCR types will tend to behave similarly. The relevant data are presented in Figure 3, which compares the two groups across Year 1 and Year 2. A "*" beneath the period marker indicates that the 90 percent confidence intervals for the two group means are separated. These data address more clearly the Year 1–Year 2 differences noted above. In every period for both Year 1 and Year 2, the high MPCR types on average contributed more than the low MPCR types. (Recall that the high MPCR types of Year 1 are the low MPCR types of Year 2, and vice versa). However, the two groups are statistically distinct in only one of the ten periods of Year 1, while they are distinct in all ten periods in Year 2. While in a statistical sense we cannot reject the null hypothesis that the groups are the same (Hypothesis 4) in most periods of Year 1, there is strong support for the alternative hypothesis that the groups are different in every period of Year 2.

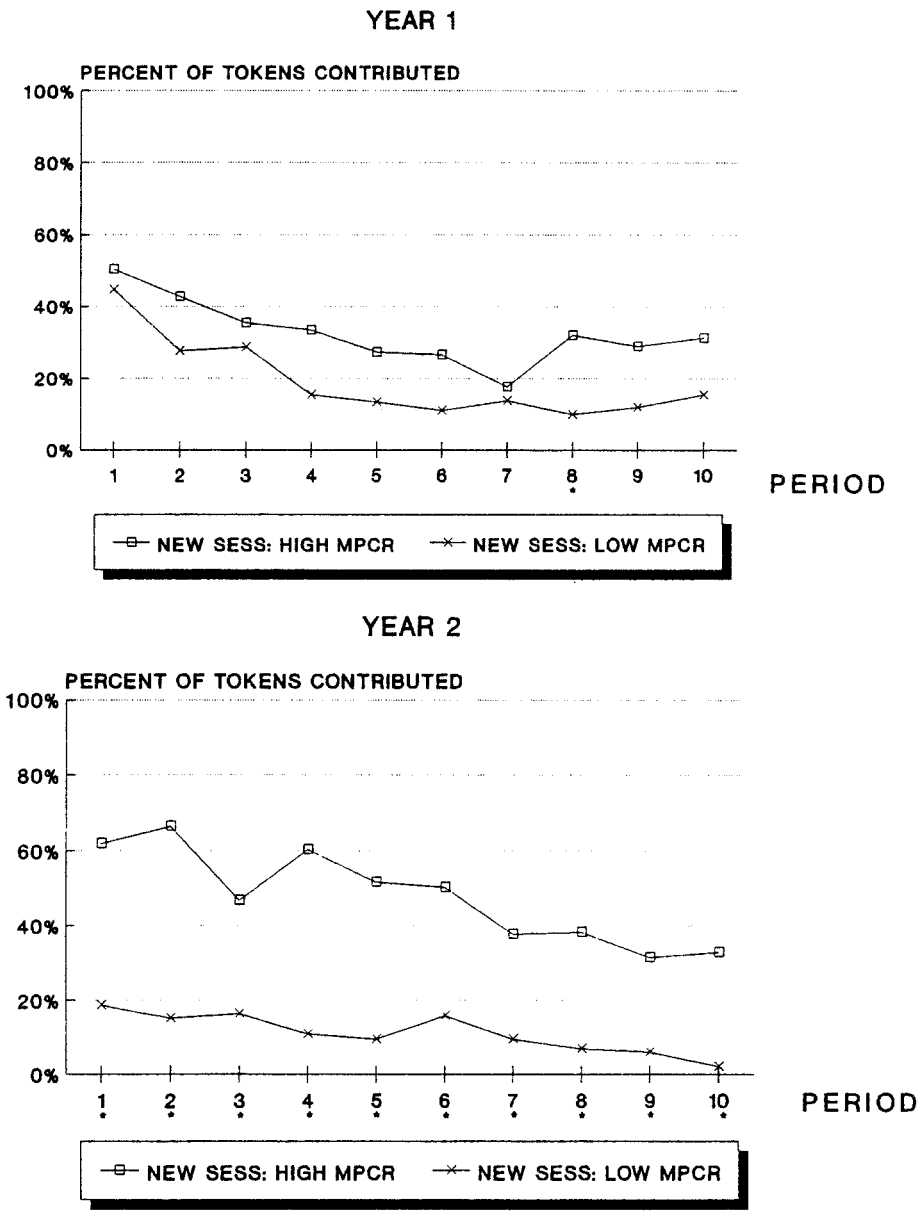


*Figure 3.* Contributions to the group good: Means of High MPCR types VS Low MPCR types.

Given that high and low MPCR types exhibit different behavior, particularly in Year 2, how does their behavior track the homogenous baseline groups? Unfortunately, the results are not clear cut.

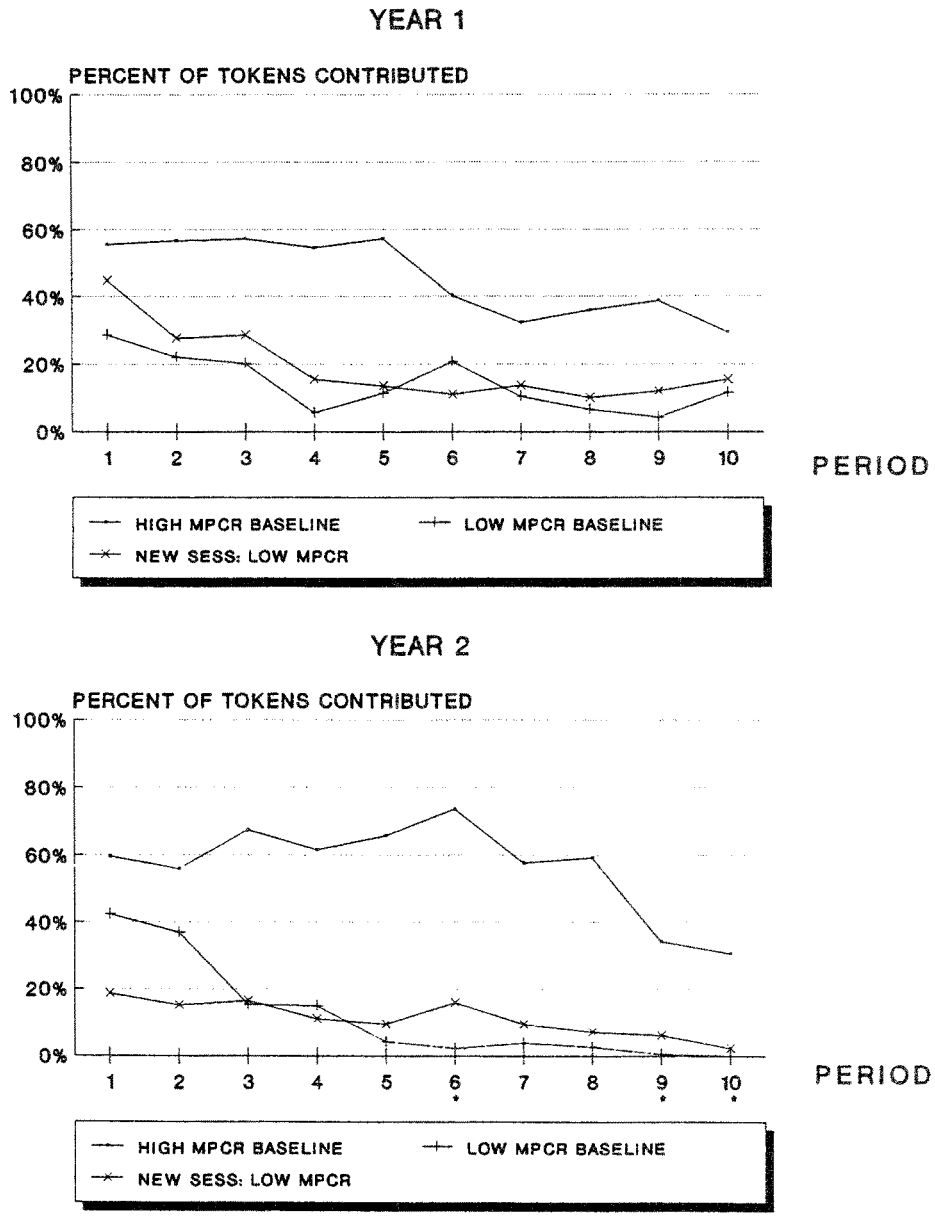


*Figure 4.* Contributions to group good Low MPCR types from asymmetric groups compared to High and Low MPCR types from baseline groups.

**Conclusion 3.** There is weak evidence evidence that when participating in mixed groups, high MPCR types show a tendency of lower contributions than do their equivalents in homogenous high MPCR groups. Likewise, there is

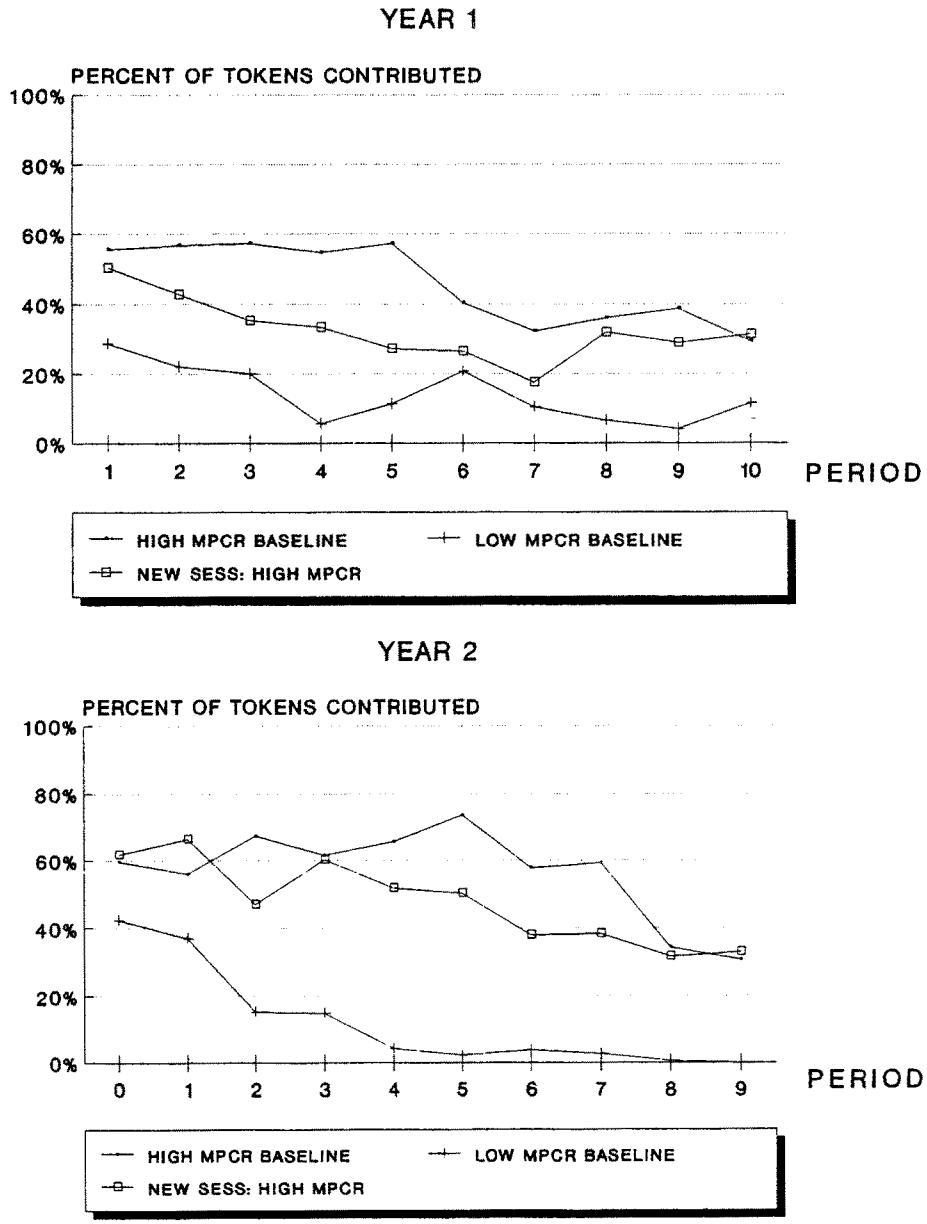


*Figure 5.* Contributions to group good High MPCR types from asymmetric groups compared to High and Low MPCR types from baseline groups.

some weak evidence that, when participating in mixed groups, low MPCR types show a tendency of higher contributions than do their equivalents in homogenous low MPCR groups.

For more insight on this question, we turn to Figures 4 and 5. Figure 4 (5) tracks the contribution decisions of the low (high) MPCR types as compared with the low and high MPCR homogenous baseline results. A "*" in Figure 4 indicates that there is a separation in the 90 percent confidence intervals between the means for low MPCR types in the asymmetric versus symmetric groups. Figure 5 carries analogous notation for the high MPCR types. Unfortunately, the conclusions regarding this disaggregate data are not as straightforward as in the prior discussion. For both the low and the high MPCR types, the results are somewhat ambiguous.

First, consider the data from the low MPCR types (Figure 4). Eyeballing the data does not suggest a large difference between the behavior of low MPCR types from the heterogenous groups and the low MPCR types from the baseline homogenous groups. In 15 of 20 periods, however, the mean is higher in the heterogenous group and that difference is statistically significant in 3 periods (these are three periods in which there is virtually no variance in the baseline data).

The data are similarly ambiguous for the case of the high MPCR types. In 16 of 20 periods the mean level of contribution is less for high MPCR types from the heterogenous groups relative to the high MPCR types from the baseline homogenous groups. However, in none of the 20 periods is this difference statistically significant.[10]

A similar situation can be observed by using a classification scheme for individual contribution decisions. Following Isaac, Walker, and Thomas (1984), call an individual a "strong free rider" if he contributes fewer than one-third of his tokens to the group good. We calculated the proportion of individuals acting as strong free riders for both the baseline homogenous groups and the new heterogenous group sessions, based upon the average of an individual's contributions across all periods in which (s)he is in a particular MPCR condition. Among high MPCR types, the proportion of strong free riders increased from 25 percent in the baseline to 37.5 percent in the heterogenous groups. Among low MPCR types, the proportion decreased from 91.6 percent in the baseline to 87.5 percent in the heterogenous groups. However, neither of these differences in proportions is statistically significant at the 90 percent level of confidence.

## 5. Interpretation and closing thoughts

The most striking feature of our results is the continued strong effect of an individual's own MPCR as an explanation for the level of contributions to the group good. Even in heterogenous groups, low MPCR types contributed less than high MPCR types.

These data also offer no strong evidence for any significant "poisoning of the well" or "seeding" model of heterogenous group behavior. On average, these mixed groups performed at an intermediate level compared to the homogenous baseline groups.[11] The preceding statement is consistent with the claim that other people's MPCRs do not matter, but that is not the only possible rationale. An intermediate level of aggregate contributions might result if high MPCR types lowered their contributions, and low MPCR types raised their contributions, relative to homogenous groups. Our data are ambiguous on this point. There is some directional tendency for the two types to move closer together, but this effect is not, in general, statistically significant.[12]

All in all, these experiments leave us impressed with the role of one's own MPCR as an influence on the individual incentive to contribute to public goods provision in the voluntary contributions mechanism. We are more skeptical than when we started about the blanket applicability of claims such as "seeding" or "poisoning of the well" which impute the influence of others' valuations.

## Notes

1. See Kim and Walker (1984), Isaac, Walker and Thomas (1984), Isaac, McCue, and Plott (1985), Isaac and Walker (1988a, 1988b), Andreoni (1988), Bagnoli and McKee (1991), Ledyard (1993) and Isaac, Walker and Williams (1994) for representative experimental work and reviews of the literature.
2. As a benchmark, the equilibrium concepts used here will be complete information models unless otherwise specified.
3. Each individual had a downward sloping individual valuation schedule for the public good. "Low" demanders had schedules which were everywhere lower than "high" demanders.
4. Recent new research by Palfrey and Prisbrey (1992) is also relevant to questions about heterogeneity, although with a different focus than we explore here. They investigate groups of size four in which the MPCR is heterogenous and varied for each person in each decision period. While the valuation on the public good is constant across individuals, each participant receives a new private good valuation drawn from a uniform distribution in each period. The distribution of values is known to all, but the actual draws in any one period are not. Palfrey and Prisbrey observe behavior that is much closer to the standard Nash models than most of the results reported here. Specifically, for those individuals for whom free riding is a single period dominant strategy, there is substantially less of the paradoxical positive contribution than has been reported elsewhere.
5. Despite the similarity in topics, the structure of these experiments is not designed to test the so-called "crowding out" models (Andreoni, 1993, reports on some "crowding out" experiments). Crowding out models have non-dominant strategy Nash equilibrium predictions. A change in the behavior of a dominant player (e.g., the government) alters the best responses of the other players. In the experiments reported here, complete "free riding" is always a single period dominant strategy.
6. Results from the experiments with inexperienced subjects are available from the authors.
7. It was emphasized that these earning would be paid in cash at the end of their participation.

8. In I&W experiments subjects were not explicitly informed of the homogeneity of valuations. Evidence suggests the subjects assumed this was true. In recent experiments Isaac and Walker (1993) examine whether explicit revelation of the homogeneity is important. There appears to be no significant impact, aggregating across decision periods.
9. In the I&W (1988a) paper, where the MPCR was altered for everyone in the group, I&W reported paired comparisons of individual decisions. We do not report such paired comparisons here because we do not believe they are appropriate in this different environment. A participant who switches from a low MPCR in Year 1 to a high MPCR in Year 2 also faces a change in the composition of the MPCRs of others (from one low MPCR and two highs to the reverse). We believe that further experiments specifically designed to look at "self" vs. "others" effects of MPCR are called for, but are beyond the scope of this study.
10. Statistical analysis of data from sequential periods in an experimental session is frequently problematic because of the potential non-independence of the observations. In the main text, we have addressed this by conducting a different statistical test for each period. However, we have also conducted two additional tests as a check upon the robustness of our conclusions. In one test, we took as an observation the difference between one category (say, new experiments, high MPCR types) and another in a given period. Then we performed a Wilcoxon rank sum test across all the period differences in a year. We also performed a t-test on the differences in means across all ten periods in a year.

    Neither of the two alternative tests qualitatively changes any of the reported conclusions. In general, with these two tests, there are fewer cases of insignificant differences. The two most important cases in which the change in significance is noted are the following. First, when using the difference tests for the data summarized in Figure 2, the new tests find more often a significant difference between the new data and the pooled data. This is quite consistent with the number of periods in which the mean of the new experiments is below the mean of the pooled baseline. Secondly, these two tests more often find a significant difference between the new high MPCR types and the old high MPCR types from the homogenous experiment baseline. Again, this is consistent with the number of periods in which this ordering occurs.
11. There is one piece of our data which, standing alone, provides some stronger support for a "poisoning-of-the-well" effect. After we had conducted the first 5 of our 10 groups, we noticed a curious phenomenon. The Year 2 results appeared almost exactly as the aggregate for all ten groups reported here. However, the Year 1 results showed some evidence of "poisoning" *for the high MPCR types only* (this is essentially the F&S result). This was *not* present in the Year 2 data. We were concerned that perhaps subjects did not carry over from the "training" experiments the possibility that values were heterogenous, but were reminded of it again before Year 2. However, if there was an inappropriate assumption of homogeneity present in these Year 1 periods, one could argue that the errors introduced could go in either direction. Therefore, we conducted the last 5 groups with subjects being explicitly reminded of the possibility of different private returns both before Year 1 and before Year 2. In the second 5 groups, the differential poisoning effect in Year 1 was substantially gone. This greater occurrence of poisoning type behavior in *only* Year 1 with *only* the high MPCR types in *only* the first five groups remains a mystery to us.
12. It is not obvious how our results mesh with those of Palfrey and Prisbrey (1992). Their introduction of randomly assigned MPCRs produced results substantially more in line with the standard Nash model. While the different purposes of the experiments makes a direct comparison difficult, the result that a heterogenous mix of MPCRs yields a pattern that looks less like our baseline and more like the standard Nash model is on the "poisoning of the well" side of the ledger. As Palfrey and Prisbrey point out, none of the current models offer a straightforward way to argue why group heterogeneity of MPCRs matters. Therefore, we have no easy explanation for the difference in our results.

## References

Andreoni, J. (1988). Why free ride? Strategies and learning in public goods experiments. *Journal of Public Economics* 37: 291–304.

Andreoni, J. (1993). An experimental test of the public goods crowding-out hypothesis. Mimeo. University of Wisconsin.

Bagnoli, M. and McKee, M. (1991). Can the private provision of public goods be efficient? Some experimental evidence. *Economic Inquiry* (April): 351–356.

Brookshire, D.S., Coursey, D.L. and Redington, D.B. (1988). Special interests and the voluntary provision of public goods. Mimeo. November.

Fisher, J. and Schatzberg, J. (1988). Asymmetric payoffs in prisoner's dilemma games. Mimeo.

Isaac, R.M., McCue, K. and Plott, C.R. (1985). Public goods provision in an experimental environment. *Journal of Public Economics* 26: 51–74.

Isaac, R.M. and Walker, J.M. (1988a). Group size effects of public goods provision: The voluntary contributions mechanism. *Quarterly Journal of Economics* 103: 179–201.

Isaac, R.M. and Walker, J.M. (1988b). Communication and free riding behavior: The voluntary contributions mechanism. *Economic Inquiry* 26: 585–608.

Isaac, R.M. and Walker, J.M. (1993). Nash as an organizing principle .... Mimeo.

Isaac, R.M., Walker, J.M. and Thomas, S. (1984) Divergent evidence on free riding: An experimental examination of some possible explanations. *Public Choice* 43: 113–149.

Isaac, R.M., Walker, J.M. and Williams, A.W. (1994). Group size and the voluntary provision of public goods: Experimental evidence utilizing very large groups. Forthcoming, *Journal of Public Economics*.

Kim, O. and Walker, M. (1984). The free rider problem: Experimental evidence. *Public Choice* 43: 3–24.

Ledyard, J.O. (1993). Public goods: A survey of experimental research. Mimeo. California Institute of Technology, August. Forthcoming in A. Roth and J. Kagel (Eds.), *The handbook of experimental economics*.

Marwell, G. and Ames, R. (1980). Experiments on the provision of public goods. II. Resources, interest, group size, and the free rider problem. *American Journal of Sociology* 85: 926–937.

Olson, M. (1971). *The logic of collective action*. Cambridge, MA: Harvard University Press.

Palfrey, T.R. and Prisbrey, J.E. (1992). Anomalous behavior in linear public goods experiments: How much and why? Mimeo. California Institute of Technology. September.