Voluntary Contributions to Public Goods: Generosity or Attempted Cooperation

Public goods are goods which are *nonexcludable* (no one can exclude another's consumption) and *nonrival* (one person's use does not diminish another individual's consumption). Theory suggests that because agents do not consider the positive benefits realized by others when they make their decisions, they will provide less than the optimal amount of the public good. According to theorists, agents will "free-ride" on the generosity of others without returning the favour.

In public goods experiments subjects quite regularly provide a level of public good that exceeds the theoretically derived self-interested level. This result has been repeated many times under a variety of experiment treatments. This has lead to the hypothesis that people might gain some utility from the act of contributing to the group good; that is, they may obtain utility from being generous.

This paper presents the results of a series of laboratory sessions which distinguish between the effects of generous behaviour and attempts at cooperative behaviour of a group of subjects participating in a voluntary contributions public goods setting. If generosity has value, then generous subjects should be willing to incur a cost to be generous. Cooperators, on the other hand, do not give for the benefit of others but instead understand the benefits of mutual cooperation as being in their own self-interest. Results suggest that subject behaviour may better be characterized as attempted cooperation rather than generosity.

Robert Moir
May 1995 Version

DRAFT -- Please do not quote without author's permission.

Dept. of Economics, McMaster University, Hamilton, ONT. Canada,
Phone: (905) 525-9140 x23757
Fax: (905) 521-8232
Internet: moir@mcmail.cis.mcmaster.ca

Voluntary Contributions to Public Goods: Generosity or Attempted Cooperation

Rob Moir[1]

Not he who has much is rich, but he who gives much.
- Erich Fromm.

Altruism: *The principle or practise of unselfish concern for the welfare of others.*
- Collins English Dictionary, 1989.

I. Introduction

Economists are presented with a Herculean task when asked to study altruism, as altruism does not fit well with our concepts of simple utility maximization in which it is assumed that individuals are self-interested. Altruism, modelled with interdependent utilities -- B's welfare is an argument in A's utility function -- has been used to explain positive contributions towards public goods (Barro, 1974; Becker, 1974; cited in Andreoni, 1989). However, if the utility a person receives from a public good is *not* additively separable from their private good utility, then the Nash behavioural assumption predicts positive voluntary contribution to the public good (Warr, 1983; Bergstrom, Blume and Varian, 1986) which is entirely motivated by self-interest. Thus, under plausible circumstances, the observation of voluntary contributions does not necessarily imply altruism.

Experimental evidence suggests that the Nash equilibrium prediction is not strictly adhered to; it is often the case that individuals and groups over-contribute relative to the standard Nash zero conjectural variation (ZCV) equilibrium both in interior solution (Chan, Mestelman, Moir and Muller, 1993; Li, 1991) and dominant free-riding (zero contribution) environments (Isaac, Walker and Thomas, 1984). Andreoni (1989, 1990) suggests that some agents receive a **warm glow** from being generous, and hence they over-contribute relative to the standard Nash equilibrium prediction.

The theory of warm glow has two quite important implications. First, it can radically alter a government's taxation policy, as shown by Andreoni (1989). Second, this theory also raises many questions for experimental economists. If warm glow exists, how can it be accounted for in our analysis? In most contexts, warm glow is a private valuation which cannot separated from learning, mistakes or strategy in experimental data. If we assume or conclude that it exists, then we can explain almost any positive deviation from the Nash equilibrium prediction by appealing to warm glow and negative deviations might then be described as **cold prickle** (Andreoni, 1995).

---

[1] I am grateful to David Feeny, Stuart Mestelman, Mark Isaac and Samuel Dinkin for their editorial and design comments. Andrew Muller developed the initial software used to conduct this experiment. Funding was provided by Vernon Smith at the Economic Science Laboratory at the University of Arizona and by Environment Canada's Green Plan through a grant from the Tri-Council Eco-Research Secretariat. All remaining errors are my own.

In this paper, the warm glow theory is tested using experimental economic methods. In section II, the warm glow theory and theories with similar predictions are discussed. The difficulties of analyzing warm glow and an experimental design which avoids these difficulties are presented in sections III and IV respectively. The predictions and results of this experiment are presented in sections V and VI respectively. Conclusions follow in section VII.

II. Voluntary Contribution to Public Goods: Theory

In two theory papers, Andreoni (1989, 1990) proposes a warm glow addition to agents' utility functions. Essentially, he suggests that agents may not derive benefit solely from the levels (in any combination) of the private or public good, but may also gain some utility by ***generously*** giving to the public good. Individuals who receive no warm glow should exhibit complete crowding-out (*neutrality* according to Warr, 1983; Bergstrom et al., 1986) as their own contribution and government contribution are viewed as perfect substitutes. That is, if the government levies lump-sum taxes and uses this revenue to provide the public good, individuals will decrease their voluntary contribution by the same amount as they were taxed. However, for agents who obtain a warm glow from generosity, this perfect substitution result does not follow and contributions are no longer completely crowded-out. Thus, because "people derive some utility from the act of giving" itself, "warm glow makes private gifts imperfect substitutes for gifts from other sources" (Andreoni, 1989: 1457) and governments can then tax individuals with warm glow to finance the provision of public goods, and increase the total (voluntary plus taxed) contribution to the public good.

Also important to warm glow theory is that the standard Nash equilibrium prediction obtained without accounting for generosity in the utility function is less than the modified Nash equilibrium prediction made when an individual's utility from generosity is taken into account. Andreoni (1989) proposes that experimental evidence of agents contributing more than the Nash equilibrium prediction is because subjects do not maximize their payoff but rather a utility function which includes a warm glow from generous contribution to the public good. Unless the theory of warm glow giving is somehow upper-bounded by taxed contribution levels, generosity should be evident at all levels of contributions because of the imperfect substitution property.

A number of models other than warm glow suggest that contributions to the public good will be greater than the simple Nash equilibrium prediction. A model of interdependent utility which captures altruism, suggests that individuals will over-contribute relative to the simple Nash equilibrium prediction based on own returns. Furthermore, interdependent utility also predicts imperfect crowding-out[2], because, like the generosity model, own contribution and government contributions are not viewed as perfect substitutes. If subjects feel any spite or guilt in their contribution to the public good, then the standard Nash equilibrium prediction may not be

[2] Imperfect in the sense that there may be incomplete crowding-out if others' utilities are negatively related to yours (which happens to be consistent with Andreoni's (1993) data) or excessive crowding-out if other's utilities are positively related to yours.

attained, and crowding out will be imperfect (Chan, Godby, Mestelman and Muller, 1994).

Alternatively, if we do *not* assume that individuals make zero conjectural variations (ZCV) then the standard ZCV Nash equilibrium is not correct. Bergstrom et al. (1986) have shown that such a model retains the neutrality property. However, if a number of agents have positive conjectural variations, then the self-interested equilibrium prediction (accounting for the positive conjectural variations) would exceed the ZCV Nash equilibrium prediction and lead to a similar conclusion as Andreoni's warm glow theory.

III. Standard Models and Tests

It is impossible to distinguish between the effects of generosity, the effects of strategic action (signalling, punishment/reward strategies, non-ZCV, etc.), or the effect of interdependent utility when voluntary contributions are primarily in the region between the simple Nash equilibrium and the equal-contribution Pareto optimal solution. Thus, one must test for generosity (and its opposite, ***vindictiveness***) at these particular boundary points, where there is no reason to expect additional (lower) contributions if non-ZCV, signalling, or interdependent utility models are the true behavioural models, yet warm glow (cold prickle) would predict additional (lower) contributions.

To date, public good experiments have modelled the payoff from a public good using three different methods. First were the linear public goods in which it is a dominant strategy to contribute nothing to the public good, and socially optimal to contribute all endowments to the public good (Isaac, Walker and Thomas, 1984; Isaac and Walker, 1988). Second, there are experiments using payoff structures with interior (non-dominant) Nash equilibrium and socially optimal contribution levels (Li, 1991; Isaac and Walker, 1992; Andreoni, 1993; Chan et al., 1993). In these cases, agents at either contribution level consume a mix of both the private and public good. Finally, there exist experiments which look at provision points (Bagnoli and McKee, 1991). I will not consider the third type of experiment.

In the first two experiment types, there is room for the experimenter to effectively provide the public good through taxation by requiring individuals to contribute some level towards the public good (make the voluntary contribution mechanism less than voluntary). One could speculate that once individuals are informed that they are being so taxed, then if they are taxed to the Nash equilibrium contribution level, any remaining contribution might be considered warm glow. This myopic view is not necessarily the case though.

First, agents are known to learn over the course of an experiment. During this learning process, they can either make mistakes, or make educated guesses which provide them with a greater understanding of the laboratory environment. It will always be difficult to remove this effect from any experiment.

Second, by taxing individuals up to the Nash equilibrium, you reduce the cognitive complexity of the experiment and further make cooperative signalling a more effective tool by reducing the ability of others to "free-ride" upon a cooperators good nature[3]. Boundary effects may become important (Chan et al., 1994). This may explain the results of Andreoni's (1993) test of the crowding-out hypothesis which found incomplete crowding-out. He concludes that,

> While this experiment does not provide direct evidence for the motives that people may have for contributing more in the presence of taxation, the behaviour in the experiment is broadly consistent with the hypothesis that people get pleasure from the act of contributing to the public good.
>
> (p.1325)

The subjects in these experiments however, showed little evidence of Nash over-contribution in the no-tax experiments. This is incompatible with the predictions of the warm glow model.

Third, there is both experimental evidence and theoretical suggestion (see Radner, 1980; Friedman, 1971) that even in finite games, agents may show some sort of strategy. This strategy may involve signalling and punishment in order to lead a group to the social optimal level of contribution. Indeed, at any group contribution level exceeding the Nash but less than the socially optimal, group returns are greater than at the Nash. If all agents contribute at a level slightly greater than their Nash equilibrium predicted levels, then each individual has a higher payoff than what would have been received had all played strictly Nash. Theorists argue that rational backwards induction should prevent signalling in any finite game, thus preventing collusion from taking place. However, if the group is rational enough to use backwards induction, then why are they not rational enough to see through the fallacy, ignore the formal rationality, and gain the highest returns possible through cooperation? This question leads to an infinite regress problem which is left to the theorists to discuss.

The effects of the first problem, learning, on the interpretation of experimental results are often insurmountable. They are usually accounted for by deleting the first few observations. However, it is possible to remove elements of the second problem and nullify the third problem altogether by taxing at the socially optimal level and observing agents' contributions to the public good. At this point, all strategic possibilities are removed. In the first type of free-riding experiments, this is impossible, as the social optimum involves all subjects contributing all of their endowment to the public good. However, in a non-linear public good (i.e. a case where the public good and the private good have interactive effects in people's utility), neither the socially optimal nor the Nash equilibrium contributions are necessarily bounded by an individual's endowment. Any contribution to the public good, above the taxed amount, reduces individual returns by more than it improves the social gain[4]. An agent's consistent positive

---

[3] When the decision space is reduced by such a tax, the *strategy space* has been reduced. Subjects cannot free-ride upon the cooperative nature of others which implies that cooperation has become easier. This might explain some of the non-neutrality that Andreoni (1993) finds.

[4] In this case, the contribution of no more than the mandatory number of tokens is a *super dominant* strategy. Unlike a prisoner's dilemma in which the payoff for both cooperating exceeds the payoff from both cheating, here payoff is at its highest. Any additional contribution reduces own payoff, and cooperation among individuals to increase group contribution reduces individual payoff. We see that when individuals are required to make a contribution no larger than

contribution to the public good then can be viewed as a measure of their generosity.

It is not the purpose of this experiment to discuss the relevancy of the interdependency model. Instead, this model is presented as an alternative explanation for the commonly observed over-contribution results experimentalists have observed in public good environments. The theoretical predictions of both models seem to agree with previous experimental results. However, at the social optimum, the interdependency model reasonably calls for no additional contribution[5].

It is only fair to test for generous behaviour under conditions which give the theory a fairly high chance of success. As such, agents must know, or at least be reasonably sure of, the effects of their contributions upon others[6]. If an environment existed where one's additional contribution inflicted damage upon others (e.g. a negative externality such as a public bad or in a common-pool resource environment), then a reasonable interpretation of generosity would be ɔwer contribution levels. Furthermore, Hume's theorem as discussed by Roumasset (1991) suggests that agents might find it easier to cooperate in an environment in which the number of agents is relatively small. Both of these features are captured in the design below.

V. Experiment Design

As suggested in the previous section, this experiment involves a tax, earmarked for the public good, which is set at the socially optimal level of contribution. In addition, it is of interest to discover the relationship between generosity and vindictiveness for individuals. In order to test this, a ABC and BAC design is suggested. Subjects either play the generosity (A) phase or the vindictiveness (B) phase first, which is then followed by a series of periods (C) with no restrictions placed on their contribution. The vindictive environment is created by placing a ceiling upon an individual's per period contribution at the Nash equilibrium. They may contribute any amount up to and including this amount. Any contribution less than the ceiling on a consistent basis, costs an individual more than it hurts the group, and thus we conclude that individual behaves in a vindictive manner. This design is referred to as the *partners* design.

In an alternate *strangers* design, subjects were randomly rematched every period to form ew groups. These groups were randomly assigned to treatment A, B or C each period. All ther design features were constant across experiments. The warm glow theory would predict

---

ʰe Nash equilibrium contribution, meeting the mandatory requirement is also a *super dominant* strategy.

[5] Essentially this requires that no individual cares more for others than she does for herself.

[6] That is, subjects must be sure that this is indeed a public good for *all* people, and not a public bad for some. This can best be operationalized by showing subjects that their payoffs are exactly the same. Palfrey and Prisbrey (1993) perform a linear public goods experiment in which the returns to private and public good contributions are random each period. They observe low levels of 'splitting' -- investment in both goods -- in this experiment and conclude that warm glow and altruism do not exist. However, even a small amount of splitting may provide evidence for either warm glow or altruism in a very harsh, low-information environment.

no difference between these two designs in any of the phases. However, alternate theories about ublic good contribution predict no additional contribution in the A and B phases and a reduction of cooperation in the unrestricted phase in the strangers design.

The payoff function, described in detail below, is common to all subjects, and all subjects are informed that this is the case. Agents are provided with a payoff table on which their decision space is clearly marked. They are told that other agents have the exact same payoff table. Furthermore, all agents are made aware that the minimum (maximum) contribution restriction applies to them and to all members of their group. The conversion ratio for all individuals is one lab dollar is equivalent to $0.005 US/CDN, and this is common knowledge.

The particular form of the payoff function used for this experiment is presented below. There are n=3 agents each endowed with w_i=24 tokens. Let g_i be agent i's contribution to the public good, and x_i be his/her contribution to the private good. The payoff for subject i in lab dollars is then;

$$\pi_i = x_i + G + (\tfrac{1}{n}) x_i G \tag{1}$$

where $G = \sum_{k=1}^{n} g_k$. Notice that $\frac{\partial \pi_i}{\partial G_{-i}} = 1 + (\frac{1}{n}) x_i > 0$ so no individual is ever harmed by another person's contribution. Also note that the Nash equilibrium contribution of 6 is less than the focal point $\frac{w_i}{3} = 8$.

This continuous payoff function has a ZCV best response function,

$$g_i^* = \frac{w_i - G_{-i}}{2} \tag{2}$$

which solves to $g_i^* = 6$[7] and $G^* = 18$ if all like agents behave alike. The resultant return per agent is then $\pi_i^* = 144$ per period in lab dollars.

The social optimum contribution level is of the form

$$G^s = \frac{n(n-1)}{2} + \frac{W}{2} = 3 + \frac{W}{2} \tag{3}$$

for n=3 and $W = \sum_{k=1}^{n} w_k$. With W=72, then any combination of g_i^s's summing to G^s=39 is Pareto optimal. The symmetric Pareto optimum is $g_i^s = \frac{G^s}{3} = 13$. The per period profit of agent i is then $\pi_i^s = 193$ in lab dollars.

The *partner* experiments were run in three parts. In the ABC design, part one uses the

[7] When a discrete allocation mechanism is used (agents must contribute an integer level of tokens), another Nash equilibrium exists at {5,6,7}. However, this does not follow the 'like act alike' rule, and it is expected that fairness heuristics will dominate this equilibrium and lead individuals to the {6,6,6} solution.

taxation procedure described above, and lasts for 8 periods. Part two then consists of 8 periods in which the ceiling on contributions is applied, and part three follows with 8 periods, but the contribution restrictions are removed. In the *strangers* experiments, subjects faced each of the 3 treatments a total of 8 times. These treatments were faced in a random order for a total of 24 periods.

The experiments were conducted using public goods software developed at McMaster University and run on a UNIX network. The first set of experiments were conducted in May 1994 at the Economic Science Laboratory at the University of Arizona. Subjects were recruited from undergraduate economics courses. The October 1994 experiments were conducted at the McMaster Experimental Economics Laboratory. Subjects were recruited from the general student population. Instructions[8] were read aloud to the subjects and a short question and answer session was conducted to ensure they knew how to use the payoff tables. Subjects were informed that their groups would remain constant, and they traded payoff tables to confirm that all payoff tables were alike. In the *partners* experiment, a highlighted version of the payoff table was given to the subjects at the beginning of each sub-section and instructions concerning that portion of the session were read aloud. In the *strangers* treatment, subjects were provided with 3 payoff tables (two of them highlighted) which corresponded to the three different treatments. Subjects knew their allocation of tokens and the combined allocation of tokens to Market 2 made by the other members of their group.

V. Predictions

If generosity characterizes subject behaviour, then it must have some value. Subjects should be willing to pay money to be generous. Only when contributions are restricted to be at or above the Pareto optimal (13) can the effects of generosity be unambiguously separated from strategic behaviour. The first prediction states:

> **(1) Mean individual contributions to the public good, when corrected for upper (24) and lower (13) boundary effects, should exceed 13 in the minimum contribution phase if generosity characterizes subject behaviour.**

This is a direct test of the existence of the warm glow described by Andreoni (1989, 1990). When required contributions are 13, a subject who expects others to contribute 13 each, will earn 92¢ instead of 97¢ if they contribute 14 tokens. This subject incurs a cost of 5¢ to satisfy their warm glow, and their action cannot be interpreted as strategic signalling. To test this prediction, the tobit regression:

$$g_{it}=\beta_0+\beta_1 t+\beta_2 D+\beta_3 S+\epsilon_{it} \tag{4}$$

with the upper and lower censoring bounds appropriately set, is conducted. In this regression t is the normalized period, D is a design dummy (1 for ABC and 0 for BAC), and S is a dummy

[8] Copies of the instructions are available from the author upon request.

for the *strangers* experiment[9]. Assuming that period and design have negligible effects, then prediction (1) implies that we will reject $H_0: \beta_0=13$ in favour of $H_A: \beta_0>13$. Furthermore, central tendency statistics and frequency of plays should also confirm that the majority of subjects contribute more than 13 tokens if the prediction holds true.

The opposite of generosity, vindictiveness, can be identified if subjects are constrained to contribute no more than the Nash equilibrium number of tokens -- 6 each. Significant deviations below six would indicate that subjects are willing to pay to make others worse off.

(2) Mean individual contributions to the public good, when corrected for upper (6) and lower (0) boundary effects, should be below 6 during the maximum contribution phase if vindictiveness characterizes subject behaviour.

If some people exhibit generosity, then there may be others who exhibit vindictiveness. To test this prediction, the tobit regression:

$$g_{it}=\alpha_0+\alpha_1 t+\alpha_2 D+\alpha_3 S+\epsilon_{it} \quad (5)$$

with the upper and lower censoring bounds appropriately set, is conducted. Assuming that period and design have negligible effects, then prediction (2) implies that we will reject $H_0: \alpha_0=6$ in favour of $H_A: \alpha_0<6$. Furthermore, central tendency statistics and frequency of plays should also confirm that the majority of subjects contribute less than 6 tokens if the null hypothesis is true.

(3) Subjects are predisposed to generosity; they are not predisposed to vindictiveness. That is, if subjects are on average generous, they will deviate from 13 when the minimum contribution is 13 *more often* than they will deviate from 6 when the maximum contribution is 6.

The model of generosity in voluntary contributions to public goods explained the over-contribution relative to the ZCV-Nash equilibrium that is evident in so many experiments. If generosity is the correct explanation, then the majority of the population should not show vindictiveness as a trait. In order to evaluate this prediction, the mean deviation from the two boundaries must be calculated. The frequency at each level of token contribution is calculated and multiplied by the absolute deviation from the boundary. The sum of this value is the mean deviation from the boundary. For the minimum required contribution of 13, the calculation would be:

[9] Tobit modelling assumes that the error in the 'latent' linear model is normally distributed. It then corrects for the fact that the data has been censored.

$$\overline{D_{13}}=\sum_{i=13}^{24} f_i \cdot |(i-13)| \qquad (6)$$

where $f_i = \frac{\text{number of plays at } i}{\text{number of observations}}$. A corrected value accounts for the fact that between 13 and 24 the decision space is 12, while between 0 and 6 the decision space is 7. The mean deviation can be expressed as a percentage of the decision space.

> **(4) Grouped data from the unrestricted contributions portion of the experiment will show results similar to that of past public goods experiments. Contributions will initially exceed the Nash equilibrium and then decay towards the Nash equilibrium.**

This prediction tests whether the participants in this experiment combined with the use of these particular parameters and instructions leads to the standard results found in similar experiments. It also tests for the standard over-contribution relative to Nash result consistent with the warm glow theory (among others). In order to capture the dynamic effects of time, the variable t^{-1} is used. A standard linear regression model:

$$G_{jt}=\gamma_0+\gamma_1 t^{-1}+\gamma_2 D+\gamma_3 S+\varepsilon_{jt} \qquad (7)$$

will be used in which ε_{jt} is assumed identically and independently distributed.

> **(5) Individual contributions in the unrestricted phase consistently deviate in a positive direction from the standard Cournot prediction.**

Nash equilibrium is a static rather than intertemporal concept. As such, its rejection as a description of human economic behaviour is almost certain because human behaviour evolves over time. The Cournot model of sequential decision-making provides a dynamic scheme whereby two individuals might reach an equilibrium. This model, however, does not involve complex strategic play (e.g. signalling for cooperation/non-cooperation or learning about the nature of other economic individuals). Instead, it assumes that every individual reacts to the observed actions of others in the preceding period, by maximizing their payoff assuming others will not change their actions. If a model of generosity characterizes subject behaviour, then there should be significant *positive* deviation away from the ZCV Cournot prediction which does not decay over time. Specifically, this result should hold for the strangers experiment when there are no gains to be made from signalling. It may be argued that the minimum required contribution restriction only tests for a strong version of warm glow ($\frac{\partial WG_i}{\partial g_i}>0$ and $\frac{\partial^2 WG_i}{\partial g_i^2}\geq 0$). This analysis tests for a weaker version of warm glow in which warm glow declines as own contributions increase ($\frac{\partial WG_i}{\partial g_i}>0$ and $\frac{\partial^2 WG_i}{\partial g_i^2}\leq 0$).

We can calculate subject i's deviation from the Cournot predicted contribution in period t using the following formula:

$$dev_{it}=g_{it}-(\frac{w_t-G_{-i(t-1)}}{2}). \tag{8}$$

When $G_{-i(t-1)}=0$ and $g_{it}=0$ then dev_{it} is lower bounded at -12. A tobit regression:

$$dev_{it}=\eta_0+\eta_1 t+\eta_2 D+\eta_3 S+\epsilon_{it} \tag{9}$$

is used to test this prediction. Prediction (5) suggests that $H_0: \eta_0=0$ will be rejected in favour of $H_A: \eta_0>0$ <u>and</u> $H_0: \eta_1=0$ will not be rejected in favour of $H_A: \eta_1 \leq 0$.

> **(6) If subjects on average exhibit generosity, the *truncated* frequency distribution of individual contributions from the unrestricted phase, should be shifted to the left of frequency distribution of individual contributions from the minimum contribution phase. Additionally, the *truncated* unrestricted data frequency distribution should lie to the right of the maximum contribution data.**

This prediction suggests that the deviations made at either of the restricted phases cannot be explained in terms of individuals' actions in an unrestricted environment. In the unrestricted phase, individuals may show generous behaviour or cooperative attempts. However, in the maximum contribution phase, cooperative attempts are futile so only generous behaviour remains as an explanation. This implies that the frequency distribution of token contributions in the unrestricted phase should lie to the left of the minimum contribution phase. Similar logic would suggest that if subjects behave generously in the unrestricted contributions phase then the censored unrestricted data should lie to the right of the minimum contribution data.

> **(7) The design difference (ABC versus BAC) shows no effect in terms of individual contributions in either the minimum or maximum contribution phases.**

These predictions can be analyzed directly from the regression results in regressions (4), (5), (7) and (8). Prediction (7) can be tested more directly if we analyze the mean individual contribution, by design in each period, for the two phases. As this data is truncated at both ends and highly skewed towards 13 or 6, we know that the normality assumption is violated, so the t test for difference in means is invalid in finite samples. An exact randomization test for difference in means is thus used (Kennedy, 1995; Moir, 1995).

VI. Results

The results of the experiment are often reported in two ways. Both entire phase (all eight periods) and final period (period 8) data are utilized. In the *partners* experiment, subjects always knew that each phase was 8 periods long, and that there were 3 phases. Furthermore, they were informed that groups would remain constant. This ensures that strategy and generosity are open options to them and also that priors over the (un)cooperative nature of others are being updated and not re-initialized. However, this leads to a problem when one considers the minimum required contribution phase (A) in both the ABC and BAC designs. Subjects know that the experiment will continue for another 16 or 8 periods respectively. Although they do not know what the conditions will be in the next phase, they may wish to signal their cooperativeness by contributing more than 13 tokens in period 8. A similar strategy is available to subjects in period 8 of the maximum contribution phase, but contributions of less than 6 show only non-cooperativeness which has no reasonable strategic value. Thus, for the minimum required contribution phase, period 7 data is used as a proxy for the final period data. This is not a problem in *strangers* experiment as treatment and group composition were randomized each period. Subjects in this experiment were made aware that the experiment would last 24 periods. Table 1 summarizes the data for the entire experiment.

TABLE 1

t	Mean	Std Dev.	Median	Mode	Decision Space	Obs
1	14.344	2.224	13	13	12	90
2	14.378	2.564	13	13	12	90
3	14.011	2.096	13	13	12	90
4	13.889	2.358	13	13	12	90
5	14.000	2.522	13	13	12	90
6	13.567	1.629	13	13	12	90
7	13.622	2.118	13	13	12	90
8	13.756	2.485	13	13	12	90
1	5.107	1.538	6	6	7	75
2	5.560	1.093	6	6	7	75
3	5.787	0.741	6	6	7	75
4	5.667	1.070	6	6	7	75
5	5.693	1.026	6	6	7	75
6	5.613	1.293	6	6	7	75
7	5.600	1.103	6	6	7	75
8	5.520	1.408	6	6	7	75
1	8.467	6.000	10	-	25	75
2	7.893	5.562	8	-	25	75
3	7.147	5.336	7	-	25	75
4	7.120	5.824	6	-	25	75
5	6.533	5.145	6	-	25	75
6	6.227	4.666	6	-	25	75
7	6.787	4.998	6	-	25	75
8	6.133	4.545	6	-	25	75

In Table 1, time periods are normalized at 1 for the start of each phase of the experiment. In the final phase, the results are quite consistent with past experimental evidence ·f both mixed and pure public goods. The decision space is a measure of the number of tokens subjects could actually choose to donate. In the *partners* experiment, 12 groups of 3 participated in the ABC design and 6 groups of 3 participated in the BAC design. The session conducted on May 19, 1994 suffered from a hardware problem, so data was only collected for the A part of the ABC design. In the *strangers* experiment, 12 groups of 3 participated. The data in Table 1 are also summarized in Graph 1.

Result 0

There is no significant experiment effect between the partners and strangers treatment, specifically during the minimum required and maximum restricted contribution phases.

In either experiment, there is no strategic reason to contribute more than 13 (less than 6) tokens in any period when contributions are restricted. Because the data has been censored at a boundary, typical parametric tests cannot be used. An approximate randomization test fails › reject the null that the aggregate final period contribution of the 18 partners groups and 12 trangers groups are equal in either of the restricted phases at the 10% level of significance. Visual inspection of Graph 2 also suggest this conclusion. When necessary, the data analysis below includes a dummy for the strangers experiment, which is included for completeness. Within the analysis of either the minimum required or maximum restricted contribution phase this variable is not significant.

Result 1

On average, subjects do not contribute more than 13 tokens when their contribution must be 13 or greater.

As the data in Table 1 suggest, contributions did not stray far from 13 throughout the minimum r[illegible]ired contribution phase of the experiment. Averages exceed the value 13, but any [illegible] ny period by any subject will necessarily mean that this is the case. The median [illegible] always 13. Of all 720 individual observations in this phase, 527 (72.9%) were 3. In the final period, 79 out of 90 individuals (87.8%) were 13. Furthermore, 28 out of 90 subjects (31.1%) played 13 for the entire 8 periods. One individual contributed more than 13 in every period. The results of the tobit regression of contribution upon time, design and xperiment type are contained in Table 2.

TABLE 2

Tobit Regression of Individual Contribution Upon Time and Design
Upper Bound 24, Lower Bound 13

Variable	Coefficient	t-stat	p-value

t	- 0.61773	- 4.541	0.000
D	2.71077	3.327	0.000
S	- 1.32911	- 1.542	0.124
constant	11.01678	12.322	0.000

It is certainly impossible to reject the null that the constant is equal to 13 in favour of the alternative that the constant is greater than 13. Both time and design have significant effects. As usual, contributions decay over time. The ABC design positively effects contributions. This may be due to learning. If so, we would expect a negative relation between design and contributions in the tobit analysis of the maximum contribution phase.

These results suggest that the generosity model is not on average a strong predictor of economic behaviour. Subjects do not appear to gain utility just out of giving for the sake of giving.■

Result 2

> **On average, subjects do not contribute less than 6 tokens when their contribution is restricted to be 6 or less.**

Similar to Result 1, the data in Table 1 suggest that contributions did not stray far from 6 throughout the B phase of the experiment. Again, the average is less than 6 but the median and mode are both 6. Of all 600 individual observations in this phase, 508 (84.7%) were 6. In the last period, 65 out of 75 individuals (86.7%) were 6. Furthermore, 44 out of 75 subjects (58.7%) contributed 6 tokens in each of the 8 periods. Two individual always contributed less than 6 in every period. The results of the tobit regression of contribution upon time, design and experiment type are contained in Table 3.

TABLE 3

Tobit Regression of Individual Contribution Upon Time and Design
Upper Bound 6, Lower Bound 0

Variable	Coefficient	t-stat	p-value
t	0.36042	2.428	0.015
D	- 1.51960	- 1.694	0.091
S	0.90870	1.049	0.295
constant	10.15435	9.844	0.000

Again, it is impossible to reject the null that the constant is equal to 6 in favour of the alternative that it is less than 6. Design has a moderately significant negative effect upon contributions. As suggested before, this might indicate subject learning about dominant strategies.

Result 2 suggests that vindictiveness does not characterize the behaviour of subjects in his experiment. Subjects do not appear to be willing to forgo income in order to cause others to lose income.■

Result 3

Subjects are not predisposed to generous contributions over vindictive withholding of contributions.

The mean deviations from the boundaries in both the minimum and maximum contribution phases are presented in Table 4. Results for the entire phase and last period are presented in terms of average deviation and percentage of decision space[10].

TABLE 4

	All Periods		Final Periods	
	Mean Deviation	% Decision Space	Mean Deviation	% Decision Space
Minimum Phase (13-24)	0.946	7.9	0.700	5.8
Maximum Phase (0-6)	0.432	6.2	0.480	6.9

After correcting for the size of the decision space, it is not evident that subjects deviate from the boundary any more in the minimum required contribution phase than in the maximum contribution phase. When a tobit regression is run upon a constant, we can calculate a corrected mean contribution under each of the two conditions. These results are presented in Table 5.

TABLE 5

	Mean	Std. Dev.	n
Minimum Phase (13-24)	8.837	12.754	720
Maximum Phase (0-6)	11.730	18.080	600

This supports the earlier findings in Results (1) and (2). In fact, a t-test for difference in means rejects the null that these means are equal at the 1% level of significance. It suggests that when subjects are restricted to contribute at least the symmetric Pareto optimal token contribution, they would, left to their own decisions, prefer to cooperate less than when they are restricted to

[10] See equation (6).

contribute at most the Nash equilibrium.■

Result 4

> **The group total contributions in the unrestricted phase are similar to the results found in other experiments. Specifically, group total contributions initially exceed the Nash equilibrium prediction and decay towards the Nash equilibrium over time.**

This result can be seen by visual inspection of Table 1 and Graph 1 at the individual level. Individual contributions on average exceed the standard ZCV Nash equilibrium but decay towards it over time. In Graph 3, a histogram of group contributions for the entire experiment is presented. The mean contribution of 21.1 is significantly greater than 18 (p-value=0.000).

Table 6 contains the results of a linear regression of group total contribution to the public good upon the reciprocal of time, design and experiment type.

TABLE 6

Unrestricted Linear Regression of Group Contribution on Time^{-1}, Design and Experiment

Variable	Coefficient	t-stat	p-value
t^{-1}	7.58348	3.348	0.000
D	2.37202	1.384	0.168
S	1.93750	1.258	0.210
constant	16.94448	11.529	0.000

$\overline{R}^2$=0.0508

The constant value of 16.9 is the average aggregate contribution to the group good after accounting for time dynamics, design and experiment type. The null hypotheses that this value is 18 (3*6) cannot be rejected (F-test p-value=0.4735). The t^{-1} coefficient is significant and positive. This suggests that, like other public good experiments, these subjects over-contributed in the early periods and their contribution decreased as time passed. In other words, the results from the restricted contribution phases are not being generated by an anomalous subject pool, at least as far as these characteristics are concerned.

Result 5

> **Significant deviations away from the standard zero-conjectural variations Cournot model are made, but these deviations decline towards zero over time.**

As is evident in Graph 4, the histogram of deviations from the Cournot prediction are all centred around zero. The variance in deviations declines when only the final period is considered. In Graph 5, the time path of deviations is presented. This information is ·ımmarized in Table 7 below.

TABLE 7

Tobit Regression of Deviations from Cournot Predictions
Lower Bound -12

Variable	Coefficient	t-stat	p-value
t	- 0.56751	- 4.358	0.000
D	2.35258	3.247	0.001
S	1.53838	2.362	0.019
constant	3.43056	4.081	0.000

Although the constant is positive, the time coefficient is significantly negative suggesting that people's over-contribution disappears over time. This directly contradicts the warm glow theory. ı a period by period test of the mean deviation in both the ABC and BAC designs of the ɔ*rtners* experiment and in the *strangers* experiment, it is impossible to reject the null of zero deviations by period 5,3 and 4 respectively. This data does not support even a weaker model of warm glow.

There is a fundamental difference between the *partners* and *strangers* data in the unrestricted contribution phase (Graphs 6 and 7). In both experiments, the largest number of individual contributions occur at 0 tokens. The number of 0 plays far outweighs the number observed in the maximum contribution phase, and can be labelled strong free-riding behaviour instead of vindictiveness. In the *partners* experiment, peaks occur at 0, 6 and 12 (and 13 if all periods are considered). In the *strangers* experiment, peaks occur at 0, 4 and 12. In the unrestricted phase of either experiment any contribution between 0 and 12 is rationalizable (Osborne & Rubenstein, 1994)[11]. Thus, in the *partners* experiment, subjects seem to be using cooperative signalling strategies (13), some variant of a reactive Nash strategy (6), a strong rationalizable strategy (0 and 12), and a free-riding strategy (0). However, the strategies used the *strangers* experiment seem to be limited to strong rationalizable (0 and 12), free-riding ı) and an unidentified strategy (4). Partners seem to attempt to cooperate **and** free-ride more than strangers in this environment. Neither group shows a disposition towards warm glow or cold prickle.■

[11] A rationalizable solution is iteratively dominant solvable and generally includes many more solutions than the Nash. For instance, subject 1 believes the other two subjects will play 12 each, then their best response is to play 0. Subject 1 is correct in this belief if subject 2 believes 1 and 3 will contribute 0 each and subject 3 believes that 1 and 2 will each contribute 0. Similar reasoning shows that any triplet of individual contributions of 0 through 12 is rationalizable with some belief structure.

Result 6

> **Individual contributions in the unrestricted phase do not entirely account for deviations in either of the restricted phases.**

The data for the entire experiment is presented in Graphs 8 (a, b, c). Graphs 9 (a, b, c) contain similar graphs for the final period contributions. First notice in Graph 8a the large frequency of plays at 0 token contribution (18.17%). This result is also evident in Graph 9a (18.67%). Although this suggests that the censored unrestricted contribution data is shifted to the left of the minimum (13) contribution data (i.e. suggesting that generosity exists) -- see Graph 8b -- it also suggests that the censored unrestricted contribution data is shifted to the left of the maximum (6) contribution data (i.e. suggesting that the unrestricted data shows an even larger amount of vindictiveness) -- see Graph 8c. When only final period data is considered, the censored unrestricted contribution distribution closely matches the minimum contribution data - see Graph 9b. However, similar analysis with the maximum contribution data again shows the censored unrestricted contribution distribution shifted to the left -- see Graph 9c.

The results reaffirm the conclusion in Result 5 that there are significant deviations away rom the ZCV Cournot prediction. Subjects repeatedly contributing 0 tokens may have negative conjectural variations, anti-altruism, may have rationalized that 0 is the 'best' action, or may be content to free-ride upon the contributions of others and hope that others will contribute on their behalf. The effective strategy space (see footnote 2) within this experiment is between 0 and 13 tokens. In the unrestricted contribution data only 7.01% of all observations (4.00% of final period observations) exceed 13. This data suggests that they certainly act within the strategic space. Furthermore, notice the additional (small) peaks in the unrestricted contribution distribution at token levels of 12 and 13 in the inclusive data (Graph 8a) and 12 in the final period data (Graph 9a). For the most part, token contributions of 12 are associated with large positive deviations from the ZCV Cournot prediction and may be interpreted as signals for desired cooperation. Thus, this data suggests that individuals behave in a conscious manner attempting to illicit cooperative behaviour from others while in an anonymous, no communication environment.■

Result 7

> **There are significant design effects in overall average contribution during each of the minimum/maximum contribution phases. Specifically, subjects in the ABC design contribute more tokens per period on average in the minimum contribution phase and less tokens per period on average in the maximum contribution phase, than subjects in the BAC design.**

For each phase the null hypothesis is $H_0: \overline{g_u^{ABC}} = \overline{g_u^{BAC}}$ with $H_A: \overline{g_u^{ABC}} \neq \overline{g_u^{BAC}}$. In order to test this hypothesis, average contributions in each design are calculated for each of the eight periods. An exact randomization test is used to test whether the overall average period contribution is significantly different between designs. In the case of the minimum contribution phase, the difference in means (ABC - BAC) is 0.8194 which has an associated p-value of 0.0035. Similarly, in the case of the maximum contribution phase, the difference in means (ABC - BAC) is -0.2927 which has an associated p-value of 0.0216. In both cases, these are sufficiently low to reject the null hypothesis that the two means are equal.

These results would seem to suggest that the subjects in the ABC design are on average both more generous and more vindictive. If a lack of understanding about the payoffs were the sole reason why subjects might contribute some number other than 13 or 6 then we would expect the ABC mean to be higher during the minimum contribution phase and the BAC mean lower during the maximum contribution phase. Certainly this is not the case. Subject pool effects may be evident, as the majority of the ABC design was run with undergraduate students at the University of Arizona, while the BAC design was run entirely with a mix of graduate and undergraduate students at McMaster University.■

VII. Conclusion

The warm glow theory of public good voluntary contributions predicts the general tendency of over-contribution relative to the Nash equilibrium (Andreoni, 1989, 1990). Furthermore, it predicts incomplete crowding-out which has important tax policy implications. However, a number of other theories share these predictions.

The effects of generosity cannot be separated from the effects of strategic signalling, non-zero conjectural variations, free-riding, punishment strategies, etc., when voluntary contributions lie between Nash and Pareto equilibrium predictions. Some subjects may be predisposed towards cooperation. Thus, a reasonable method for testing for generosity is to require individuals to contribute at least the Pareto amount and then see if they voluntarily contribute more. Only then are we able to see if they are willing to pay to be generous.

The results of this experiment suggest that subjects do not on average exhibit noticeable levels of generosity or vindictiveness. When subjects make strictly voluntary contributions, they use the entire decision space but only a negligible amount of contributions above the Pareto equilibrium are observed. This might suggest that subject types are quite diverse in terms of their cooperative and free-riding natures, and that learning about others is quite difficult in an anonymous, non-communicative environment. *Homo economicus* is a complex individual operating within a social network. Each individual strives to maximize its own utility while realizing that concerted group effort may ultimately improve its own rewards. Some individuals in the species show signs of cooperativeness, strategic play, altruism, and group thinking, but there is little evidence to support the claim that they are generous. For policy, this implies that mechanisms or institutions should be designed which enhance cooperation instead of relying upon taxation's incomplete crowding out effects.

Bibliography

Andreoni, J. (1988). Privately provided public goods in a large economy: The limits of altruism. Journal of Public Economics 35: 57-73.

Andreoni, J. (1989). Giving with impure altruism: Applications to charity and Ricardian equivalence. Journal of Political Economy 97(6): 1447-1458.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm glow giving. The Economic Journal 100(104): 464-477.

Andreoni, J. (1993). An experimental test of the public-goods crowding-out hypothesis. American Economic Review 83(5): 1317-1327.

Andreoni, J. (1995). Warm-glow versus Cold-prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments. Quarterly Journal of Economics 60(1): 1-21.

Bagnoli, M. & McKee, M. (1991). Voluntary contribution games: Efficient private provision of public goods. Economic Inquiry 29: 351-366.

Barro, R.J. (1974). Are government bonds net wealth? Journal of Political Economy 82: 1095-1117.

Becker, G.S. (1974). A theory of social interactions. Journal of Political Economy 82: 1063-1093.

Bergstrom, T.C., Blume, L. & Varian, H. (1986). On the private provision of public goods. Journal of Public Economics 29: 25-50.

Chan, K., Mestelman, S., Moir, R. & Muller, R.A. (1993). Voluntary Contributions to Public Goods Under Varying Endowments. *Forthcoming* in the Canadian Journal of Economics.

Chan, K., Godby, R., Mestelman, S. & Muller, R.A. (1994). Spite, Guilt and the Voluntary Provision of Public Goods when Income is not Distributed Equally. Department of Economics Working Paper #94-08. McMaster University.

Friedman, J.W. (1971). A non-cooperative equilibrium for supergames. Review of Economic Studies 38(1): 1-12.

Isaac, R.M., Walker, J.M. & Thomas, S.H. (1984). Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations. Public Choice 43(2): 113-149.

Isaac, R.M. & Walker, J.M. (1988). Communication and free-riding behaviour: The voluntary contribution mechanism. Economic Inquiry 24: 585-608.

Isaac, R.M., & Walker, J.M. (1992). Nash as an organizing principle in the voluntary provision of public goods: Experimental evidence. Manuscript.

Kennedy, P.E. (1995). Randomization tests in econometrics. Journal of Business and Economic Statistics 13(1): 85-94.

Li, Ya. (1991). A game theoretic model of public good provision with positive Nash equilibria and its experimental evidence. Department of Economics Discussion Paper #91-10. University of Arizona.

Moir, R. (1995). A Monte Carlo Analysis of the Small Sample Properties of Parametric, Nonparametric and Exact Randomization Tests. *Dissertation*. McMaster University.

Osborne, M.J. & Rubenstein, A. (1994). A Course in Game Theory. MIT Press: Cambridge, Massachusetts.

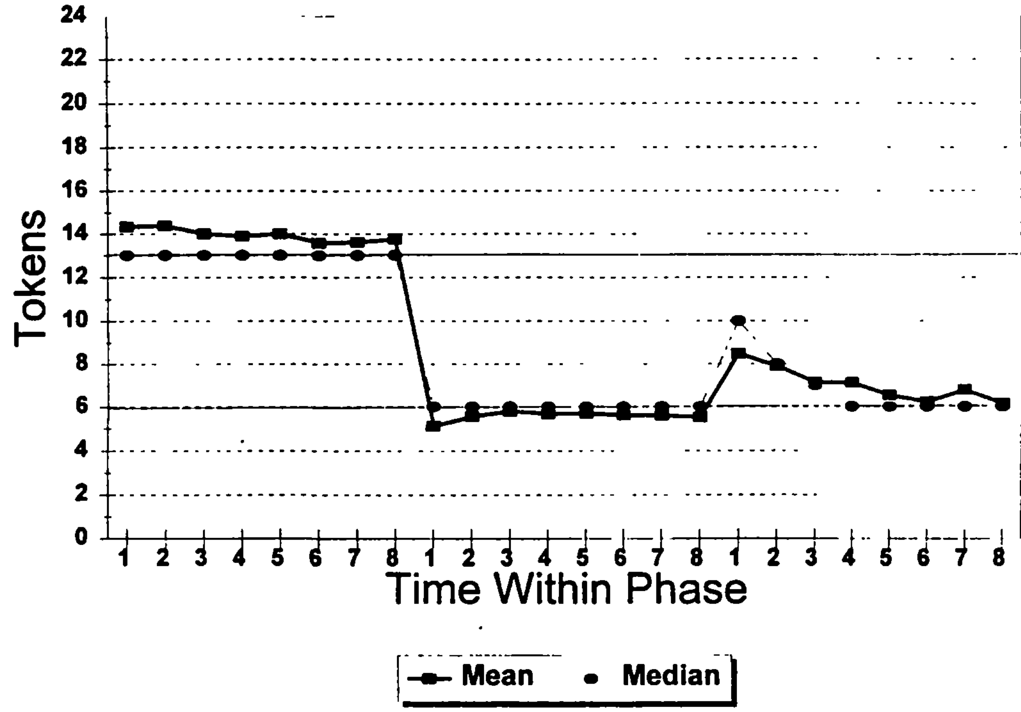
Palfrey, T. & Prisbrey, J. (1993). Altruism, reputation, and noise in linear public goods experiments. Unpublished Manuscript.

Radner, R. (1980). Collusive behaviour in noncooperative epsilon-equilibria of oligopolies with long but finite lives. Journal of Economic Theory 22(2): 136-154.

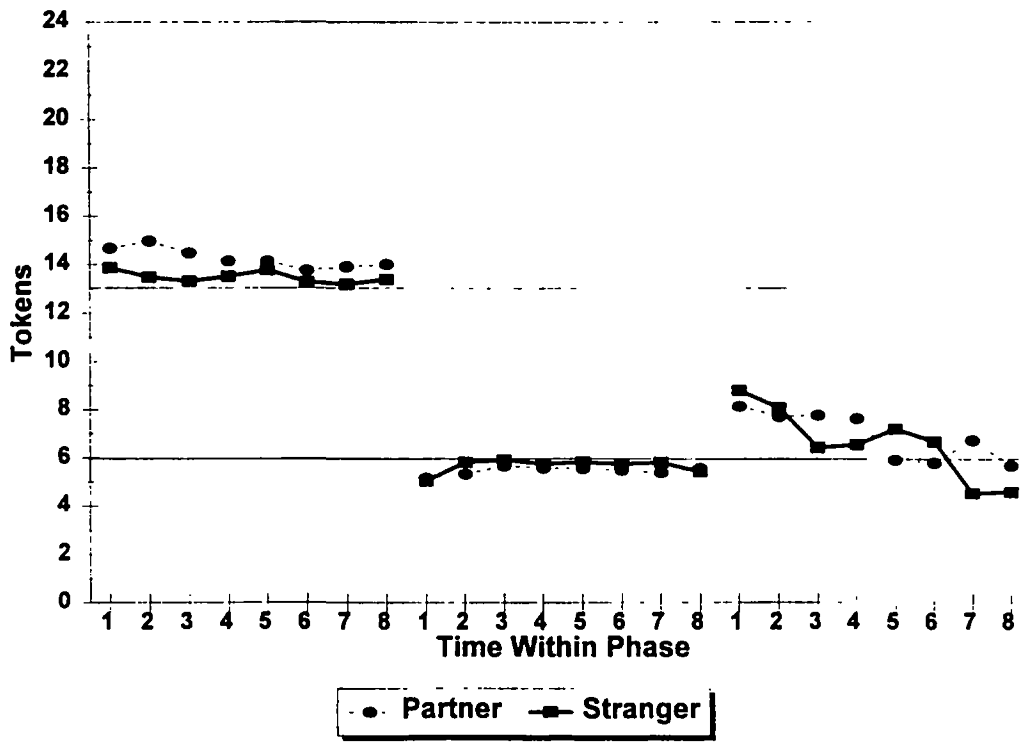
Roumasset, J.A. (1991). Constitutional management of free riders and common property resources. Osaka Economic Papers 40(3,4) (March): 322-330.

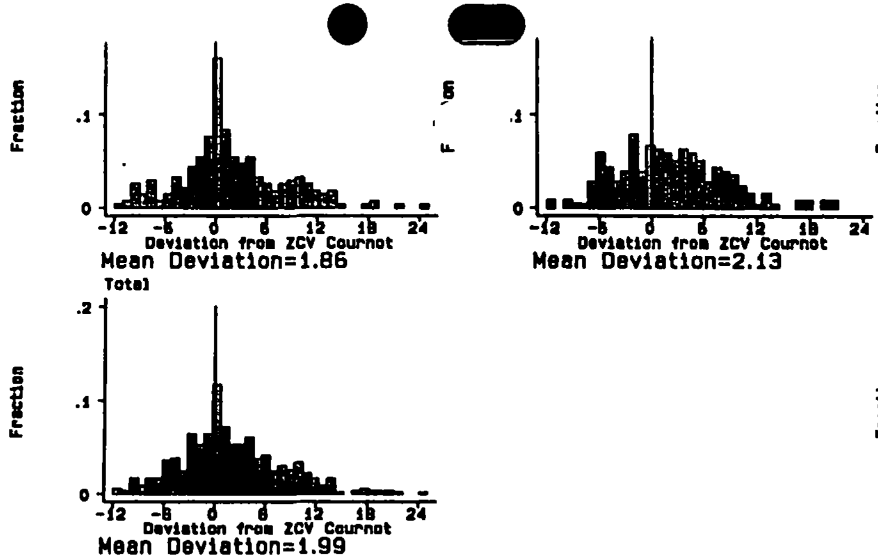
Warr, P.G. (1982). Pareto optimal redistribution and private charity. Journal of Public Economics 19: 131-138.

GRAPH 1: Mean and Median Contributions by Phase and Time

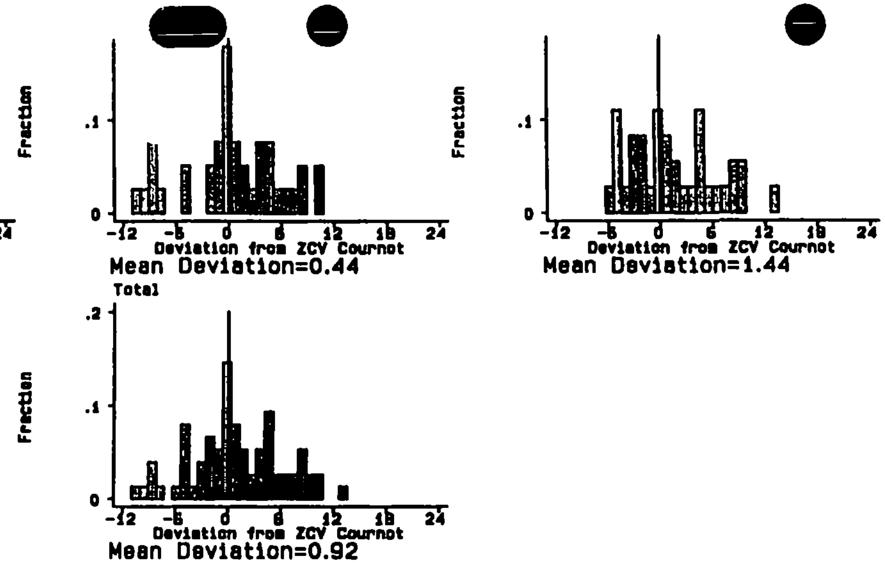


GRAPH 2: Mean Individual Contribution by Time and Phase



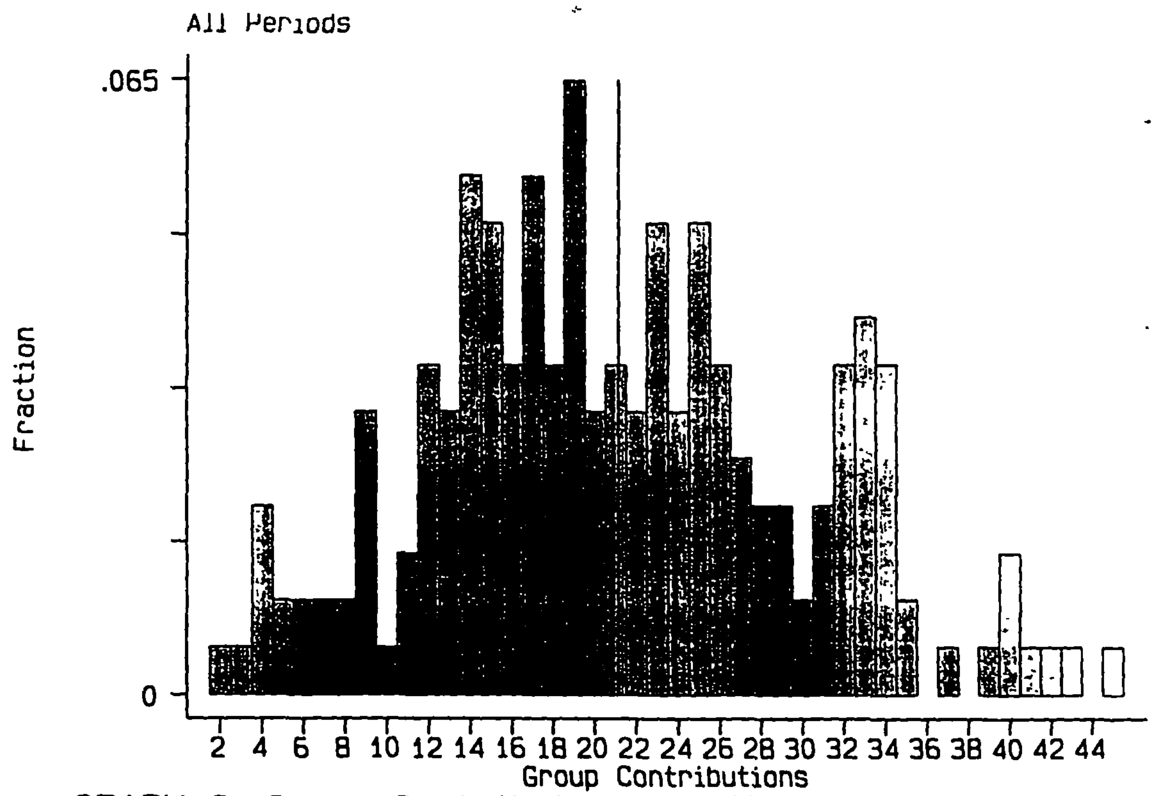


Periods 2-8

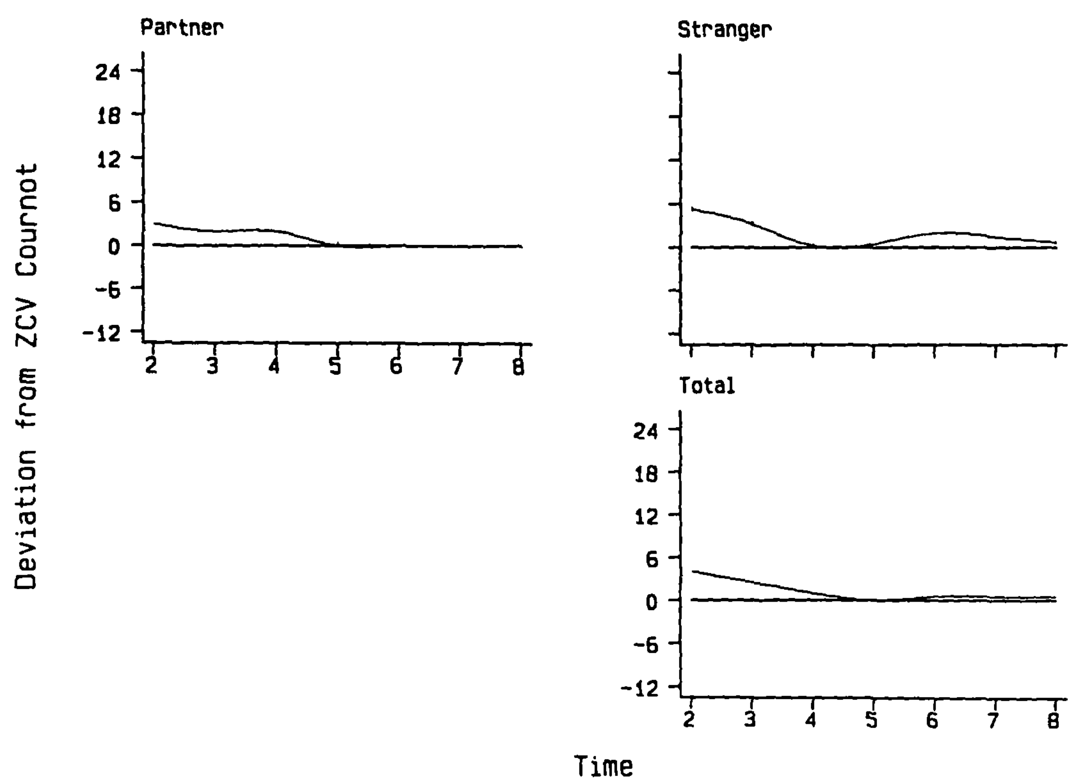


Period 8

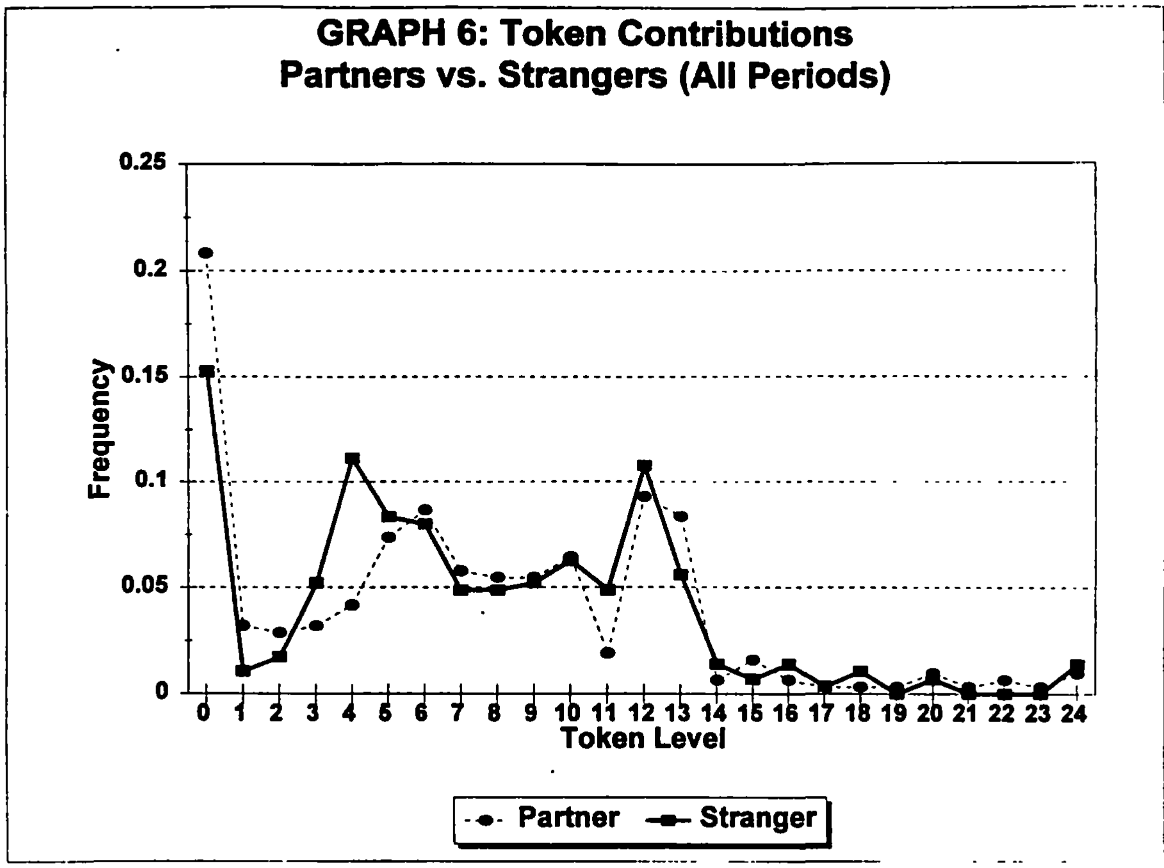
GRAPH 4: Deviations from ZCV Cournot Prediction

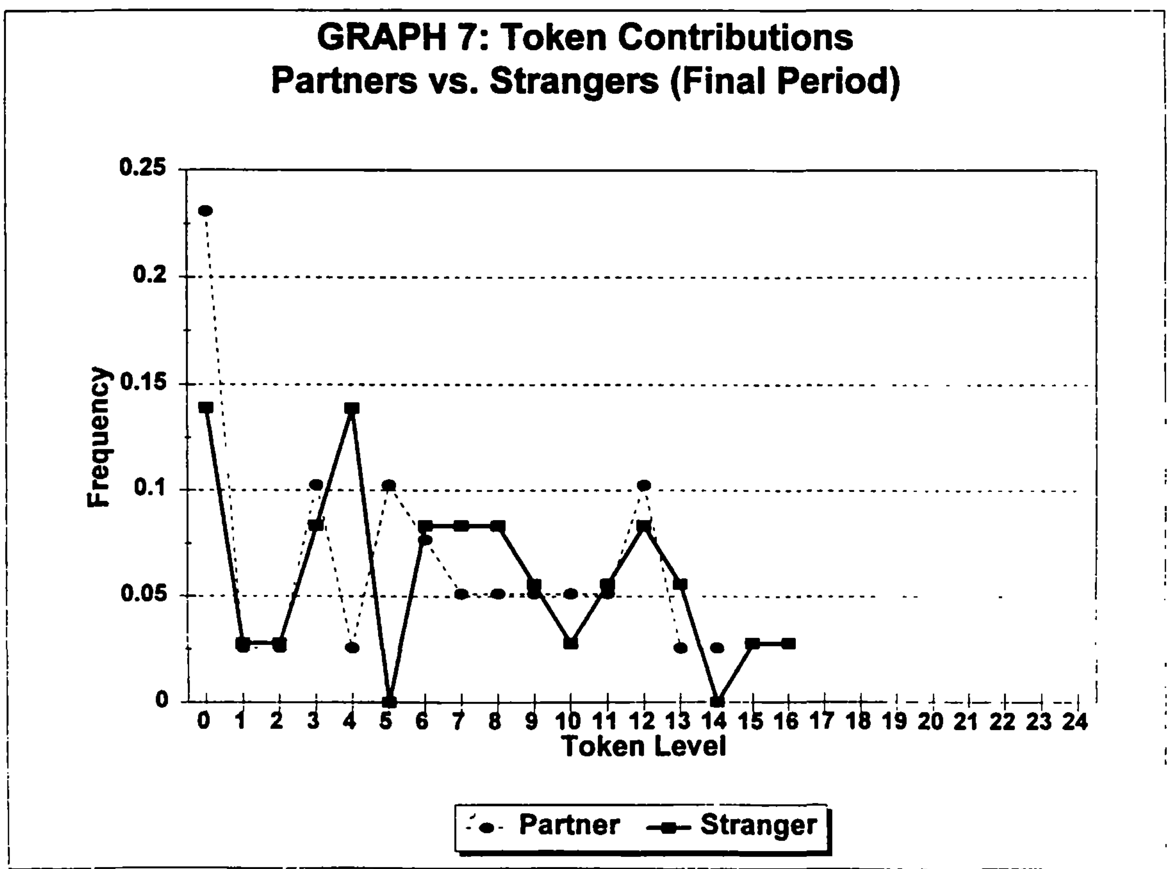


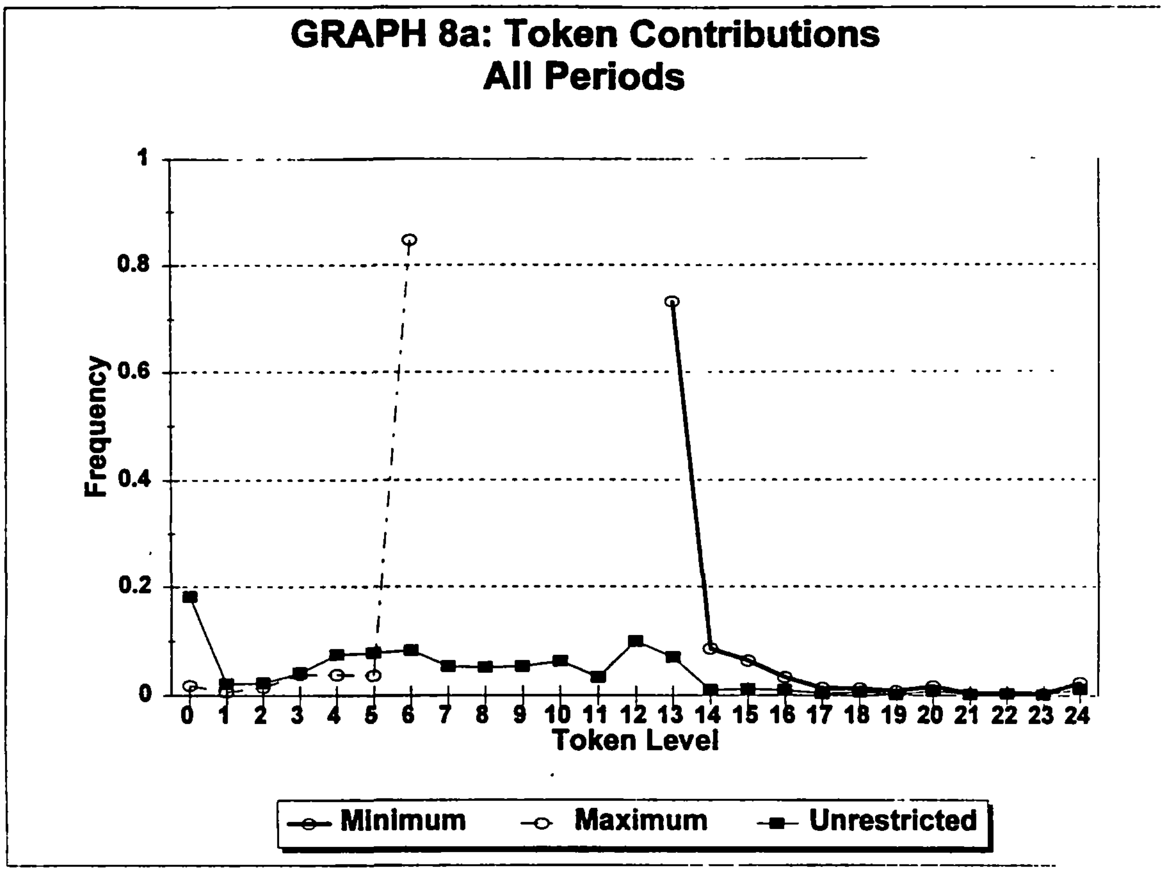
GRAPH 3: Group Contributions in Unrestricted Phase

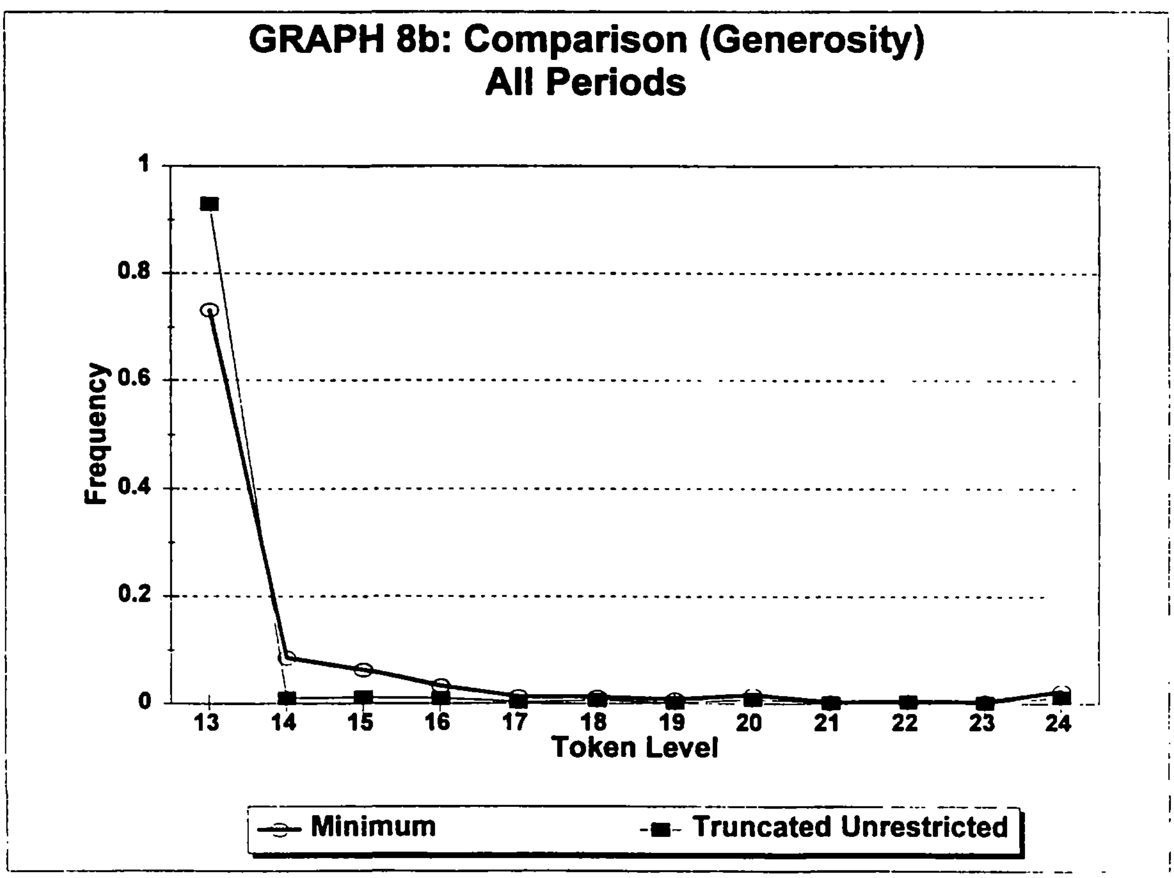


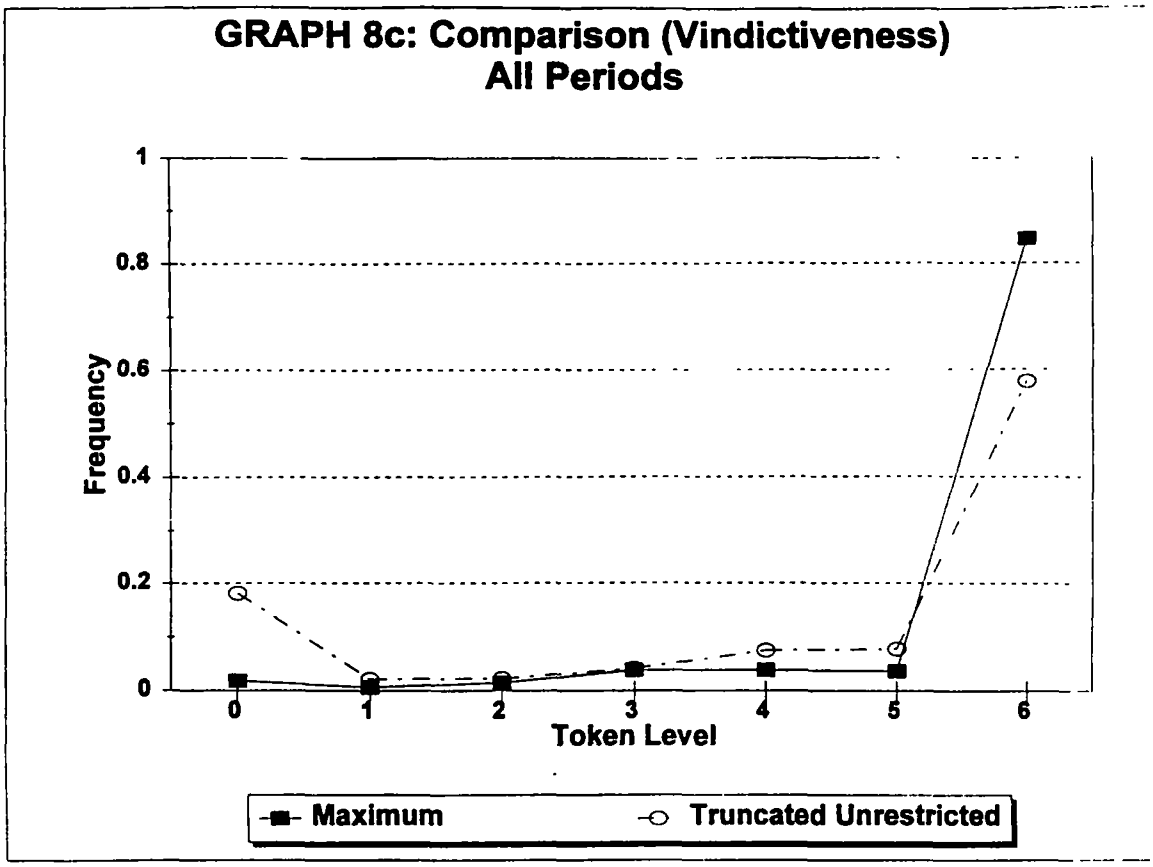
GRAPH 5: Deviations Through Time

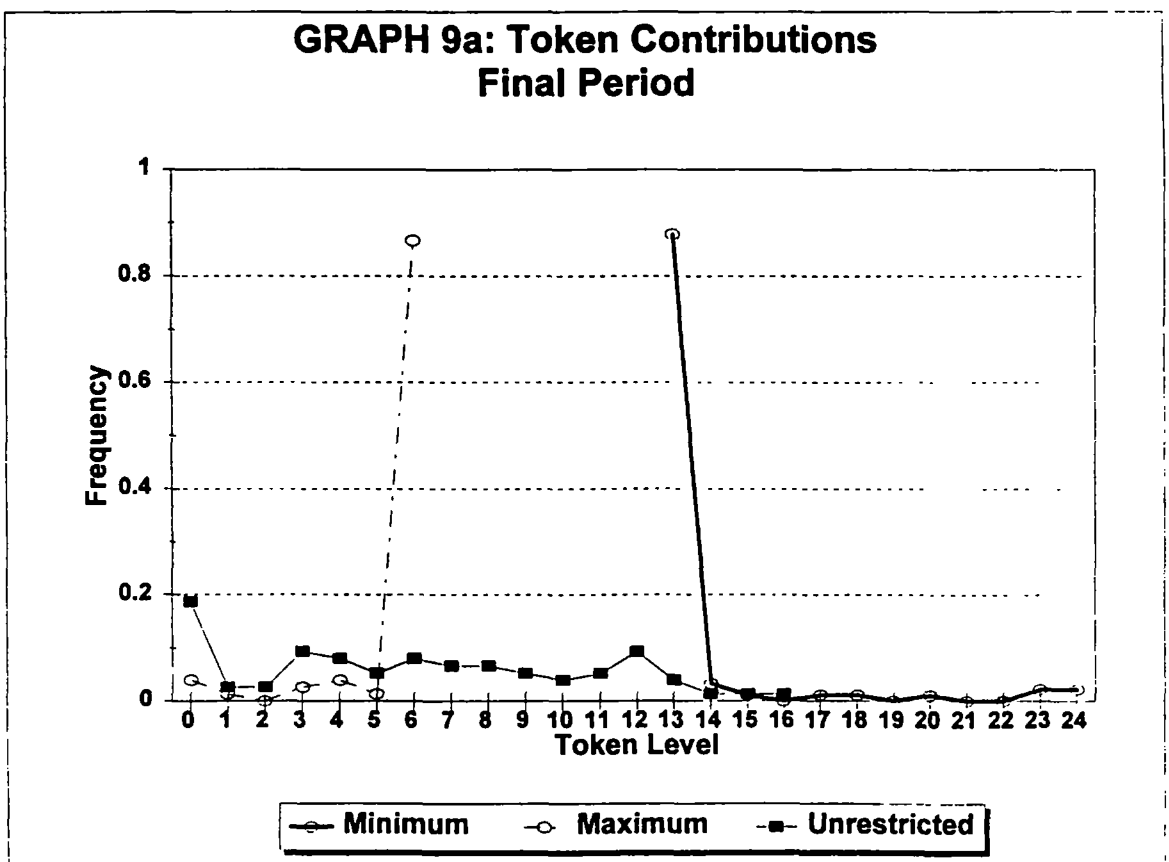

GRAPH 6: Token Contributions
Partners vs. Strangers (All Periods)
Frequency
0.25
0.2
0.15
0.1
0.05
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Token Level
Partner
Stranger

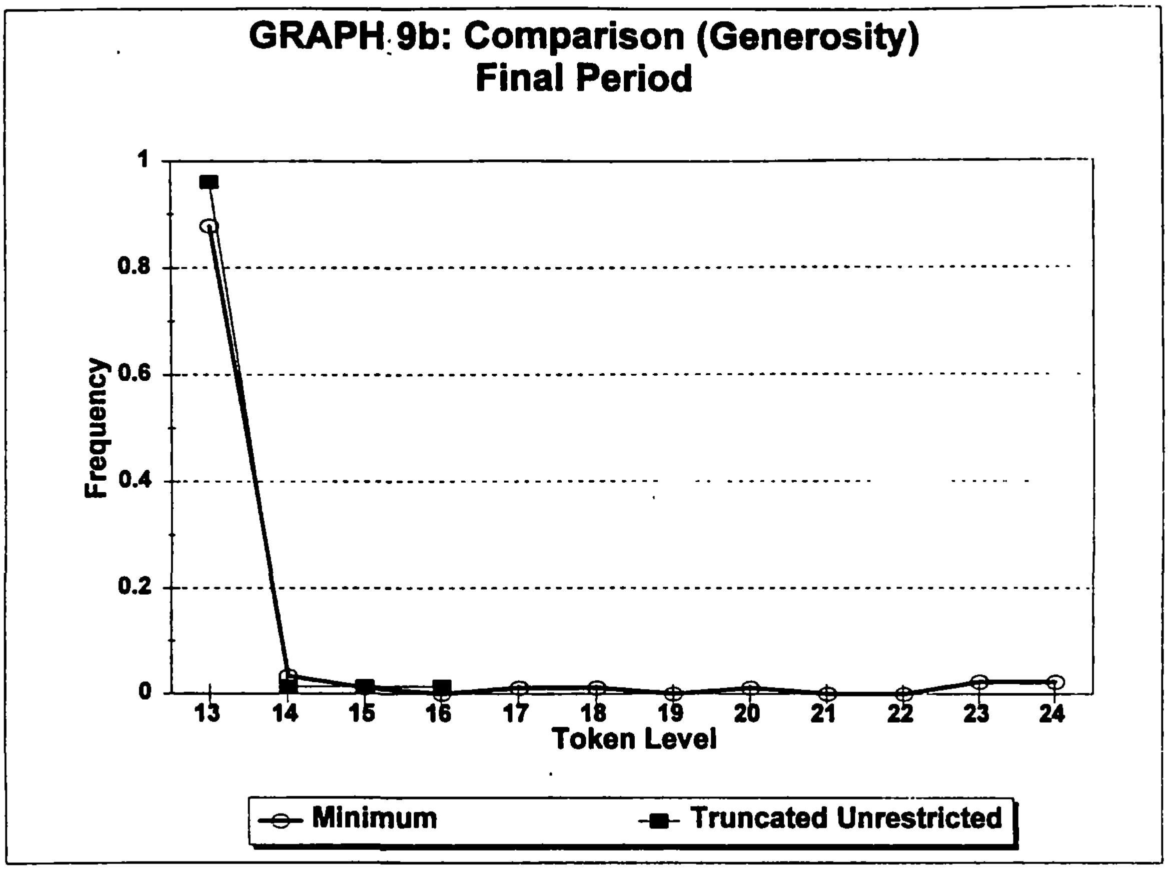

GRAPH 7: Token Contributions
Partners vs. Strangers (Final Period)
Frequency
0.25
0.2
0.15
0.1
0.05
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Token Level
Partner
Stranger


GRAPH 8a: Token Contributions
All Periods
Frequency
1
0.8
0.6
0.4
0.2
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Token Level
Minimum
Maximum
Unrestricted

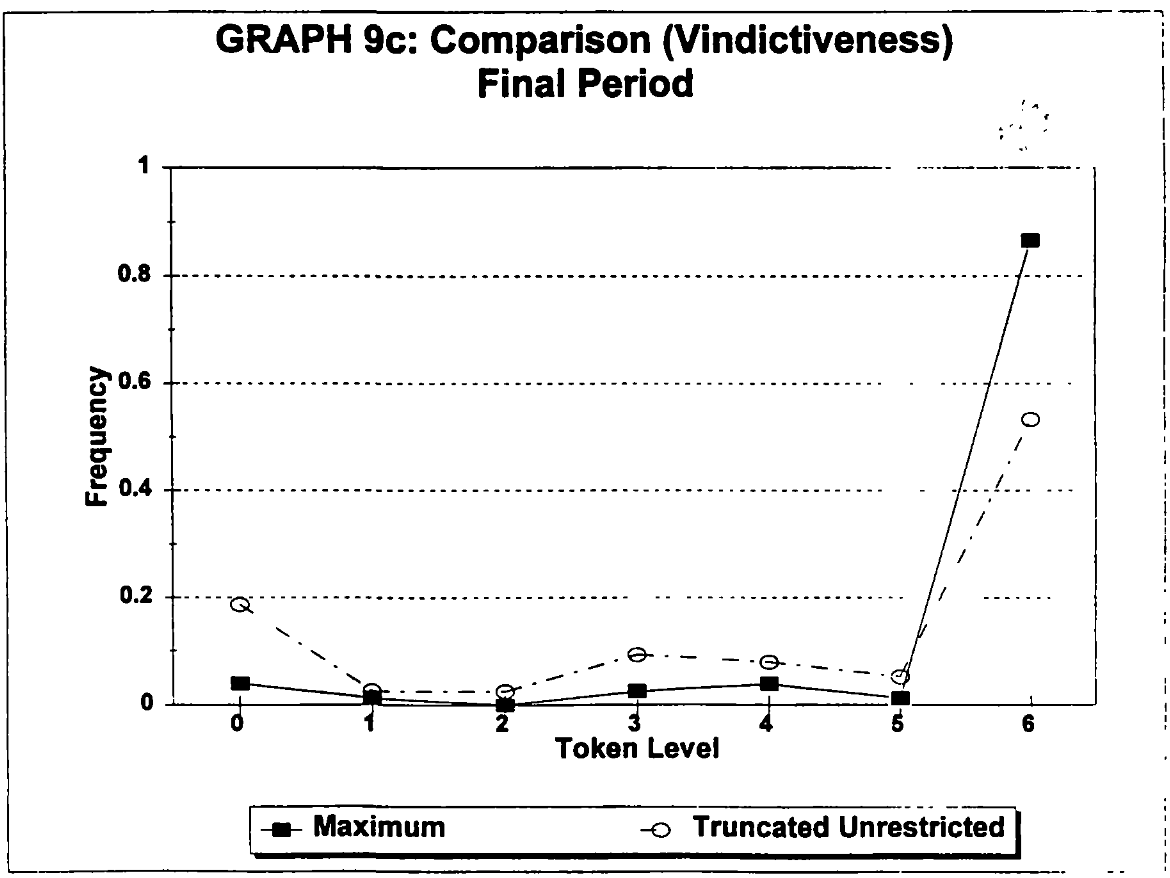

GRAPH 8b: Comparison (Generosity)
All Periods
Frequency
1
0.8
0.6
0.4
0.2
0
13 14 15 16 17 18 19 20 21 22 23 24
Token Level
Minimum
Truncated Unrestricted


GRAPH 8c: Comparison (Vindictiveness)
All Periods
Frequency
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
Token Level
Maximum
Truncated Unrestricted


GRAPH 9a: Token Contributions
Final Period
Frequency
1
0.8
0.6
0.4
0.2
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Token Level
Minimum
Maximum
Unrestricted



GRAPH 9b: Comparison (Generosity)
Final Period
Frequency
1
0.8
0.6
0.4
0.2
0
13
14
15
16
17
18
19
20
21
22
23
24
Token Level
Minimum
Truncated Unrestricted



GRAPH 9c: Comparison (Vindictiveness)
Final Period
Frequency
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
Token Level
Maximum
Truncated Unrestricted