
95-03-04

Rachel Croson
A NOTE ON AND REPLICATION
OF ANDREONI'S

Why Free Ride? Strategies and Learning
in Public Goods Experiments:
PARTNERS AND STRANGERS REVISITED

Rachel T. A. Croson

95-03-04

Department of Operations and Information Management
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6366

Also available as a working paper of the Risk Management and Decision Processes Center
A Note on and Replication of Andreoni's
Why Free Ride?: Strategies and Learning in Public Goods Experiments:
Partners and Strangers Revisited*

Rachel T.A. Croson
CrosonR@opim.wharton.upenn.edu
October, 1993
(Revised March, 1995)

Abstract

This note reports experimental results from a replication of Andreoni (1988) public goods experiments. Two hypotheses had been proposed to explain declining contributions in laboratory voluntary contribution mechanisms: simple learning and strategies. Andreoni's results were not consistent with either hypothesis. The results reported here are not consistent with the simple learning hypothesis, but are compatible with the strategies hypothesis.

*The author thanks Jim Andreoni, Jon Baron, Jerry Green, Mark Isaac, Eric Maskin, Jeff Prisbrey and George Wu for helpful comments and Jim Andreoni for providing data from his 1988 experiments. All omissions or mistakes are the responsibility of the author. Funding of experiments from the Economic Science Lab at the University of Arizona is gratefully acknowledged.
1. Introduction

US individuals made over 100 billion dollars of philanthropic contributions in 1993 (Giving USA, 1994)\(^1\) This behavior is inconsistent with economic theory in that the recipients of these contributions usually provide public goods. The public goods literature shows that optimal contributing behavior often involves “free-riding” on others’ contributions by contributing nothing. Much research in economics has focused on developing incentive-compatible mechanisms designed to overcome this free-riding problem.

Experimental economists have studied the free-riding hypothesis extensively. When subjects play a repeated public goods game, contributions to the public good start at around half the experimental endowment and decrease over time. Although contributions reach their lowest point in the last period of the game, they do not quite reach the one-shot dominant-strategy equilibrium of full free riding. Two hypotheses have been proposed to explain this decline in contributions: *simple learning* and *strategies*.

The simple learning hypothesis suggests that subjects learn the incentives of the game throughout the experiment. Since subjects learn at different speeds, average contribution should decay over time. If the game is repeated for long enough, all subjects should learn to free ride.

The strategies hypothesis suggests, instead, that subjects are playing with an eye toward influencing their partners’ actions. Kreps *et al.* (1982) describe this process for the finitely repeated Prisoners’ Dilemma. If subjects believe that others don’t understand the free-riding equilibrium, it may be rational for sophisticated players to contribute in order to avoid educating

---

\(^1\) The actual amount contributed by individuals in 1993 was $102,550,000,000. This number excludes charitable giving by corporations, foundations and bequests.
their partners. Similarly, if subjects believe that others are playing a reciprocal or trigger strategy, it may be rational for them to contribute in order to avoid triggering their partners’ defection. Toward the end of the game, however, the value of this strategic play decreases; in the last round, the free-riding strategy is optimal. Strategic play against possibly non-payoff-maximizing opponents (or opponents believed to be non-payoff-maximizing) will thus produce declining contributions.

Andreoni (1988) discusses these two theories and designs an experiment to separate them. Andreoni’s data supports neither the simple learning nor the strategies hypothesis. This note reports an attempted replication of his experiment with slightly different outcomes. The data reported here provide support for the strategies hypothesis only.

2. Public Goods and the Voluntary Contribution Mechanism

Pure public goods are defined as goods that are both nonrivalled and nonexcludable. That is, multiple players can consume the good at the same time (nonrivalled) and it is not possible to exclude players who did not pay for the good from consuming it (nonexcludable). There are many mechanisms which can fund public goods. The one examined in this study—the Voluntary Contribution Mechanism (VCM)—most closely parallels philanthropic giving or contributing behavior in a very stark setting.

Assume each player $i$ in a group of $N$ identical players has some endowment $E$, which can either be contributed to a group account and used to produce units of a public good or can be privately consumed and converted to cash. Call the amount contributed to the group account by $i$, $x_i$. The individual’s earnings from private consumption is simply the amount consumed ($E_i - x_i$). The individual’s earnings from contributions to the group account is a function of the sum of contributions by all participants $P(\sum x_i)$. The group’s earnings is the sum of the individual
earnings and the payouts from the public account $\Sigma_i(E_i x_i) + NP(\Sigma x_i)$. Each individual chooses $x_i$ to maximize his earnings $(E_i x_i) + P(\Sigma x_i)$. We say there is a pure public goods problem when two conditions are satisfied.

Condition 1: Contributions to the private exchange are individually optimal.

$$1 > P(\Sigma x_i)^2 \quad \forall x_i, \forall x_i$$

(1)

Thus regardless of the contributions of the other players, player $i$ never wants to contribute to the group account.

Condition 2: Contributions to the private exchange are not optimal for the group.

$$NP(\Sigma x_i) > 1 \quad \forall x_i, \forall x_i$$

$$P(\Sigma x_i) > \frac{1}{N} \quad \forall x_i, \forall x_i$$

(2)

Thus regardless of the contributions of the other players, the group as a whole earns more when player $i$ contributes to the group account than when he contributes to the private account.

When a game which includes a pure public goods problem is played once there is a unique dominant strategy equilibrium in which all players fully free ride (contribute zero).

When the game is repeated finitely many times (with endowments expiring at the end of each period), contributing zero in all periods is the unique subgame perfect equilibrium. These strong equilibrium predictions are, however, not observed when subjects play public goods games.

3. Previous Experiments

A. Voluntary Contribution Mechanisms

Marwell and Ames (1979, 1980, 1981) were the first to test public goods provision behavior in the VCM. They find that when subjects play a one-shot, context-free public goods

---

2 $P(\Sigma x_i)$ is often called the marginal per capita return (MPCR) and is the marginal return on contributions to the group exchange.
game they contribute around half their endowment to the public good and consume the rest.

When subjects play a repeated public goods game (with a subgame-perfect equilibrium of full free riding), contributions in the first period look similar to those observed in Marwell and Ames, but they decrease over time toward the free-riding solution. Although contributions reach their lowest point in the last period of the game, they do not quite reach the one-shot dominant strategy equilibrium of full free riding (Davis and Holt, 1994) Chapter 6. Two possible causes of the decrease in contributions, simple learning and strategies, were examined experimentally by Andreoni (1988).

B. Testing Simple Learning and Strategies

Andreoni’s experimental design directly tests both the simple learning and the strategies hypotheses. The simple learning hypothesis is tested by the addition of a “restart” in the standard game. Subjects were told they would play a game of ten rounds. At the end of the ten rounds, they were told that there was just enough time to play another game of ten rounds. Since the simple learning hypothesis suggests that the declining pattern of contributions is a result of subjects learning how to play the free-riding equilibrium, it predicts that contributions in the second game of ten rounds should be at or below contributions in the last round of the first game. That is, subjects have learned the incentives in the first game, so in the second game contributions should continue to be low.\(^3\) Were the simple learning hypothesis correct, contributions should decline period-by-period throughout the full twenty rounds.

The variable Andreoni designed to test the strategies hypothesis involves the process governing the matching between members of the contributing group. One treatment, *partners*, mirrors past experiments in that members of a subject’s group remain fixed throughout the

\(^3\)Andreoni truncated the second 10 rounds after 3 rounds of play for financial reasons. This experiment reports all 10 “restart rounds.”
experimental session (including the restart rounds). In this treatment, we would expect to observe strategic play, and consequently declining contributions as predicted by the strategies hypothesis. In another treatment, strangers, members of a subject’s group are randomly re-selected from the pool of subjects before each of the 10 rounds in each game. In this treatment, any given subject’s gains from playing strategically are significantly diminished. Since declining contributions are assumed to be a result of strategic play, we expect to observe uniformly low contribution (i.e., little strategic play) in the strategies treatment.\footnote{Andreoni recognizes that there is a chance for a subject’s actions to feed back to him, however, the incentive to play strategically is significantly diminished in the strangers’ treatment (Andreoni, 1988, p. 294 note 5).} For an in-depth discussion including the details and motivation of this experimental design, see Andreoni (1988).

\section*{C. Prior Experimental Results}

\subsection*{1. Simple Learning}

The results of Andreoni’s study were inconsistent with the simple learning hypothesis. This hypothesis predicted declining contributions over all rounds of play. Instead Andreoni observed a strong “restart effect” in the partners treatment—contributions increased sharply from the last round of the original game to the first round of the restart game. There was a similar but weaker restart effect in the strangers treatment. This effect is not consistent with the simple learning hypothesis, which predicts that subjects, having learned the optimal way to contribute in the original game, should not alter their behavior in the restart game. This restart effect had not been replicated before this study. Here, the restart effect is significant in the partners treatment, and insignificant but arguably present in the strangers treatment, consistent with Andreoni’s results and inconsistent with the simple learning hypothesis.
2. Strategies

The more surprising result of Andreoni's study was the evidence against the strategies hypothesis. Rather than finding strangers (groups re-randomized over time) contributing less than partners (groups constant over time), as the strategies hypothesis would suggest, he found strangers contributing more than partners in all ten rounds of the original game. In addition, fewer strangers free rode (contributed zero) than did partners. In contrast, the data from this replication suggests a statistically significant difference in contributions between the two treatments that is consistent with the strategies hypothesis. Strangers give less than partners and free-ride more.

One other study has since examined the partners/strangers treatment (Weimann (1994)). There were, however, a number of procedural differences between Weimann's and Andreoni's original study. In Weimann's experiment the partners treatment was run in a laboratory and involved direct contact with an experimenter and the other subjects involved. However, the strangers treatment was run over the phone and involved no personal contact with either the experimenter or (more importantly) with the other subjects. In addition, half of Weimann's subjects received information about the individual contributions of the other members of their group, rather than information only about the total contribution of the other members as in Andreoni and previous experiments. Weimann goes on to show that the average contributions in these two treatments are the same. However, related work by Sell and Wilson (1991) which has been replicated by Croson (1995) shows that the variance of contributions in these two treatments differs substantially. Thus pooling the observations from these treatments (as Weimann does) is potentially misleading.\(^5\) Weimann's overall results (after pooling the

---

\(^5\) A control treatment of partners contacted by phone was run in order to compare the differences between phone contact and personal contact. However, this treatment was run in the individual-contribution feedback treatment only and not in the (more standard) total-contribution feedback treatment.
observations from different feedback treatments) suggest no differences between partner and stranger contributions.

One other study which used a partners/strangers design in slightly different settings reports higher variance among contributions by strangers than partners (Palfrey and Prisbrey (1993)). An additional test comparing the variances of contributions in the two treatments in this study shows the variance of strangers contributions to be statistically significantly lower than that of partners overall.

4. Experimental Design and Procedure

Each treatment in this experiment involved 24 subjects arranged in six groups of four. The subjects were recruited from economics classes at the University of Arizona summer school. They were paid a five dollar show-up fee along with their earnings in the experiment. Each treatment took less than one hour to run. Average earnings of the subjects were $12.87 (plus the $5 fee). The entire experiment was computerized. Instructions were given through the computer screen; subjects entered their contribution decisions through the keyboard and, at the end of each period, feedback about their earnings was displayed on the screen.

Each treatment involved a ten-period voluntary contribution mechanism (the original game) with another ten-period game played directly after (the restart game). At the beginning of each period, all subjects were endowed with 25 tokens which could be allocated either to a private account, which paid 2¢ per token to the individual only, or to a group account (the public good), which paid 1¢ per token to each of the four members of the individual’s group. At the

---

6 In addition there were three practice periods before the first game began to familiarize the subject with the computer program and the process. Subjects were not paid their earnings during the practice period. There were no practice periods before the restarted game. Contribution and estimate results for the practice periods as well as the instructions used are available from the author.
end of each period, subjects were told the aggregate contribution of the other three members of group, the total group contribution and their earnings.

Notice that this VCM incorporates a pure public goods problem. Regardless of the decisions of the other players, each individual strictly prefers to place each of his tokens in his private account, earning 2¢, than in the group account, earning 1¢ (Condition 1). However the group as a whole earns 4¢ when a token is placed in the group account (1¢ to each of the four members) but only earns 2¢ when the token is placed in a private account (private earnings of the individual). Thus Condition 2 is satisfied as well.

In the partners treatment the six groups of four which were randomly chosen in the first period of the game remained together throughout both the original game’s 10 rounds and the restart game’s 10 rounds. In the strangers treatment, group membership was re-randomized at the beginning of each period.

5. Experimental Results and Analysis

Figure 1 shows the average subject contribution in each treatment over ten actual rounds and ten restart rounds. The increase in average contribution at the beginning of the second ten rounds (the restart effect) was observed in Andreoni’s experiment as well as this one and is inconsistent with the simple learning hypothesis. In contrast to Andreoni’s results, strangers contributed less than partners in this experiment, with the difference between the two narrowing as the end of each game approached. Both these observations are consistent with the strategies hypothesis.7

---

7 If subjects in the partners treatment are playing strategically, their contributions should decrease toward the end of the game. In the last period, the decision partners are making is identical to that of strangers—there will be no further interaction in either treatment among the members of the group. Thus contributions in the two treatments should be identical in the last period.
A. Simple Learning

Table 1 reports average contributions across the actual and restart rounds. The simple learning hypothesis suggests that contributions will decrease throughout all 20 rounds. Thus contributions in all the restart rounds should be lower than those in round ten of the first game. This is clearly not the case in either the partners or the strangers data.\textsuperscript{8}

Contributions in round one of the restart game are higher than those in round ten of the original game, at the 5\% level of significance for partners and at a nonsignificant level for strangers.\textsuperscript{9} The former result is inconsistent with the simple learning hypothesis. The nonsignificant result in the strangers treatment is consistent with the strategies hypothesis; strangers have no strategic incentive to contribute more in the first period than in any other, thus no significant restart effect should be observed.

However, the data may be consistent with more sophisticated models of learning. Contributions in the restart game are in general lower than those in the original game. A

\textsuperscript{8}A two-sided Wilcoxon test cannot reject at the 5\% level the hypothesis that contributions in a given round are indistinguishable from contributions in that same-numbered restart round with one exception. Contributions in round 7 and contributions in restart round 7 of the partners treatment are significantly different at the 5\% level. This test was run only on independent observations; average group contributions in the partners treatment (n=6, m=6 for each period) and average session contributions in the strangers treatment (n=2, m=2 for each period).

\textsuperscript{9}Using a one-sided Wilcoxon test on independent observations (n=6 m=6 for partners, n=2 m=2 for strangers), we can reject the hypothesis that partners contributed more in round 10 of the original game than in round 1 of the restart game at p<.05 (U=4.00 p=.013). In the strangers treatment that hypothesis cannot be rejected (U=0 p=.167).
blocking test on contributions pooled over all periods suggests this difference to be statistically significant.\textsuperscript{10} Although contributions are lower in the restart rounds than they were in the actual rounds, the strong restart effect in the partners treatment suggests that simple learning is not driving the decline in contributions over the ten round game.

B. Strategies

Table 2 reports the comparison between contributions by partners and strangers. The strategies hypothesis suggests that partners will contribute more than strangers, and that the difference will decrease over the course of the game. In this study, partners always gave more than strangers. Furthermore, the difference between the two treatments did decrease toward the end of each game.

---

Insert Table 2 here

---

Although the difference between contributions by partners and strangers is not significant in any given round it is highly significant over the first ten rounds (the original game), the second ten rounds (the restart game) and over all twenty rounds.\textsuperscript{11} If we pool over rounds and compare average contributions across each entire game, we can reject the hypothesis that partners and strangers contribute the same amount at the 1\% level.

\textsuperscript{10}For each period for each individual, subtract their contribution in the restart game from their contribution in the original game in that same period. Then test these differences to see if they are significantly different from zero. Pooling over all periods, all participants, both partners and strangers gave significantly more in the original game than in the restart game at the 1\% level (partners $z=4.95$, strangers $z=5.82$ n=240 for each). Alternatively we can use a 2-sided Wilcoxon test on contributions to reject at the 5\% level the hypothesis that contributions in the actual game are the same as those in the restart game when pooled over all 10 periods for both partners ($z=3.27, p=.0011$) and strangers ($z=3.31, p=.0009$). This test uses average group and session contributions, here for partners n=60 m=60 and for strangers n=20 m=20.

\textsuperscript{11}In each period (n=6, m=2) contributions are not significantly different between treatments. For the original game (n=60, m=20) a two-sided Wilcoxon test confirms that partners contribute more than strangers do at the 1\% level ($z=2.60, p=.0093$). For the restart game, the same test is similarly significant ($z=2.70, p=.0067$). Over both games using all 20 rounds (n=120, m=40), partners contribute significantly more than strangers ($z=3.31, p=.0009$).
We can also look at the proportion of free riders (subjects contributing zero tokens) in each treatment. Table 3 compares these proportions over the partners and strangers treatments. In 11 out of the 20 rounds significantly more strangers than partners free ride. Rounds at the end of each ten-round game are not statistically different, which is consistent with the strategies hypothesis (the treatments should converge toward the end of the game). Additionally, the proportion of free riders in each ten-round game and over both games is significantly different.\textsuperscript{12}

\begin{center}
Insert Table 3 here
\end{center}

In this replication, unlike Andreoni's original experiment, the difference between strangers and partners contributions is consistent with the strategies hypothesis. Partners contribute significantly more and free ride significantly less than strangers do, as the strategies hypothesis predicts. The differences between the treatments decrease over time, again as predicted by the strategies hypothesis. In section 6 below, I discuss some procedural differences between the two experiments. However, additional replication is needed to test the strategies hypothesis in this setting.

C. Variance of Contributions

Another experiment which used a strangers treatment under different conditions has reported higher variance in contributions by strangers than by partners (Palfrey and Prisbrey

\textsuperscript{12}Using a t-test of proportions on the proportion of free riders, we conclude that those proportions are different at the 5% level in 11 out of the 20 rounds. The t-test examines differences in proportions, if \(p_1\) is the proportion of free riders in treatment 1 and \(n_1\) is the number of observations in the treatment then

\[ t = \frac{p_1 - p_2}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \]

Here, \(n_i\) for each period is 24 individual contributions. Over each game 240 observations were used. Over both games 480 observations were used. The proportion of free riders in each treatment averaged over the actual rounds, the restart rounds and all the rounds differ at the 5% level using the t-test.
(1993)). Table 4 below reports the standard deviation from each period’s contributions in this study. In only three of the 20 periods did the variance of contributions by strangers exceed that of partners.\textsuperscript{13}

Insert Table 4 here

We can use an F-test to examine whether the variances in contributions are different in the two treatments. The results suggest rejecting the hypothesis that the two series have the same variance at the 1% level. Contributions by partners, exhibit significantly more variance than those by strangers, contrary to the results of Palfrey and Prisbrey. Table 5 reports the standard deviations over the original game, the restart game and all 20 periods as well as the associated F-statistics.\textsuperscript{14}

Insert Table 5 here

The results reported in this replication of Andreoni’s study provide evidence against the simple learning hypothesis by validating the restart effect. They also provide strong and robust support for the strategies hypothesis (partners contribute more than strangers, free ride less than strangers and the differences between the treatments decrease over time) in contrast to Andreoni’s results. The next section discusses procedural differences between this study and Andreoni’s experiments.

\textsuperscript{13}Following Weimann (1994) here I report standard deviations of individual contributions (n=24, m=24). The statistical tests below use independent observations (n=6, m=2).

\textsuperscript{14}An F-test for equality of variances uses average group and session observations (n=60, m=20 for the original and restart games; n=120, m=40 for both games).
6. Procedural Differences

Although the experimental design in this note and Andreoni’s study were the same, some differences in parameter values and procedures are discussed in this section. A direct comparison of data between two different experiments is usually problematic; nonetheless Figures 2 and 3 provide graphical representations of contributions in the two experiments. Preliminary statistical analysis indicates cross-experiment agreement on partner contributions but not on those of strangers.\textsuperscript{15}

\hspace{1in} Insert Figure 2 here

\hspace{1in} Insert Figure 3 here

A. Subjects

The group sizes in both treatments of this study were slightly smaller than in Andreoni’s which might affect contributions. Table 6 below depicts the differences.

\hspace{1in} Insert Table 6 here

\textsuperscript{15}Contributions are transformed into contributions as a percentage of endowment. Using a 2-sided Wilcoxon test and restricting our attention to those periods that were played in both treatments and by both experimenters we can conclude that there is no significant difference between contributions in the partners treatment in the two experiments, but that strangers contributed significantly more in Andreoni’s experiment than in this one. In the partners treatment the test is run on average group contributions pooled over the periods (n=60 m=60) and we cannot reject the null hypothesis of same contributions at the 5% level (z=1.449 p=.148). In the strangers treatment the test is run on average session contributions pooled over the periods (n=20 m=20) and we can reject the null hypothesis of same contributions at the 5% level (z=2.245 p=.024). Many thanks to Jim Andreoni for providing the data from his 1988 experiment for comparison.
Andreoni’s study involved four strangers groups of five subjects each and three partners groups of five subjects each, all playing a normal ten-round game. Another four groups of five strangers and three groups of five partners played a ten-round game plus a three-round restart (n_s=40, n_p=30 for ten original rounds, n_s=20, n_p=15 for three restart rounds). This study used two sessions for each treatment, with three groups of four in each session, all of who played both an original game and a full ten-round restart game (n_p=24, n_s=24 for ten original rounds and ten restart rounds).

In the strangers treatment, the smaller group size in this study implies that the probability of meeting someone you have met before is slightly higher than it was in Andreoni’s study. If anything, this should have led strangers to play more strategically (and thus more like partners) in this study than in Andreoni’s. However, this effect was not observed; instead, strangers played nonstrategically in this study, contributing less than partners.

The two experiments also involved a change in subject pool. Subjects in Andreoni’s experiment were recruited from undergraduate Introductory Economics courses at the University of Michigan. Subjects in this experiment were recruited from summer Economics courses at the University of Arizona.

B. Earnings

Participants in this experiment were paid a show-up fee of $5, plus their earnings. Andreoni used no show-up fee, but paid subjects entirely from their earnings in the experiment. In Andreoni’s study, subjects were endowed with 50 tokens in each round and earned 1¢ per token when investing in the private good and (along with each other member of the group) 1/2¢ per token when investing in the public good. In the current experiment, subjects were endowed with 25 tokens and earned 2¢ per token when investing in the private good and (along with each other member of the group) 1¢ per token when investing in the public good. This study lasted
under one hour. Subjects earned $5 for showing up and an average of $12.87 for their decisions.

The marginal per capita return (MPCR) which has been shown to be an important variable in voluntary contribution mechanisms (Isaac and Walker (1988)) was the same in both studies (2:1)\(^6\)

C. Computers

This study was completely computerized. Subjects signed in, collected their show-up fee and sat at a computer terminal. Instructions were given via the computer and contribution decisions were entered via keyboard. Andreoni’s subjects recorded their decisions by hand; a computer was used for the randomization and to calculate earnings.

D. Practice Rounds and Instructions

Instructions for the two experiments are available from the respective authors and are roughly compatible. In this experiment, subjects played three practice rounds to familiarize themselves with the workings of the program. In Andreoni’s study, no practice rounds were played, but instructions were read aloud and examples provided.

7. Conclusion

After reporting his experimental results, Andreoni concludes that “neither strategies nor learning can be supported as explanations of decay in public goods experiments.” (p. 300). This note reports a replication of Andreoni’s experiment in which strategies can be supported as an explanation of decay in public goods experiments, although simple learning cannot.

The results reported here do not support simple learning as an explanation of declining contributions in public goods experiments. The restart effect is significant (p<.05 in the partners treatment), and the distribution of offers in nine rounds of the actual game could not be

\(^6\) The MPCR is the ratio between a subject’s payoff from the private account and the public account.
distinguished from the distribution of offers in their restart counterpart rounds. These results are similar to Andreoni's. However, the results may be consistent with more complicated forms of learning. Overall contributions in the restart game were significantly lower than those in the original game.

The results reported here are consistent with strategies as an explanation of declining contributions in public goods experiments. The differences in contributions between partners and strangers are in the correct direction (partners contribute more than strangers do) and are significantly different in each individual game and in both games combined (p<.01 for all three comparisons). Strangers free ride (contribute zero) significantly more than partners do as well. The strategies hypothesis predicts that contributions in the two treatments will converge toward the end of the game, which, in this experiment, they do. In contrast, Andreoni found strangers contributing more than partners in each round of the original game, with the difference between the two increasing toward the end of the game.

Palfrey and Prisbrey (1993) find more variance in strangers contributions than in partners in various settings. This study finds higher variance in contributions by strangers than partners in only three of the 20 rounds. An F-test on contributions finds significantly more variance in partners contributions than in strangers over each game and both games pooled (p<.01 for all three comparisons).

This study provides evidence against the simple learning hypothesis and for the strategies hypothesis, in contrast to Andreoni (1988) who found evidence against both. The data may also be consistent with more sophisticated models of learning. The difference in results between studies highlights the importance of replication for economics experiments. We are searching for robust behavior: only by replicating experiments can we safely draw conclusions about our investigations.
References


Figure 1
Average Contributions

Number of tokens

Period #

- □ Partners
- ■ Strangers
Figure 2
Average Contributions by Partners

Number of tokens

Period #

- Andreoni
- Croson
Figure 3
Average Contributions by Strangers

Number of Tokens

Period #

Andreon - Croson
Table 1
Average investment in public good per subject

<table>
<thead>
<tr>
<th></th>
<th>Round</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>All</td>
</tr>
<tr>
<td>Restart Partners</td>
<td>11.54</td>
<td>11.33</td>
<td>10.29</td>
<td>7.88</td>
<td>7.33</td>
<td>6.88</td>
<td>4.21</td>
<td>6.50</td>
<td>4.25</td>
<td>2.67</td>
<td>7.28</td>
</tr>
<tr>
<td>Difference</td>
<td>2.42</td>
<td>1.50</td>
<td>1.13</td>
<td>4.46</td>
<td>5.00</td>
<td>5.00</td>
<td>5.71 *</td>
<td>1.29</td>
<td>4.79</td>
<td>1.88</td>
<td>3.32 *</td>
</tr>
<tr>
<td>Strangers</td>
<td>10.38</td>
<td>10.13</td>
<td>8.71</td>
<td>7.83</td>
<td>6.46</td>
<td>7.92</td>
<td>6.21</td>
<td>4.79</td>
<td>5.42</td>
<td>2.58</td>
<td>7.04</td>
</tr>
<tr>
<td>Restart Strangers</td>
<td>6.21</td>
<td>5.38</td>
<td>5.67</td>
<td>4.38</td>
<td>3.54</td>
<td>4.67</td>
<td>2.21</td>
<td>1.79</td>
<td>1.46</td>
<td>1.79</td>
<td>3.71</td>
</tr>
<tr>
<td>Difference</td>
<td>4.17</td>
<td>4.75</td>
<td>3.04</td>
<td>3.46</td>
<td>2.92</td>
<td>3.25</td>
<td>4.00</td>
<td>3.00</td>
<td>3.96</td>
<td>0.79</td>
<td>3.33 *</td>
</tr>
</tbody>
</table>

*different at the 5% level (two-sided Wilcoxon test)
Table 2
Average investment in public good per subject

<table>
<thead>
<tr>
<th>Round</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>All</th>
<th>Games</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strangers</td>
<td>10.38</td>
<td>10.13</td>
<td>8.71</td>
<td>7.83</td>
<td>6.46</td>
<td>7.92</td>
<td>6.21</td>
<td>4.79</td>
<td>5.42</td>
<td>2.58</td>
<td>7.04</td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>3.58</td>
<td>2.71</td>
<td>2.71</td>
<td>4.50</td>
<td>5.88</td>
<td>3.96</td>
<td>3.71</td>
<td>3.00</td>
<td>3.63</td>
<td>1.96</td>
<td>3.56 *</td>
<td></td>
</tr>
<tr>
<td>Restart Partners</td>
<td>11.54</td>
<td>11.33</td>
<td>10.29</td>
<td>7.88</td>
<td>7.33</td>
<td>6.88</td>
<td>4.21</td>
<td>6.50</td>
<td>4.25</td>
<td>2.67</td>
<td>7.28</td>
<td></td>
</tr>
<tr>
<td>Restart Strangers</td>
<td>6.21</td>
<td>5.38</td>
<td>5.67</td>
<td>4.38</td>
<td>3.54</td>
<td>4.67</td>
<td>2.21</td>
<td>1.79</td>
<td>1.46</td>
<td>1.79</td>
<td>3.71</td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>5.33</td>
<td>5.96</td>
<td>4.63</td>
<td>3.50</td>
<td>3.79</td>
<td>2.21</td>
<td>2.00</td>
<td>4.71</td>
<td>2.79</td>
<td>0.88</td>
<td>3.58 *</td>
<td></td>
</tr>
</tbody>
</table>

*different at the 1% level (two-sided Wilcoxon test)
Table 3
Proportion of subjects free riding

<table>
<thead>
<tr>
<th>Round</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>All</th>
<th>Both Games</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partners</td>
<td>0.08</td>
<td>0.13</td>
<td>0.21</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.25</td>
<td>0.42</td>
<td>0.42</td>
<td>0.54</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>Strangers</td>
<td>0.42</td>
<td>0.38</td>
<td>0.38</td>
<td>0.42</td>
<td>0.50</td>
<td>0.46</td>
<td>0.46</td>
<td>0.54</td>
<td>0.54</td>
<td>0.75</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>-0.33 **</td>
<td>-0.25 *</td>
<td>-0.17</td>
<td>-0.25 *</td>
<td>-0.33 **</td>
<td>-0.29 **</td>
<td>-0.21</td>
<td>-0.13</td>
<td>-0.13</td>
<td>-0.21</td>
<td>-0.23 *</td>
<td></td>
</tr>
<tr>
<td>Restart Partners</td>
<td>0.13</td>
<td>0.21</td>
<td>0.13</td>
<td>0.21</td>
<td>0.33</td>
<td>0.21</td>
<td>0.42</td>
<td>0.29</td>
<td>0.42</td>
<td>0.63</td>
<td>0.30</td>
<td></td>
</tr>
<tr>
<td>Restart Strangers</td>
<td>0.50</td>
<td>0.50</td>
<td>0.46</td>
<td>0.58</td>
<td>0.54</td>
<td>0.58</td>
<td>0.63</td>
<td>0.63</td>
<td>0.54</td>
<td>0.71</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>-0.38 **</td>
<td>-0.29 *</td>
<td>-0.33 **</td>
<td>-0.38 **</td>
<td>-0.21</td>
<td>-0.38 **</td>
<td>-0.21</td>
<td>-0.33 **</td>
<td>-0.13</td>
<td>-0.08</td>
<td>-0.27 *</td>
<td></td>
</tr>
</tbody>
</table>

*different at the 5% level (t-test)
**different at the 1% level (t-test)
Table 4
Standard Deviation of Investment in Public Good

<table>
<thead>
<tr>
<th></th>
<th>Round</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Strangers</td>
<td>10.12</td>
<td>10.08</td>
<td>9.25</td>
<td>9.26</td>
<td>8.66</td>
<td>9.54</td>
<td>7.76</td>
<td>7.06</td>
<td>7.93</td>
<td>6.45</td>
</tr>
<tr>
<td>Difference</td>
<td>-0.49</td>
<td>0.33</td>
<td>0.78</td>
<td>0.52</td>
<td>1.12</td>
<td>-0.10</td>
<td>0.40</td>
<td>2.76</td>
<td>2.48</td>
<td>0.88</td>
</tr>
<tr>
<td>Restart Strangers</td>
<td>8.37</td>
<td>7.88</td>
<td>7.18</td>
<td>7.28</td>
<td>6.22</td>
<td>7.75</td>
<td>3.24</td>
<td>2.78</td>
<td>2.13</td>
<td>4.31</td>
</tr>
<tr>
<td>Difference</td>
<td>0.97</td>
<td>1.52</td>
<td>1.05</td>
<td>0.80</td>
<td>1.75</td>
<td>-0.76</td>
<td>3.15</td>
<td>5.11</td>
<td>4.00</td>
<td>1.55</td>
</tr>
</tbody>
</table>
### Table 5
Standard Deviation of Investment in Public Good and F-Tests

<table>
<thead>
<tr>
<th></th>
<th>Original Game</th>
<th>Restart Game</th>
<th>Both Games</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partners</td>
<td>5.95</td>
<td>5.53</td>
<td>5.95</td>
</tr>
<tr>
<td>Strangers</td>
<td>3.09</td>
<td>1.95</td>
<td>3.06</td>
</tr>
<tr>
<td>F-statistic for</td>
<td>3.71 =F(59,19)</td>
<td>8.07 =F(59,19)</td>
<td>3.8 =F(119,39)</td>
</tr>
<tr>
<td>significance level</td>
<td>0.001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
Table 6
Experiments and Treatments

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Treatment</th>
<th>Subjects per Group</th>
<th>Number of Groups</th>
<th>Rounds 1-10 original game</th>
<th>Rounds 1-3 restart game</th>
<th>Rounds 4-10 restart game</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andreoni (1988)</td>
<td>Partners</td>
<td>5</td>
<td>3</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Andreoni (1988)</td>
<td>Partners</td>
<td>5</td>
<td>3</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Croson (1995)</td>
<td>Partners</td>
<td>4</td>
<td>3</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Croson (1995)</td>
<td>Partners</td>
<td>4</td>
<td>3</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Andreoni (1988)</td>
<td>Strangers</td>
<td>5</td>
<td>4</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andreoni (1988)</td>
<td>Strangers</td>
<td>5</td>
<td>4</td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Croson (1995)</td>
<td>Strangers</td>
<td>4</td>
<td>3</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Croson (1995)</td>
<td>Strangers</td>
<td>4</td>
<td>3</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>