Public Goods and the Prisoners Dilemma: Experimental Evidence

Liam Delaney – Junior Sophister

Detective Curran: You like playing Games, don't you?
Catherine Tramell: I have a degree in psychology, it goes with the turf…Games are fun. *(Basic Instinct, 1992)*

Like Ms. Tramell, Liam Delaney is interested in games. In this essay he discusses game theory, experimental economics and what an understanding of psychology can bring to this area. In particular the "free rider" problem is interpreted in terms of established theory. He then examines how his own experimental findings support these views.

In this essay I examine the topic of public goods and demonstrate how the use of experimentation can shed light on many crucial issues, which have, in the past, been the domain of introspection, philosophy and economic theory.

The free rider problem results from the fact that public goods are both non-exclusive and non-rivalrous in consumption. Thus, once provided, there are no dividing lines about how the benefits from consumption should be allocated. This leads, in a competitive market, to under-provision, as the rational agent will under-contribute to the costs of provision and *free ride* off the contribution of others, thus enjoying the benefits of consumption but not accruing the costs. If strict rationality assumptions are imposed and traditional assumptions about agents' utility functions are valid, then pure public goods will not, except by nature, be provided at all in a market system.

The notion of the Prisoner's Dilemma has been employed as a useful analogy to explain the free rider problem. First introduced into the literature by Melvin Dresher and Merrill Flood in 1950, the layout of the Prisoners Dilemma is as follows: Two suspects are charged with a crime. They are placed in separate interview rooms for questioning. Communication between the two is not possible. Each suspect has two choices, either confess or claim innocence. If both confess then both go to prison. If both claim innocence, then both get a lesser jail sentence. However, if one confesses and the other claims innocence then the confessor is rewarded with an even lesser sentence while the unrepentant is given a greater sentence than he would have got had both of them confessed. We are left with a situation where, although the best result for the two comes from them both claiming innocence, the nature of the dilemma makes it optimal for each individual to confess.

Johansson[1] gives an introduction as to how this can be applied to explain market failures. The example he gives takes two individuals that can take up production of a pure public good, which yields to each a benefit of six pounds but costs a total of eight pounds to produce. If the two individuals co-operate, they both earn two pounds, which is the Pareto-efficient outcome. However, if we look at *Table 1*, keeping the following points in mind

1. Communication does not occur

2. It is not possible to make binding contracts

Table 1:

	Net benefit	
	Player 1	Player 2
Both co-operate	2	2
Player 1 but not Player 2 co-operates	-2	6
Player 2 but not Player 1 co-operates	6	-2
Non-co-operation	0	0

We see that the dominant strategy for each individual is to not co-operate. Non co-operation is a dominant strategy in that no matter what the other individual does, non co-operation will yield a higher payoff than co-operation. Thus, if each player plays rationally, non co-operation results in gain for neither party.

However useful an analogy the Prisoner's Dilemma may be, we are still a long way from understanding the processes which lead to non co-operation in such situations. Indeed we have yet to establish that non co-operation actually results in practice. The importance of understanding such phenomena cannot be understated. Ledyard gives a particularly eloquent account of the importance of such research:

> *"Some of the most fundamental questions about the organisation of society center around issues raised by the presence of public goods. Can markets provide optimal allocations of public goods such as air pollution or health? How well do current political institutions perform in the production and funding of public goods such as space exploration and national defence? How far can volunteerism take us in*

[1] Johansson (1991)

> *attempts to solve world environmental problems? If existing institutions, thrown up in the natural evolutionary processes of history, do not produce desirable results, can we discover other organisational arrangements that would better serve the interests of society? At an even more fundamental level public goods raise questions about the very nature of humans. Are we co-operative or selfish?"* [2]

While traditionally such topics would have been subjected to a philosophical or introspective analysis, from the 1950s onwards economists began to make use of experimental methods in order to glean some insight into questions posed by the existence of public goods. Experimental research differs from the traditional non-experimental work in economics in that the scientist has direct influence over the independent variables. This allows for a greater degree of control and increases the confidence with which one can make causal, rather than co-relational inferences. Economics has not generally been viewed as a science which leant itself to experimental methods. As Samuelson and Nordhaus put it:

> *"One possible way of figuring out economic laws is by controlled experiments. Economists [unfortunately] cannot perform the controlled experiments of chemists and biologists because they cannot easily control other factors. Like astronomers or meteorologists, they generally must be content largely to observe."* [3]

This view has certainly not gone unchallenged and as the pages of modern day journal articles will testify, economics is becoming an increasingly experimental science. The experimental research in to public goods, in particular, now spans volumes. Rather than trying to summarise it all here, the remainder of the essay will be devoted to some excellent examples of how experimental economics continues to shed light on the questions outlined above.

We begin by considering the layout of a typical public goods experiment. The basic game runs as follows: A group of students are brought in to a room and given an amount of money to invest. They are told that they may keep the money and not invest it or that they may place a proportion of the money to the group fund. Whatever is in the group fund at the end, the subjects are told, will be doubled up and divided evenly among the group. Thus each participant benefits equally from

[2] Ledyard (1995)
[3] Samuelson & Nordhaus (1985)

the group contribution, regardless of his/her individual contribution. Furthermore contributions are made simultaneously and the entire process is anonymous and without overt inter-personal communication. At the end of the game, subjects are paid in cash the amount they have earned.

We can see that this situation is very similar to the Prisoner's Dilemma and, like the Prisoner's Dilemma, the dominant strategy for each individual is to contribute nothing while the Pareto-efficient outcome is for each group member to contribute everything. Thus, in this example, the outcome predicted by individual rationality will lead to the group earning half what it could have earned had co-operation occurred. In practice, however this result seldom occurs.
Ledyard gives the following review of what actually occurs in practice in such studies, a review, which is consistent with other reviews of the field.[4]

1. In one-shot trials and in the initial stages of finitely repeated trials, subjects generally provide contributions halfway between the Pareto-efficient and the free riding level.

2. Contribution declines with repetition, and

3. Face to face communication improves the rate of contribution.

Thus people are neither completely self-serving nor completely co-operative but begin to converge on to Nash equilibrium towards the end of finitely repeated trials.
We can now begin to go beyond the economic theory and ask some questions as to what is actually going on in Prisoner's Dilemma situations. James Andreoni, a university of Texas professor, is a name synonymous with this type of research. In one of his early experiments[5], Andreoni tested two competing hypotheses, which had been offered to explain the decay effect in public goods experiments. The first is known as the *learning hypothesis*. This maintains that, in the initial stages of these experiments, subjects are unclear as to the nature of the payoffs involved. With each repetition subjects become gradually more conscious of the fact that it is to their advantage to be non-co-operative and, in this way, behaviour gradually converges on to rational Nash equilibrium strategies.

[4] See, for example, Kim & Walker (1984), pp. 113-49.
[5] Andreoni (1988)

An equally plausible hypothesis was offered by Nash himself and is known as the *strategies hypothesis*. For Nash, *finitely repeated games* are misleadingly named and are, in actuality, just one big game rather than a finite number of separate and distinct games. Thus the subjects are affected not only by information about what the other subjects will do in current rounds. They are also influenced by what they believe will be the effects of their own actions on the behaviour of the other subjects in future rounds. The basic claim of the strategies hypothesis is that subjects contribute relatively generously in earlier rounds in order to promote such behaviour among the other subjects. Once the strategic motivation for doing so is gone i.e. towards the end where it is no longer possible to influence future behaviour, the rate of contribution decreases concomitantly.

To test this, Andreoni first devised a standard prisoner's dilemma game, which took the following format. Subjects were placed in groups of five and given fifty tokens to invest. The choice was to invest tokens *privately* or *publicly*. Each token invested privately yielded a return of 5% to the individual. Each token invested publicly yielded a return of 2.5% to each member of the group. Thus with complete co-operation the total group yield will amount to 12.5%. However with completely selfish behaviour the total group yield will only amount to 5%. Using undergraduates as his subject pool, Andreoni replicated the basic result that contributions are initially halfway between free-riding and Pareto-efficient levels and decline with repetition.

However, the main purpose of the study was to test the strategies and learning hypotheses. To test the strategies hypothesis he removed the postulated strategic motivation for contribution. He did this by randomly assigning subjects to two groups *Partners* and *Strangers*. Subjects assigned to the *Partners* condition were informed that they would be in the same group for each repetition of the trial. Subjects in the *Strangers* condition were told that they would be reassigned to different groups with each repetition of the trial and that the likelihood that any two given subjects would be in the same group in different rounds was extremely small. Thus, with the strategic motivation for contribution removed, one would expect initial rates of contribution to be significantly lower in the *Strangers* condition than in the *Partners* condition.

To test the learning hypothesis Andreoni unexpectedly announced a restart at the end of the tenth (supposedly the last) trial. Subjects were told that the experiment would be repeated under the exact same conditions. Now, if the learning hypothesis is correct then the restart should not significantly affect the rate of contribution in

either the *Partners* or the *Strangers* condition. Thus contribution rates in Round 1 of the restarted trial should not be significantly different from those in Round 10 of the original game. The restarted game, instead of running for the full ten rounds, was then unexpectedly stopped after three rounds.

Andreoni's results beg many interesting questions as to what is actually going on in these experiments. The *strategies* hypothesis was found to be unable to account for the decay effect. There was a significant difference in inter-round contribution rates between the *Partners* and the *Strangers*, however, in the opposite direction to what the strategies hypothesis would predict. *Strangers*, who had no strategic motivation to do so, actually contributed more in every round of the trial? Likewise, the "learning" hypothesis could not account for the data. The restart did significantly increase the rate of contribution in both the *Partners* and the *Strangers* condition. Had learning been responsible for the decay observed in the first ten rounds, then there is no reason to suppose that subjects would not continue to act selfishly in what was, in effect, the 11th round.

One explanation for the results may be that Andreoni focused on the wrong forms of strategy and learning. Perhaps, subjects learn the single-shot equilibrium, hence the low contribution rate on round 10 but do not understand either the backwards induction necessary to understand the full game equilibrium or the nature of the strategies involved. This would explain why contribution rates jump back up in Round 11. Selten and Stocker[6] tested this by running a series of super-games using the same subjects. They found the same restart phenomena even after 25 runs of the super-game listed above. As Andreoni points out, once free-rider effects are given close experimental study, the results defy the neat categories of economic theory. A large body of experimental research exists which attempts to explain the free rider phenomenon in terms of non-standard behaviour. Some of this research is beginning to yield both theoretical and practical significance.

Experimental work on the provision of public goods has yielded valuable information about the conditions under which voluntary co-operative behaviour will occur. Andreoni[7] showed that in standard dictator games i.e. where one person has a choice to maximise his/her benefit at the expense of others, there were significant interaction effects between the degree of selfishness and the gender of the dictator. Thus while there was no linear relation, it was found that males were more likely to

[6] Selten & Stoeker (1986)
[7] Andreoni (1999)

be either completely selfish or completely selfless whereas females were generally found to be egalitarian, distributing the effects of the wealth evenly. The significance of such results for, as Andreoni puts it '*those who broker in altruism*' i.e. charity organisations, is clear. Practical understanding of co-operative and charitable behaviour can greatly facilitate an increase in its rate of its occurrence.

It is through this type of channel that experimental economics is developing in to a flourishing sub-discipline and incorporating many methods and facets of human behaviour that have long been neglected by traditional economic theory. In this regard economics has become closely inter-related with academic sub-disciplines within psychology such as Social Psychology[8] and Learning Theory. Again the increasing volumes of work testify to a union between the two disciplines. As this essay has tried to show this union is essential if we are to understand vital issues such as free rider affects which, although they are made more understandable by economic theory, cannot be explained by resorting to standard rational economic assumptions.

As a brief endnote, my own attempt to shed some light on public goods phenomena looked at individual differences in certain personality traits. My aim was to reduce some of the noise in these simulations by trying to establish, a priori, who would be likely to show a high rate of contribution and who would be likely to show a low rate of contribution? The questionnaire measured the personality traits of *altruism* and *concern for money*, the questions being taken from a popular test used for recruitment and selection purposes by many firms.[9] As it turned out these personality traits did not predict mean contribution rates. The usual decline effect was evident but mean contribution rates were higher in eight of the ten rounds than the subject pool in the *Partners* condition of Andreoni's experiment, this probably being due to the greater degree of anonymity in the Andreoni study.

[8] See, for example, Akerlof & Dickens (1982), pp. 307-19 for an excellent example of how "*psychological variables*" can be rigorously incorporated into inter-temporal models whilst still guaranteeing the existence of the resulting equilibrium.
[9] Hunter & Roberts (1989)

Bibliography

Akerlof, G. & W. T. Dickens (1982) "The Economic Consequences of Cognitive Dissonance" in American Economic Review, 3, June.

Andreoni, J. (1988) "Why Free Ride? Strategies and Learning in Public Goods Experiments" in *Journal of Public Economics*, 37, pp. 291-304.

Andreoni, J. (1999) "Which is the Fair Sex?" "Gender Differences in Altruism". Available from James Andreoni official website.

Johansson, P.O. (1991) *An Introduction to Modern Welfare Economics*. Cambridge University Press: London.

Kim, O. & M. Walker (1984) "The Free Rider Problem: Experimental Evidence" in *Public Choice*, 43.

Ledyard, John (1995) "Public Goods: A Survey of Experimental Research" in Kagel, J. and A.E. Roth (eds.) *Handbook of Experimental Economics.* Princeton University Press: New York.

Samuelson & Nordhaus (1985) in Friedman, D. & S. Sunder (eds.) (1994) *Experimental Methods: A Primer for Economists.* Cambridge University Press: London.

Selten, R & R. Stoeker (1986) " End Behaviour in Sequences of Finite Prisoner's Dilemma Supergames: A Learning Theory Approach" in *Journal of Economic Behaviour and Organization.*