Journal of Public Economics 47 (1992) 85–106. North-Holland

Voluntary provision of public goods
The multiple unit case*

Mark Bagnoli

Indiana University, Bloomington, IN 47405, USA

Shaul Ben-David

University of New Mexico, Albuquerque, NM 87131, USA

Michael McKee

University of New Mexico, Alberquerque, NM 87131, USA

Received August 1989, revised version received October 1990

This paper reports the results of a series of experiments designed to test the predictions of a model of voluntary provision of public goods through private contributions. The particular voluntary contribution game implements the core in successively undominated perfect equilibria, but the behavioral question is whether the agents adopt strategies which support this refinement to the Nash equilibrium. The experimental evidence suggests that they do not: core allocations do not consistently occur in the laboratory markets.

1. Introduction

Economists have long held the belief that free riding will cause public goods to be undersupplied when funded with voluntary contributions. This suggested that, if public goods were to be provided at efficient levels, then the government's power to command payments through the tax system was needed. Such a pessimistic view was challenged by researchers who proposed several incentive-compatible mechanisms for public good provision [see Clarke (1971), Groves (1973), Groves and Ledyard (1977), Tideman and

*An earlier version of this paper was presented at the Economic Science Association Fall Meetings, October 1988, in Tucson. We wish to thank the participants of our session and Mark Isaac, in particular, for helpful comments. The exposition was enormously improved thanks to the careful comments of two anonymous referees. We also wish to thank Ray Prince for discussions during the design and conduct of the experiments and Roy Gardner for several helpful comments and advice. Funding was provided by the Council for Research and Creative Work at the University of Colorado and by Resources for the Future.

Tullock (1977), and Smith (1980) for early examples]. These mechanisms yield efficient outcomes but they are generally quite complex. Thus, they are sometimes viewed as an alternative to the tax system which the *government* may use to directly provide public goods not as private provision mechanisms. Recently, there has been a renewed search for more easily implemented mechanisms.

Among the simpler mechanisms that have been studied, both theoretically and experimentally, are voluntary contribution games. In a series of laboratory experiments, Isaac and Walker (1988) have investigated what they term the 'voluntary contribution mechanism' in which individuals voluntarily pay money into a group fund. The total return on the fund is divided equally among the members of the group independent of their individual contributions. This setting is essentially a prisoner's dilemma in which the dominant strategy is to contribute zero. They report a positive level of contributions in all of their sessions, although total contributions were always well below the Pareto efficient level.[1]

Other similar voluntary contribution game mechanisms are more successful in generating efficient allocations in the laboratory. Robyn Dawes and several of his colleagues [see, for example, van de Kragt et al. (1983) and Dawes et al. (1986)] have explored 'minumum contributing set' regimes in which each of the members of the group decide whether to contribute (a fixed amount) or not. If the number contributing meets or exceeds an announced threshold, then the public good is provided to the group, and all the members of the group receive a prescribed share of the good. Palfrey and Rosenthal (1984) analyzed one such game and found that it is capable of generating both efficient and inefficient outcomes.

Recently, Bagnoli and Lipman (1989) have investigated two contribution games in which the individuals decide whether to contribute *and* the level of their own contributions. Both contribution games involve posting voluntary contributions under the provision rule that the good (or the next unit of the public good) will be provided when the total group contributions meet or exceed a threshold – otherwise the contributions are refunded. In the first game, the group must decide whether or not to provide a public good. In the second, not only must the group decide whether or not to provide the public good, but also how much (number of units) to provide. Bagnoli and Lipman showed that their first contribution game fully implements the core in *undominated perfect equilibria*,[2] while the second game fully implements the

[1]Isaac and Walker found that the total contributions are strongly dependent on the marginal per capita return to contributions to the public good.

[2]This means that the set of core outcomes coincides with the set of undominated perfect equilibrium outcomes. The latter are the trembling hand perfect Nash equilibrium outcomes in the game after the removal of all dominated strategies.

core in *successively undominated strictly perfect equilibria* (SUSPE).[3] As a practical matter, the version in which the group must determine the level of provision is the more interesting. In this paper we report on the results of a series of laboratory experiments designed to evaluate Bagnoli and Lipman's predictions in this setting.

Laboratory experiments allow considerable control over incentives and the information provided to the subjects. Consequently, they are a very useful device for testing Bagnoli and Lipman's theoretical predictions. In this case the desirability of running experiments follows from two behavioral aspects of their mechanism. First, to implement the core they employed refinements of the Nash equilibrium and so one might wonder whether the particular refinements are appropriate. Second, even after employing a refinement, there are multiple equilibria and so one might also wonder whether or not the players can focus on one of them. Laboratory experiments have been shown to be particularly useful in addressing the broad question of the robustness of refinements to Nash equilibrium; our second objective for this paper.[4]

For the simple problem of whether or not to provide the public good, Bagnoli and Lipman's predictions are supported by the experimental results reported by Bagnoli and McKee (1991). For the more complicated problem of the level of provision (investigated in the current paper), our results suggest that one should not rely on the proposed mechanism. Based on our results, it appears that the refinement employed to implement core allocations is not an adequate representation of individual behavior. We also conclude that individuals have much more difficulty focusing on an equilibrium in the multiple units setting than they did in the single unit setting.[5] This result is not surprising since the multiple unit game is considerably more complex, admits many more equilibrium strategies, and the refinement

[3]Briefly, successively undominated means one strips out all of the dominated strategies, checks to see if this process has created more dominated strategies, and continues to remove these dominated strategies until all of the remaining strategies are undominated. One then seeks the strictly trembling hand perfect Nash equilibria to the reduced game. The difference between strictly trembling hand perfect and trembling hand perfect is that, for the latter, the strategies must be robust to some set of small trembles, while for the former the strategies must be robust to *all* sets of small trembles.

[4]There is a growing literature reporting on the use of laboratory experiments to evaluate refinements to the Nash equilibrium. For example, Camerer and Weigelt (1988) report on a series of laboratory experiments designed to test the predictive power of sequential equilibrium refinements. Other examples are provided by the work of Cooper et al. (1990) and van Huyck et al. (1990). It may be that the only place that such behavioral phenomena may be investigated is in the controlled environment of the laboratory.

[5]There is, of course, an alternative hypothesis to the failure of the mechanism. The subjects may have been unable to achieve the efficient outcome due to defects in the experimental design or implementation. We cannot rule out this possibility entirely but we have taken care to implement the theory as closely as possible in our experiments. There is also the possibility that the subjects would be more able to achieve core allocations with greater experience. In particular, it may be interesting to bring subjects back into the laboratory to participate in the same institution with different parameters. This is a topic for future research.

necessary to implement the core is much stronger than that used for the single unit setting. However, we do find that the subjects are able to achieve the predicted equilibrium with some degree of regularity even in this complex setting.

2. Theoretical foundations[6]

The basic structure of Bagnoli and Lipman's contribution game is relatively simple. Contributions to the provision of a public good are solicited from the agents in the economy and the public good is provided if the total contributions are sufficient to cover the cost. Otherwise the contributions are returned. The theoretical predictions have been developed for games of complete information. Thus, the cost of the good, the initial wealth of all agents, and the agents' valuations for the public good are common knowledge.[7] All agents are required to post their contributions to the public good without knowing the contributions made by the other agents.

It is convenient to describe the multiple unit contribution game in the context of an example using the parameters employed in one of our experimental sessions. Consider a group of five individuals each of whom is endowed with 55 tokens. The group is faced with the task of providing up to 4 units of a public good through voluntary contributions from the members of the group. The public good is produced at a constant per unit cost of 50 tokens. Each person in the group values the first unit of the public good at 20 tokens (aggregate valuation is 100 tokens), the second at 18, the third at 16, and the fourth at 10. All of this is common knowledge to the members of the group.

Only the first three units of the public good yield positive surplus to the

[6]In this section we provide the barest outline of the game investigated by Bagnoli and Lipman. The reader is referred to their paper for the full details.

[7]The level of information availability might seem excessive and such that no field applications of the contribution game mechanism could be forthcoming. Yet, Bagnoli and McKee provide anecdotal evidence of such contribution games being successfully employed to generate voluntary provision of public goods. For example, in 1985 the New Democratic Party (NDP) in Manitoba, Canada, sent letters to its larger contributors soliciting funds to mount a coming election campaign. The letters described those being canvassed (large donors), explained the issues in the coming election and the NDP's proposed policy stance, and explained how the money was to be used. Furthermore, the letters stipulated that a target had been set ($250,000) and that the NDP would refund all contributions if the target was not reached by a stated date. In terms of the information available to the potential donors, the announced threshold, and the provision rule, this campaign paralleled the contribution game setting. We would predict the level of contributions received to just meet the threshold. Total contributions were $251,300, only 0.5% over the target.

group. Thus, core allocations consist of three or four units of the public good being provided with total contributions of 50 tokens for each unit and no individual contributing more than his or her valuation for any unit.

Consider the following sequential structure. In the first round the agents each contribute some non-negative amount to the provision of the first unit of the good. If total contributions are less than the cost of the first unit, the contributions are returned and no units of the public good are provided. If the contributions meet or exceed the cost of the first unit, the group proceeds to the second unit. If the contributions sum to at least the cost of the second unit the group moves on to the third unit, and so on. The process stops when the contributions in a given round do not sum to the cost of an additional unit of the public good. Bagnoli and Lipman prove (their Theorem 2) that this structure implements the core in successively undominated strictly perfect equilibria and this is the structure of the multiple unit contribution game we provide in our laboratory setting.

While it is clear that the multiple unit contribution game is similar to the single unit game, there is a crucial difference. In the latter, the citizens need only decide whether or not to provide an exogenously fixed level of the public good while in the former, they must decide how much, if any, to provide. The problem of obtaining the efficient outcome is much more difficult and implementing the core requires a very strong refinement notion; successively undominated strictly perfect equilibria. As we have noted, the use of such a strong refinement immediately raises the question of its behavioral realism. Furthermore, since the equilibrium is not unique,[8] whether the agents can focus on a particular equilibrium is another open question. Finally, the additional complexity of the mechanism raises questions about its ability to present actual behavior. These questions are addressed by the empirical evaluation presented in this paper.

3. Experimental design

The experimental instructions (see appendix B) provide the subjects with the information prescribed by the theory; the threshold value (cost) for each unit, the incomes of the members of the group, and the valuations of the members of their group for all units. The continuation rule was explained as follows. The individuals were to post their contribution to the provision of the good at stage 1 (unit 1). If the sum of the contributions was not sufficient to cover the cost, then the contributions would be returned and the game

[8]Any vector of contributions summing to the cost and with no individual's contribution exceeding his valuation is an equilibrium. There are, of course, many such equilibria.

would end. If the sum met or exceeded the cost (threshold), the unit was supplied, and the group went on to the next unit and so on until the sum of the contributions in that stage was insufficient to have the unit supplied. For this unit the contributions were returned and the game ends. At each stage or unit the subjects are informed of the total contributions of their group, whether the unit has been provided, and their own remaining income for the round.

In the single unit case reported in Bagnoli and McKee, subjects posted contributions to the provision of one unit of the public good. The current experimental setting repeats the single unit game over a sequence. An equilibrium to this game, satisfying the refinement imposed by Bagnoli and Lipman, has a succession of decision rounds with one additional unit purchased at each round and with contributions adding to exactly the marginal cost. In this equilibrium, units are added so long as the group valuation exceeds the marginal cost.

For all sessions, there was a maximum of 4 units of the public good that could be provided and the parameters were set such that the core allocation provided for 3 or 4 units of the good. Subjects were assigned to groups of five persons in such a manner that they could not learn the identity of the others in their group. To investigate the effects of heterogeneity among the group members, we employed several income and valuation distributions. In the scrambling sessions we assigned all individuals the same incomes but the valuations for the public good differed across individuals. The experimental design parameters regarding incomes, valuations, and thresholds (or costs) are given in table 1.

In many experimental settings repetition is necessary to allow the subjects to 'learn the game' [see Isaac, McCue and Plott (1985)]. Repetition is accomplished by having the subjects face the same decision over several replications of the experimental setting. In addition to permitting learning, repetition allows one to investigate conditions that affect the speed of convergence to equilibrium by observing behavior over several rounds of an experiment. For example, in the single unit setting, Bagnoli and McKee found that convergence was slower for groups comprised of individuals having unequal incomes or valuations of the public good than for groups in which all members had the same income and valuation.

With repetition comes a choice of experimental designs. One can assign the subjects to the same group for the duration of the experimental session, in which case the subjects must be viewed as playing a finitely repeated game. Alternatively, one can scramble the subjects among the groups prior to the start of each period, in which case the subjects are playing a series of one-shot games.

We elected to conduct two series of experiments, one in which the subjects are reassigned to different groups between periods (denoted the S – for

Table 1

Experimental design parameters.
Initial data: Incomes and valuations

The distributions of incomes and valuations are generated within the software according to the following mapping:

Income distributions
(shares of total income – 275 tokens per group)

	Subject number				
Distribution	1	2	3	4	5
(a)	0.20	0.20	0.20	0.20	0.20
(b)	0.28	0.28	0.15	0.15	0.14
(c)	0.28	0.24	0.20	0.14	0.14
(d)	0.22	0.22	0.22	0.22	0.14

Payoff distributions
(shares of total payoff for each unit of the public good)

	Subject number				
Distribution	1	2	3	4	5
(x)	0.20	0.20	0.20	0.20	0.20
(y)	0.30	0.30	0.15	0.15	0.10
(z)	0.30	0.20	0.20	0.20	0.10

In all sessions there are two or three groups, each having five individuals, participating simultaneously. For a session, the groups may be identical or they may differ in the distribution of income or valuations.

In all cases the total income for a group of five subjects was set at 275 tokens and the cost of each unit of the public good was set at 50 tokens. The income–payoff combinations are selected as initial conditions by the experimenter. The following design was used in the experiments conducted:

Income/payoff combination no. 3 is comprised of income distribution (a) with payoff distribution (z).

Income/payoff combination no. 4 is comprised of income distribution (a) for group 1 of the session, (b) for group 2, and (c) for group 3 with payoff distribution (x).

G1 no scrambling (NS), G2 NS, and G3 NS used income/payoff combination no. 4.

G4 NS, G5 NS, and G6 NS as well as all scrambling (S) sessions used income/payoff combination no. 3.

The aggregate valuations of the units for all sessions were: unit 1 – 100 tokens; unit 2 – 90 tokens; unit 3 – 80 tokens; unit 4 – 50 tokens.

scrambling – series) and one in which the subject remained in the same group for the entire session (denoted the NS – for no scrambling – series).[9]

Since Bagnoli and Lipman studied a single-shot game and since without scrambling the subjects must be viewed as playing a finitely repeated game,

[9]We wish to express our thanks to an anonymous referee whose comments led us to a more careful exposition of the case for our use of both scrambling and no scrambling sessions.

technically only the final round of the NS experiment is a test of the theory as developed. Therefore the experiments in our S series and the final round of the NS series may be considered as being directed Roth's (1987) first motive – testing the predictions of established theory.

Roth's second motivation for experiments is that they provide observations on behavior for which the theory may be silent. Such observations may be used to direct future theoretical investigations. The experiments in our NS series are also designed for this purpose. In finitely repeated games, one subgame perfect equilibrium has the players playing the one-shot equilibrium in each period. By scrambling the subjects, one cannot learn whether the subjects playing the game are playing the one-shot equilibrium over and over. The no scrambling (NS) sessions may shed light on this issue and provide suggestions for further theoretical work.

Our use of scrambling should not be considered an experimental treatment in the sense of, say, Andreoni (1988) since we did not hold all other conditions of the experiment constant between the no scrambling and the scrambling sessions. In particular, the sessions in which we scrambled subjects were conducted for six to eight periods (plus two practice periods) whereas the no scrambling sessions lasted 15 periods. Furthermore, the time taken to scramble the subjects between rounds could be argued to have provided the subjects with additional time to think about their strategies.

All experimental sessions were conducted in the LEAP (Laboratory for Economics and Psychology) facility at the University of Colorado. This facility consists of a dedicated MicroVAX and 16 terminals (one for the monitor). The terminals are located in booths which prevent the subjects from observing their neighbors' screens. Subjects were recruited from principles and intermediate economics classes. The instructions were read aloud while the subjects followed along on their copies and questions were answered before the session began. With the payoff structure we utilized, participants would earn approximately \$1.30 per period if no units were supplied and upwards of \$2.45 if 3 units were supplied. In practice, the subjects earned between \$12.00 and \$20.00 for their participation in a 1-hour session and appeared to be quite highly motivated.

4. Hypotheses and empirical evaluation

All groups contained five individuals. We have data on 11 groups from the scrambling (S) sessions and six groups from the no scrambling (NS) sessions. The round-by-round contributions for the groups are reported in appendix A, while the summary data are reported in table 2. The data for the individual subjects are available from the authors.

Bagnoli and Lipman predict that the players will adopt strategies that

Table 2

Average group contributions and frequency of provision of 3 units.[a]

	Group contributions				
	Units				Frequency of provision
Groups	1	2	3	4	of 3 units
G1 S	57.19	51.90	50.87	–	3 of 8
G2 S	59.67	54.93	51.81	–	4 of 8
G3 S	57.17	52.40	51.50	–	3 of 6
G4 S	54.00	54.38	51.67	–	5 of 6
G5 S	53.80	54.67	53.55	–	4 of 6
G6 S	56.29	52.07	52.12	–	4 of 7
G7 S	55.28	54.13	48.96	–	2 of 7
G8 S	54.93	52.67	49.83	–	2 of 7
G9 S	54.39	51.08	52.23	–	3 of 8
G10 S	50.13	52.86	52.25	–	3 of 8
G11 S	53.86	52.81	52.75	–	5 of 8
All S	55.10	53.11	51.68		
Last period only					
G1 NS	54.00	60.00	48.00	–	no
G2 NS	56.50	58.25	53.50	39.00	yes
G3 NS	55.30	52.70	54.50	49.50	yes
G4 NS	51.00	51.00	49.00		no
G5 NS	61.50	54.50	64.00	34.20	yes
G6 NS	51.00	52.00	53.00	39.50	yes
All 15 periods					
G1 NS	53.01	44.42	–	–	2 of 15
G2 NS	52.15	53.97	52.26	–	9 of 15
G3 NS	55.02	54.02	51.87	49.92	9 of 15
G4 NS	51.67	51.00	51.33	48.10	10 of 15
G5 NS	57.00	55.96	57.00	46.81	12 of 15
G6 NS	51.33	52.32	52.25	–	8 of 15
All NS	53.37	52.20	52.99	–	50 of 90

[a]In the S sessions the composition of the groups was altered at the start of each round. Since each such group played the game only once (no group stayed together for more than one period), we actually have individual observations for 79 groups. We report the results as shown for ease of reference and so that the reader may discern any learning that has occurred since when a group reaches period 3 *all* members of that group have two previous trials.

produce core allocations. That is, the allocations that result will be Pareto optimal and individually rational. Explicit testable hypotheses (given the parameters of our design) are:

Hypothesis 1. The groups will provide 3 units of the good in each period.

Hypothesis 2. The contributions in each stage will sum to the threshold (announced as 50 tokens for each unit or stage).

Hypothesis 3. The agents' contributions will be individually rational. In the context of successively undominated strictly perfect equilibrium, this requires that individual contributions be less than the valuation of the good at *each* stage, or unit.

Hypothesis 1 constitutes a fairly weak version of the predictions embodied in the theory since it is possible to accept this hypothesis and have no allocations that are in the core because the aggregate contributions exceed the cost of the good.

In general, the data (table 2 and appendix A) provide weak support for Hypothesis 1. That is, the groups are moderately successful in providing the good at efficient levels.[10] The data from the S sessions indicate that three units are supplied in 38 of 79 possible cases (48%). There are seven additional cases where the contributions for the third unit are between 49 and 50 tokens, which is very close to the threshold. If, to allow for learning, we focus on the data from repetition 4 on, the results are marginally better. Now we have 3 units supplied in 23 of 46 (50%) observations with an additional three observations in which the total contributions to the third unit are between 49 and 50 tokens.

For the final period of the NS sessions we find that four of the six groups are successful at having 3 units of the good provided and that a fifth group contributed 49 tokens to the third unit.

Core allocations were much more evident in the single unit setting reported by Bagnoli and McKee. In that setting the good was provided by all seven of the five-person groups.

Hypothesis 2 describes the SUSPE equilibrium for the contribution game setting with our parameters. The data reported in appendix A show that the groups were not, in general, able to achieve the predicted equilibrium. For the S sessions we observe the contributions summing to 50 (the announced threshold) in only 5 of 237 observations. For the last period of the NS sessions we observe the contributions do not sum to 50 (the announced threshold) for any of the six groups.

In any economic setting there are many coordination problems that may make it difficult to achieve the theoretically predicted outcome exactly. If we take 47.5–52.5 as a range where the aggregate contributions are 'close' to the predicted equilibrium, then for the scrambling sessions we find the total contributions are within this band in 60 of 237 (25%) observations. For the

[10]There is no implication, in the theory, that *all* members of the group will post positive contributions. If the threshold will be met without agent j's contribution, then the individually rational action of j is to contribute zero.

last period of the scrambling sessions the total contributions fall within the band in 6 of 18 observations (33%).

As is frequently the case, the average level of contributions over all periods of the scrambling sessions is closer to the prediction than individual results. The respective averages are (standard deviations in parentheses): unit 1 – 55.10 (5.39); unit 2 – 53.11 (3.96); unit 3 – 51.68 (3.99). In all cases the average total contributions are not significantly different from 50 tokens. If we look at the last period of the NS sessions the average total contributions (over six observations) are: unit 1 – 54.88 (3.94); unit 2 – 54.74 (3.62); unit 3 – 53.67 (5.69).

An observation that the subjects are playing dominated strategies and/or strategies that are not strictly perfect is sufficient to refute Hypothesis 3. A strategy that is clearly not individually rational is to contribute more than one's valuation for each unit. Less obviously, another dominated strategy would be contributing in excess of one's valuation in an early stage (unit) in an attempt to move the group on to later stages where this overcontribution can be recouped. We find no evidence of either behavior. However, in round 4, subject 5 in Group 1 of the NS series started to play a strategy that involved posting a very large contribution (not quite his valuation) for the first unit and then posting a contribution of zero for the next unit. Interestingly, the group was unable to adjust to this behavior, with the result that it was only able to have the second unit supplied in one round after round 3. In general, we find our data support Hypothesis 3; the subjects are playing strategies that are individually rational.

An alternative way to see if the outcomes are close to the prediction is to examine group welfare. We have computed the level of group welfare attained by each group (table 3). For the last period the NS groups attained 75% of the theoretical welfare level or better, with the exception of G1 NS.[11] The results for the scrambling sessions, when we consider all periods, are less robust. Here the welfare levels are closer to 50% of the theoretical level with the exception of G11 S. If we focus only on period 4 and later, the results for the S series are much stronger. The groups attained better than 68% of the theoretical welfare levels.

In the single unit, unscrambled experiments reported by Bagnoli and McKee, we inferred that the equilibrium selected appeared to involve playing the same one-shot game equilibrium at each stage or round. For the NS series reported here, the behavior appears to be fairly stable if one looks to the aggregate data. For example, the mean total contributions by stage (unit) for the no scrambling sessions are (standard deviations in parentheses): unit

[11]This group contained the individual discussed earlier. Furthermore, in G1 NS the group was comprised of individuals with different payoffs to the public good. From the work of Bagnoli and McKee it appears that this type of heterogeneity is difficult for the group to deal with.

Table 3

Welfare levels.

	Scrambling sessions						
Round	G1 S	G2 S	G3 S	G4 S	G5 S		Theory
1	40.0	49.0	75.0	85.5	84.0		120.0
2	73.0	61.25	104.0	107.5	103.0		120.0
3	95.0	109.45	105.0	111.5	101.9		120.0
4	80.5	79.0	–	111.0	81.5		120.0
5	–	79.0	108.0	112.5	108.0		120.0
6	75.5	114.0	41.0	110.7	108.0		120.0
7	106.0	83.0	n/a	n/a	n/a		120.0
8	115.5	96.0	n/a	n/a	n/a		120.0
	Scrambling sessions						
Round	G6 S	G7 S	G8 S	G9 S	G10 S	G11 S	Theory
1	71.9	101.8	–	–	–	107.0	120.0
2	105.0	–	95.5	109.5	83.0	84.0	120.0
3	47.0	110.5	81.0	76.5	–	98.0	120.0
4	105.1	79.0	110.5	49.0	–	110.5	120.0
5	107.5	86.8	44.5	40.0	107.5	114.0	120.0
6	49.0	86.5	49.5	–	110.0	89.0	120.0
7	108.5	82.0	39.5	96.3	–	83.0	120.0
8	n/a	n/a	n/a	107.7	113.0	101.0	120.0
	No scrambling sessions						
Round	G1 NS	G2 NS	G3 NS	G4 NS	G5 NS	G6 NS	Theory
1	83.0	113.5	73.0	82.0	68.5	49.0	120.0
2	107.0	115.0	–	109.0	76.0	107.5	120.0
3	101.0	114.5	97.0	107.0	90.0	115.0	120.0
4	–	114.0	107.0	111.0	99.5	–	120.0
5	46.0	–	112.5	–	108.0	82.5	120.0
6	43.0	104.8	115.0	116.0	104.0	109.5	120.0
7	50.0	–	86.0	115.0	113.5	87.0	120.0
8	45.0	45.5	112.3	–	105.0	111.5	120.0
9	50.0	67.6	86.0	115.0	108.0	–	120.0
10	48.0	80.8	84.5	115.0	–	–	120.0
11	–	109.3	109.5	116.0	94.5	111.0	120.0
12	–	107.5	104.5	118.0	78.0	112.0	120.0
13	36.0	110.6	86.5	48.0	95.0	–	120.0
14	45.0	–	–	119.0	93.0	111.0	120.0
15	76.0	101.8	107.5	89.0	89.2	114.0	120.0
Sum 1	730.0	1184.7	1281.3	1360.0	1241.2	1110.0	1800.0
Sum 2	157.0	429.1	408.0	490.0	449.7	448.0	600.0

Note: Sum 1 refers to the aggregate over all 15 periods. Sum 2 refers to the aggregate over the last 5 periods.

1 – 53.37 (5.77); unit 2 – 52.2 (6.22); unit 3 – 52.99 (4.79). These values are not statistically different from the predicted equilibrium of 50 tokens. However, if we look at the period-by-period data in appendix A, there is

considerable volatility in the individual and group behavior. Even omitting G1 NS for reasons cited above, it would appear that in the no scrambling setting the subjects do not seem to be playing the same one-shot game equilibrium in each round. For example, there are periodic instances of subjects apparently 'experimenting' with different strategies during the no scrambling sessions and this is reflected in the period-by-period welfare levels reported in table 3.

In conclusion, our results suggest that in the multiple unit setting the subject behavior is not well captured by the stronger refinement. This contrasts with the behavior of the subjects in the single unit setting who played the strategies consistent with the invoked refinement – undominated perfect equilibria (Bagnoli and McKee).

5. Conclusions and remarks

We find limited support for the prediction that the agents will play equilibrium strategies that achieve a core allocation. Overall, the rate of success is much less than in the single unit experiments reported by Bagnoli and McKee. Indeed, from the results reported above it is clear that the multiple units game is much less likely to implement the core than the single unit game. Thus, our results raise serious questions about the predictive success of Bagnoli and Lipman's Theorem 2. For the multiple unit case, it does not appear that the contribution game they studied is likely to generate core allocations or even something close.

From a policy perspective this is a disappointing result. Since many of the other incentive compatible mechanisms have substantial administrative requirements and, in some cases, generate budget surpluses to be disposed of in a manner that will not affect the allocation of resources, the continued search for a voluntary provision scheme seems worthwhile. Our experiments show that Bagnoli and Lipman's multiple unit scheme, which *theoretically* yields efficient outcomes, appears not to work well in practice. Furthermore, the difference in practical usefulness between Bagnoli and Lipman's single and multiple unit games raises questions for the class of games investigated by Dawes and others. In previous work these latter mechanisms were applied to the single unit case and, to date, these games have not been evaluated in multiple unit settings. This is a worthy topic for future research.

It remains for us briefly to investigate whether some features of the experimental design are responsible for the subjects' failure to achieve core allocations. We begin by noting that our design has implemented the information and incentives prescribed by the theory. In addition, the monetary incentives appear to be salient by all objective criteria.

While the design used in the current paper parallels that used by Bagnoli and McKee in the single unit setting, where the subjects were very successful

in reaching efficient outcomes, there are some differences between the single unit setting and the multiple unit setting.

The subjects engaged in the NS sessions were required to make many more decisions than those in the single unit setting of Bagnoli and McKee. Each *unit* requires the same decision as each *stage* of the single unit setting. In addition, the strategy space is much larger in the multiple unit setting. Both of these factors may have led to the 'experimentation' reported above. Overall, the groups were least successful in the middle rounds of the NS sessions (table 3 and appendix A). Indeed, the average welfare level, for periods 1 through 5 was 79.62; for periods 6 through 10 was 73.93; and for periods 11 through 15 was 79.29. The welfare levels in the middle rounds are statistically lower at the 95% level or better. It appears that after some experimentation, which led to failure to implement the efficient outcome, the groups returned to their previous equilibrium strategies.

It is possible that subjects with prior experience in the multiple unit setting would be more successful since experienced subjects would be likely to have learned that experimentation is not fruitful. We leave this investigation as a possible avenue for future research. Indeed, the simplicity of the contribution game mechanism provides an appealing justification for future research into the properties of this mechanism.

Appendix A

Table A.1

Aggregate group contributions – no scrambling sessions.

	G1 NS				G2 NS			
	Units				Units			
Period	1	2	3	4	1	2	3	4
1	54.00	53.00	47.00	–	51.25	54.45	50.75	55.00
2	52.00	54.00	57.00	59.00	50.25	53.25	51.00	40.00
3	59.00	56.00	54.00	37.00	53.25	51.25	52.00	42.50
4	49.00	–	–	–	52.00	50.75	53.25	31.00
5	54.00	42.00	–	–	48.75	–	–	–
6	57.00	43.00	–	–	53.50	54.75	57.00	32.50
7	50.00	33.00	–	–	47.25	–	–	–
8	55.00	31.00	–	–	54.50	44.00	–	–
9	50.00	39.00	–	–	56.79	65.66	49.70	–
10	52.00	30.00	–	–	55.34	53.91	41.25	–
11	48.00	–	–	–	50.75	56.00	54.00	35.50
12	42.00	–	–	–	51.35	52.90	58.25	26.00
13	64.00	44.00	–	–	52.74	52.51	54.15	25.10
14	55.00	48.00	–	–	48.00	–	–	–
15	54.00	60.00	48.00	–	56.50	58.25	53.50	39.00
Mean	53.01	44.42	–		52.15	53.97	52.26	–
std.	5.11	10.05			2.92	5.05	4.46	
	G3 NS				G4 NS			
	Units				Units			
Period	1	2	3	4	1	2	3	4
1	89.00	52.50	55.50	44.50	55.00	53.00	49.00	–
2	47.50	–	–	–	55.00	53.00	53.00	41.00
3	70.50	51.50	51.00	44.50	56.00	53.00	54.00	45.00
4	53.00	58.00	52.00	46.00	54.00	51.00	55.00	38.00
5	51.00	55.50	52.00	55.00	47.00	–	–	–
6	52.00	53.00	50.00	54.00	50.00	53.00	51.00	55.00
7	51.00	53.00	49.50	–	51.00	52.00	52.00	51.00
8	52.00	53.50	52.20	52.80	49.00	–	–	–
9	51.50	52.50	48.60	–	52.00	53.00	50.00	50.00
10	50.50	55.00	47.50	–	52.00	51.00	52.00	50.00
11	50.50	54.00	56.00	49.00	51.00	52.00	51.00	50.00
12	51.50	58.00	56.00	49.00	51.00	51.00	50.00	49.00
13	50.50	53.00	49.50	–	52.00	39.00	–	–
14	49.50	–	–	–	50.00	51.00	50.00	52.00
15	55.30	52.70	54.50	49.50	51.00	51.00	49.00	–
Mean	55.02	54.02	51.87	49.92	51.67	51.00	51.33	48.10
std	10.76	2.06	2.88	4.22	2.47	3.72	1.92	5.22

Table A.1 (continued)

	G5 NS				G6 NS			
	Units				Units			
Period	1	2	3	4	1	2	3	4
1	61.00	60.50	45.00	–	51.00	46.00	–	–
2	66.50	65.00	62.50	46.00	53.00	57.00	52.50	48.00
3	59.50	56.50	65.00	61.00	52.00	53.00	50.00	48.00
4	61.00	54.50	56.00	39.50	49.00	–	–	–
5	55.00	56.00	51.00	48.00	55.00	52.50	49.00	–
6	54.50	56.50	55.00	50.00	53.00	53.00	54.50	45.00
7	52.50	50.00	54.00	57.50	51.00	52.00	48.00	–
8	53.00	53.00	59.00	52.00	52.00	50.00	56.50	41.00
9	51.00	50.00	61.00	50.00	49.00	–	–	–
10	49.50	–	–	–	49.00	–	–	–
11	63.00	52.00	60.50	41.50	54.00	52.00	53.00	43.00
12	59.50	53.50	41.00	–	52.00	53.00	53.00	37.00
13	56.50	59.50	60.00	44.00	48.00	–	–	–
14	51.00	62.00	64.00	38.00	51.00	55.00	53.00	29.00
15	61.50	54.50	64.00	34.20	51.00	52.00	53.00	39.50
Mean	57.00	55.96	57.00	46.81	51.33	52.32	52.25	
std.	5.11	4.45	7.25	7.93	1.99	2.76	2.56	

Notes:

(1) 'Units' refer to the units (1 through 4) of the public good.

(2) The figures in the tables are the total contributions by the respective group for each unit of the public good. A '–' indicates the group did not reach this unit, i.e. the group had failed to reach the threshold at an earlier unit.

Table A.2

Aggregate group contributions – scrambling sessions.

	G1 S				G2 S			
	Units				Units			
Period	1	2	3	4	1	2	3	4
1	60.00	40.00	–	–	51.00	49.50	–	–
2	60.00	57.00	48.00	–	74.50	54.25	47.20	–
3	60.00	57.00	58.00	58.00	63.05	59.50	58.00	51.00
4	58.00	51.50	49.50	–	55.00	56.00	48.00	–
5	45.00	–	–	–	60.50	50.50	45.50	–
6	63.00	52.50	45.00	–	52.30	50.20	53.50	35.00
7	58.00	53.00	53.00	50.00	67.00	61.50	58.50	43.00
8	53.50	52.30	51.70	43.50	54.00	58.00	52.00	43.00

Table A.2 (continued)

	G3 S				G4 S			
	Units				Units			
Period	1	2	3	4	1	2	3	4
1	64.00	51.00	49.00	–	53.50	51.00	49.00	–
2	59.00	56.00	51.00	19.50	57.00	53.50	52.00	–
3	55.00	56.00	54.00	31.00	53.00	52.50	53.00	48.50
4	48.00	–	–	–	53.00	54.00	52.00	55.50
5	58.00	52.00	52.00	42.00	52.00	56.50	50.00	46.00
6	59.00	47.00	–	–	55.50	58.80	54.00	48.00
	G5 S				G6 S			
	Units				Units			
Period	1	2	3	4	1	2	3	4
1	53.00	53.00	49.50	–	62.50	55.60	46.10	–
2	54.50	53.50	59.00	48.50	61.00	51.50	52.50	29.50
3	56.30	53.00	59.80	63.00	53.00	48.00	–	–
4	54.00	54.50	49.00	–	53.00	54.90	57.00	35.00
5	52.50	57.50	52.00	13.50	56.00	52.50	54.00	28.50
6	52.50	56.50	52.00	18.00	51.00	49.00	–	–
7					57.50	53.00	51.00	12.00
	G7 S				G8 S			
	Units				Units			
Period	1	2	3	4	1	2	3	4
1	62.00	61.50	54.75	45.75	48.00	–	–	–
2	44.50	–	–	–	64.00	61.50	50.00	15.00
3	54.25	53.75	51.50	34.50	50.50	58.50	47.50	–
4	58.50	52.50	44.50	–	55.50	52.00	52.00	10.00
5	52.70	50.50	47.50	–	55.50	46.00	–	–
6	52.50	51.00	46.50	–	50.50	49.50	–	–
7	62.50	55.50	49.00	–	60.50	48.50	–	–
	G9 S				G10 S			
	Units				Units			
Period	1	2	3	4	1	2	3	4
1	46.00	–	–	–	44.00	–	–	–
2	56.50	50.00	54.00	36.00	55.00	52.00	48.50	–
3	59.00	54.50	45.00	–	45.50	–	–	–
4	51.00	45.50	–	–	49.00	–	–	–
5	60.00	45.00	–	–	52.50	56.50	53.50	30.40
6	48.00	–	–	–	52.00	53.00	55.00	33.00
7	58.90	57.30	57.50	27.00	48.00	–	–	–
8	55.70	54.20	52.40	27.00	55.00	50.00	52.00	10.00

Table A.2 (continued)

	G11 S			
	Units			
Period	1	2	3	4
1	54.00	50.00	59.00	45.00
2	55.00	51.00	47.00	–
3	59.00	56.00	57.00	45.00
4	53.50	53.00	53.00	42.00
5	51.00	52.00	53.00	45.00
6	50.50	50.50	49.00	–
7	53.00	54.00	46.00	–
8	55.00	56.00	58.00	–

Note: In the S sessions the composition of the groups was altered at the start of each round. Since each such group played the game only once (no group stayed together for more than one period), we actually have individual observations for 79 groups. We report the results as shown for ease of reference and so that the reader may discern any learning that has occurred since when a group reaches period 3 *all* members of that group have two previous trials.

Appendix B: Experimental instructions – no scrambling case

Experimental instructions

This is an experiment in decision-making. Several research organizations have provided funds for this research. Read the instructions carefully. If you follow them and make good decisions, you may earn a considerable amount of money. This money will be paid to you in cash at the end of the experiment.

Organization

You have been organized into groups of five persons. Each group will consist of the same five persons for the duration of the sessions. The specific identities of the other persons in your group will not be revealed to you. You may not communicate with anyone else in the room during the session. Failure to observe this instruction will result in the termination of the experiment and the forfeiture of all monies earned.

The whole session will last for fifteen periods each of which will be comprised of several stages. At each stage during each period you will be required to make a decision and your total earnings for the session will depend on these decisions.

At the beginning of each period the screen will announce to you the income you will receive in tokens for the period. These tokens will be exchanged for money, at the rate of _____ cents per token, at the end of the

session. Also provided on the screen is the income of each of the other persons in your group. This information may vary from period to period so read it carefully each period.

For each stage you will be asked to post a contribution. If the sum of the contributions from the group meets or exceeds the threshold level reported on the screen the group will receive an additional bundle of tokens to be shared by all the members of the groups regardless of their actual contributions. The actual shares to each person are reported on the screen as part of your information.

Each period will proceed as follows. You will receive a new income in tokens. For Stage 1 you will post a contribution. If the sum of the contributions for the group meets or exceeds the threshold for the stage the additional tokens will be provided. If the threshold is met at Stage 1 you will go on to Stage 2. If the threshold is met at Stage 2 the second bundle of additional tokens will be provided and the group will go on to Stage 3 and so on unil the sum of the contributions from the groups does not meet the threshold. At this point the period ends and a new period will begin.

At each stage you will be informed of your remaining income at this stage. This is calculated by subtracting your net contributions to successful provision of the additional bundles from your initial income. Your share of the bundles of additional tokens is not provided to you until the end of the period. Thus, your share of the additional tokens cannot be used to contribute to the provision of additional bundles.

Contributions in excess of the threshold are kept by the persons running the experiment. For the stage at which the threshold is not attained your contributions are returned. Thus your income for the period is computed as follows: your initial income *plus* your accumulated shares of the additional tokens for those stages in which they are supplied *minus* the sum of your actual contributions in those periods for which the threshold is met or exceeded.

There are some simple rules regarding the contributions you may post. You may enter any contribution from zero up to the level of your income for the PERIOD minus your PREVIOUS CONTRIBUTIONS for the period. Contributions in excess of your current net income will not be accepted. Enter your contribution at the terminal in numbers. You may contribute part tokens, e.g. 0.5 tokens or 4.3 tokens. You will have two minutes to decide on your contribution and to enter it.

Once the contributions have been entered, the computer will compute the totals for each group. If the sum of the contributions meets or exceeds the threshold level for that stage the bundle of additional tokens will be provided to the group and your share will be paid to your account. You will be informed by the computer of the TOTAL contribution of your group but not the contributions of the individual members. You will be informed of the

remaining balance of each of the members of your group. If the threshold for the current stage is met or exceeded the computer will automatically proceed to the next stage.

We have provided a sample screen and session which will be presented when you have all finished reading these instructions. Any questions will be answered once the sample session has been presented.

The screen

The screen is comprised of 4 parts and is divided into boxes. The large box on the left is the Information Screen and it shows the period at the top. Next you see the income for yourself and for the other members of your group. You are informed of the group's THRESHOLD CONTRIBUTION (here 50 tokens) and your own payoff if the threshold is met or exceeded for each of the potential stages for this period. At the bottom you see the total payoff for the group for each stage.

At the top of the right side of the screen you see the box marked 'Contribution'. This informs you of the period and the stage, your current balance for the period and asks you to enter your contribution.

The middle box on the right side reports the RESULTS. Once all of the members of your group have posted their contributions the computer sums these contributions and will tell you the total and whether the additional bundle of tokens (the 'good') is provided at this stage. This box also informs you of your returns for the period.

The final box marked 'Message' is reserved for telling you when you should push the RETURN key to move along in the session.

The session

A session might proceed as follows. For Period 1, Stage 1 say you post 11 tokens as your contribution. The total for your group is 54 which exceeds the threshold so you receive the message in the RESULTS box that the additional tokens are provided. The group proceeds to Stage 2. Your current income is 19 tokens (your original income of 30 tokens minus the 11 you posted at Stage 1). You post a contribution of 8 tokens. This time the total for the group is 38 which is less than the threshold. This time the RESULTS box gives you the message that the additional tokens are not provided and your contribution is returned. One additional bundle of tokens has been provided in this period.

For Period 1 your total income is 30 plus 20 minus $11 = 39$ tokens. And this is added to your account to be paid at the end of the session.

Now you would proceed to the next period.

SAMPLE SCREEN

____INFORMATION SCREEN____
Period no. 1 ID no. 2
(5 Persons per group,
4 Stages per period)

The INCOMES for this period:

yours 30 tokens
others 30 30 30 30 tokens

THRESHOLD CONTRIBUTION of your group for each stage is 50 tokens

If this threshold is met or exceeded, the group will receive the following additional tokens

	Stage 1	Stage 2	Stage 3	Stage 4
Your Share	20	18	15	10
others'	20	18	15	10
	20	18	15	10
	20	18	15	10
	20	18	15	10
Group Total	100	90	75	50

____CONTRIBUTION____
Period no. 1, Stage no. 1
Your Balance = 30.0
Enter your Contribution
→
_____

_____RESULTS_____
The group contributed a total of _____ tokens

The good _____ provided at this stage

Your returns for the period are _____ tokens
(To be distributed at the end of the period)

_____MESSAGE_____

References

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

Bagnoli, M. and B. Lipman, 1989, Provision of public goods: Fully implementing the core through private provision, Review of Economic Studies, Oct., 583–602.

Bagnoli, M. and M. McKee, 1991, Voluntary contribution games: Efficient private provision of public goods, Economic Inquiry 29, no. 2, 351–366.

Camerer, C.A. and K. Weigelt, 1988, Experimental tests of a sequential equilibrium reputation model, Econometrica 56, 1–36.

Clarke, E.H., 1971, Multi-part pricing of public goods, Public Choice 11, 17–33.

Cooper, R.W., D.V. DeJong, R. Forsythe and T.W. Ross, 1990, Selection criteria in coordination games: Some experimental results, American Economic Review 80, no. 1, 218–233.

Dawes, R., J.M. Orbell, R.T. Simmons and A.J.C. van de Kragt, 1986, Organizing groups for collective action, American Political Science Review 80, no. 5, 1171–1185.

Dawes, R. and R. Thaler, 1988, Anomalies: Cooperation, Journal of Economic Perspectives 2, no. 3, 187–197.

Groves, T., 1973, Incentives in teams, Econometrica 41, 617–631.

Groves, T. and J. Ledyard, 1977, Optimal allocation of public goods: A solution to the 'free rider' problem, Econometrica 45, 783–809.

Isaac, R.M., K.F. McCue and C.R. Plott, 1985, Public goods provision is an experimental environment, Journal of Public Economics 26, 51–74.

Isaac, R.M. and J.M. Walker, 1988, Group-size effects in public goods provision: The voluntary contributions mechanism, Quarterly Journal of Economics 103, 179–201.
Palfrey, T. and H. Rosenthal, 1984, Participation and the provision of discrete public goods: A strategic analysis, Journal of Public Economics 24, no. 2, 171–193.
Roth, A., 1987, Introduction and overview, in: A. Roth, ed., Laboratory experimentation in economics: Six points of view (Cambridge University Press, Cambridge, UK).
Samuelson, P.A., 1954, Pure theory of public expenditures, Review of Economics and Statistics 36, 387–389.
Smith, V.L., 1980, Experiments with a decentralized mechanism for public good decisions, American Economic Review 70, 584–599.
Tideman, N. and G. Tullock, 1977, A new and superior process for making social choices, Journal of Political Economy 84, 1145–1160.
van de Kragt, A.J.C., J.M. Orbell and R.M. Dawes, 1983, The minimal contributing set as a solution to public goods problems, American Political Science Review 77, 112–122.
van Huyck, J.B., R.C. Battalio and R.O. Beil, 1990, Tacit coordination games, strategic uncertainty, and coordination failure, American Economic Review 80, 234–248.