Random Matching with Observable Past Actions: Experimental Tests of the Reputation Effect

過去行為可觀察下隨機配對對實驗對象合作行為的影響：聲譽效果的實驗驗證

doi:10.6277/ter.2003.314.5

經濟論文叢刊, 31(4), 2003

Taiwan Economic Review, 31(4), 2003

作者/Author： 徐麗振(Li-Chen Hsu)

頁數/Page： 507-529

出版日期/Publication Date：2003/12

Random Matching with Observable Past Actions: Experimental Tests of the Reputation Effect

Li-Chen Hsu*

We compare the difference between giving by Partners and giving by Strangers in experiments with different dimensions of histories of past actions and different lengths of trials. In experiments with twenty trials, our experimental evidence shows that Partners contribute more than Strangers in most rounds when the histories of group contributions are not revealed to Strangers. The difference becomes smaller if this information is made available to Strangers, and eventually disappears if both Partners and Strangers are further informed about the histories of individual contributions. However, in experiments with two sequences of ten trials, Partners still contribute more than Strangers even if the histories of individual contributions are known. These observations suggest that more symmetric and detailed information regarding past actions reduces the difference between giving by Partners and giving by Strangers, but sufficient repetitions are necessary for this difference to be markedly reduced.

Keywords: Public goods experiments, Reputation effect, Strategies hypothesis
JEL classification: C91, H41, C71

1 Introduction

It is well known that the dominant-strategy Nash equilibrium of the single-shot prisoners' dilemma game is defection. By applying backward induction, the same

*Department of Public Finance, National Chengchi University. I would like to thank two anonymous referees for very valuable comments and suggestions. I have also benefited from participants at the 2002 Taiwan Economic Association Annual Meeting and the 2003 Western Economic Association International Pacific Rim Conference. Financial support from the National Science Council (grant number: NSC 89-2415-H-004-049) is gratefully acknowledged.

經濟論文叢刊 (*Taiwan Economic Review*), 31:4 (2003), 507–529。
國立台灣大學經濟學系出版

equilibrium holds in every repetition of the finitely-repeated game. However, the experimental evidence often indicates the existence of substantial cooperation, thereby contradicting the theoretical prediction. A similar anomaly occurs in the public-good provision game. The unique dominant-strategy Nash equilibrium of a linear public-good provision game (which is an N-person version of the prisoners' dilemma game) is complete free-riding. However, experimental studies based on the single-shot game have generally found that subjects contribute about half of their endowments to the public good. In the repeated game, contributions typically declined and reached their lowest level in the final round, but the dominant-strategy Nash equilibrium seldom showed up.[1]

The above observations suggest that subjects need to learn the free-riding incentive through repetitions of the game. In a pioneering paper, Andreoni (1988) raised the strategies hypothesis and the learning hypothesis to explain the above observations and tested them in public goods experiments. The strategies hypothesis was derived from Kreps, *et al.*'s (1982) idea of reputation building. In a breathtaking paper, Kreps, et al. noted that incomplete information could provide players with an incentive to cooperate if they played the prisoners' dilemma game repeatedly. Players would try to build up a reputation by playing strategically for a period of time. During the process of reputation-building, they revealed their intentions to cooperate with their partners and in the meantime educated their partners to behave similarly. As a consequence, cooperation could appear most of the time in the finitely-repeated prisoners' dilemma game. Subsequent studies have called the effect of reputation-building on cooperation the reputation effect.

What Kreps, *et al.* (1982) suggested is that cooperation should be higher in finitely-repeated dilemma games than in single-shot counterparts, because reputation building or strategic play is unlikely to occur if players meet only once. In Andreoni's experiments, subjects in groups of five were divided into two conditions, the Partners condition and the Strangers condition, and they played the public-good game for ten rounds. Group components did not change across rounds in the Partners condition, while in the Strangers condition subjects were randomly rematched at the start of each new round. Under this setting, subjects in the Partners condition actually played the finitely-repeated game and subjects in the Strangers condition played the single-shot game with ten independent iterations. If the strategies hypothesis or reputation building is sufficient to explain

[1] See for instance, Marwell and Ames (1981), Issac, Walker, and Thomas (1984), Issac and Walker (1988), and the surveys by Davis and Holt (1993) and Ledyard (1995).

the decay in the repeated public-good game, then we should find more contributions by Partners than by Strangers. However, Andreoni found that contributions were instead higher in the Strangers condition than in the Partners condition, thus contradicting the strategies hypothesis.[2]

Andreoni (1988) tested the learning hypothesis by comparing contributions by Partners with those by Strangers in the final round and by restarting the game unexpectedly after subjects had played ten rounds. Since the incentives are equivalent for both Partners and Strangers in the final round and they have the same opportunity to learn, they should contribute the same levels at the end of the game. Furthermore, since the first round of the restart game is simply a continuation of the original game, if learning is the primary reason for the decay, then restarting should not affect both Partners and Strangers. Andreoni found that Strangers gave more and completely free rode less than Partners at the end of the game, suggesting that learning alone was not responsible for the decay. Moreover, Strangers were only temporarily affected by the restart, whereas Partners' contributions in round 11 returned to about the same level in round 1 and decreased only slightly after three rounds. This suggests that the strategic incentive is successfully subtracted from the Strangers treatment, but not from the Partners treatment. Again, learning alone cannot successfully explain these results.

Andreoni's work has given rise to many subsequent studies that have examined the reputation effect by using the Partners versus Strangers framework. Table 1 summarizes some related examples.[3] Among these, Palfrey and Prisbrey (1996) also found that Strangers gave more than Partners, which went against the reputation effect. Croson (1996), on the other hand, replicated Andreoni's experiments, but reported evidence supporting the reputation effect. In a way similar to Andreoni, however, she also restarted the game unexpectedly and found that the restart effect was significant in the Partners treatment and present but insignificant in the Strangers treatment. Other studies supporting the reputation effect included Sonnemans, Schram, and Offerman (1999) and Keser and van Winden (2000). Some other studies, e.g., Weimann (1994) and Brandts and Schram (2001), found no difference between contributions by Partners and contributions by Strangers. By employing subjects from different countries, Burlando and Hey (1997) found from Italian data that Partners free rode less than Strangers, while from UK data and the aggregate data they found that free riding

[2] As was mentioned previously, Andreoni's (1988) strategies hypothesis is derived from Kreps, *et al.*'s (1982) idea of reputation building. The strategies effect and reputation effect are interchangeable in the Partners versus Strangers literature.

[3] One may refer to an excellent survey by Andreoni and Croson (1998) for more details.

by Partners and Strangers did not differ significantly. In addition, they found a restart effect in seven out of ten sessions, the effect being rather profound in UK sessions.[4] Brandts, Saijo, and Schram (2002), using subjects from Japan, the Netherlands, Spain, and the US, tested the reputation hypothesis by using the GCI (Gross Cooperation Index) as the dependent variable, and countries, treatments (Partners/Strangers), age, the economics field, and gender as independent variables. They found no difference between Partners and Strangers and no effect of the country variables on the GCI.

These different conclusions are not surprising, since the experimental designs in the above-mentioned studies generally vary to some extent. However, these studies indeed have something in common. First, both Partners and Strangers played the public-good game for several rounds, and they were informed about the round-payoff at the end of each round. Under this design, subjects could figure out other group members' behavior in this round from their round-payoffs. Since Partners stayed in the same group until the end of the game, group contributions in all preceding rounds turned out to be public information. By contrast, Strangers were re-grouped, so they had no clue about how other group members that they met during this round behaved previously unless they were provided with this information explicitly. Therefore, besides the strategic incentive, information asymmetry may also have led to differences between contributions by Partners and contributions by Strangers.

Second, as can be seen from Table 1, all of these studies adopted four-person to six-person groups, which differed from Kreps, *et al.*'s (1982) two-person groups. Any single subject in groups of more than two members actually played the public-good game with a 'bundle' of other group members, and of course he (she) himself (herself) was involved in a 'bundle' when he (she) faced other subjects. Therefore, when one subject played strategically, trying to educate others to behave in the same way or to reveal that he (she) had been educated, this strategic behavior may have been averaged out by other group members' non-cooperative behavior. This situation became more serious if subjects could only observe the aggregate behavior of other group members but not individual behavior. As a consequence, Partners may not have contributed any more than Strangers, even if it was believed that only Partners could play strategically, but Strangers could not.

[4]Burlando and Hey's restart mechanism differs from those in Andreoni (1988) and Croson (1996) in two respects. First, they informed subjects about the restart before the original game began, and second, subjects who were Partners (Strangers) in the original game might have become Strangers (Partners) in the restart game.

Table 1: Studies on Partners versus Strangers

Study	Rounds	Group size	MPCR	History is provided?	Restart	Who contribute more?	Variation
Andreoni (1988)	10	5	0.5	no	yes (3 rounds)	Strangers	—
Brandts and Schram (2001)	10	4	various across rounds	no	no	Neither	—
Brandts, Saijo, and Schram (2002)	10[a]	4	various across situations	no	no	Neither	—
Burlando and Hey (1997)	8	6	various across rounds	no	yes (8 rounds)	UK: Neither Italian: Partners Aggregate: Neither	Strangers higher
Croson (1996a)	10	4	0.5	no	yes (10 rounds)	Partners	Partners higher
Keser and van Winden (2000)	25	4	0.5	Yes (group histories)	no	Partners	Partners higher
Palfrey and Prisbrey (1996)	10[b]	4	various	no	no[c]	Strangers	Strangers higher
Sonnemans, Schram, and Offerman (1999)	36[d]	4	0.667	no	no	Partners	—
Weimann (1994)	10	5	0.5	Yes (group and individual histories)	no	Group history: Neither Individual history: Neither	Strangers higher
This study	20, 10	5	0.5	Yes (group and individual histories)	Yes (only in the experiments of ten rounds)	20 rounds: No history: Partners Group history: Partners (in early stages of the game) Individual history: Neither *10 rounds:* Individual history: Partners	*20 rounds:* Neither *10 rounds:* Partners higher

[a] Each round involves 10 situations, with each situation having a different MRS.

[b] Four ten-round sequences were involved in the experiments.

[c] Subjects knew from the beginning that they would play four sequences.

[d] Group composition remained constant for a minimum of three and a maximum of twelve periods. Subjects who left the group would never come back.

The last common feature is that, apart from Sonnemans, Schram, and Offerman (1999) and Keser and van Winden (2000), most experiments lasted only ten

or less iterations (before the restart if there was one). If repetitions are necessary for subjects to play strategically during the game and to approach the free riding behavior when the final round eventually arrives, then an interesting question that arises here is whether ten or less trials are sufficient for subjects to learn or to react.

We would like to concentrate on the strategies hypothesis and shall only discuss subjects' learning behavior briefly in this paper. Due to the common features of these preceding studies, we examine the difference between Partners and Strangers by providing subjects with different dimensions of histories of past actions and by varying the duration of experiments. These variations help us to determine whether more symmetric or detailed information or both reduces the difference between contributions by Partners and contributions by Strangers and whether longer trials are necessary for this difference to disappear.

In experiments with twenty trials, our experimental evidence shows that Partners contributed more than Strangers in most rounds when the histories of group contributions were not revealed to Strangers. The difference became smaller as the histories of group contributions were made available to Strangers, and eventually disappeared when both Partners and Strangers were further provided with the histories of individual contributions. These observations are complementary to the findings that suggested that providing information on individual histories had no effect on Partners' behavior, but raised Strangers' contributions. Under the condition that the histories of individual contributions were public information, we ran additional experiments with two sequences of ten trials. Here we observed that Partners still contributed more than Strangers. These findings suggest that more symmetric and detailed information reduces the difference between contributions by Partners and contributions by Strangers, but that sufficient repetitions are necessary for the difference to disappear.

The rest of the paper is organized as follows. Section 2 provides the experimental design. The experimental results are presented in Section 3. Section 4 concludes.

2 Experimental Design

We employ a linear public-good design in our experiments. This design was first introduced by Marwell and Ames (1981) and has been adopted by many subsequent researchers. Subjects played the game for twenty rounds or two sequences of ten rounds, depending on which experiments they participated in. They were each endowed with 200 points per round and were instructed to allo-

cate this endowment between the Public Account (public good) and the Private Account (private good). Denote x_i as the amount of points that subject i allocates to the Private Account and $g_i (= 200 - x_i)$ as the amount of points that he (she) allocates to the Public Account. The subject's payoff per round is written as $P_i = x_i + \alpha \sum_{j=1}^{n} g_j$, where α is the marginal per capita return (MPCR) and n is the group size. We used $\alpha = 0.5$ and $n = 5$ in all experiments. Therefore, one point allocated to the Private Account earned one point for oneself, while one point allocated to the Public Account earned each group member 0.5 points. By summing the subject's payoff earned in each round we are able to obtain his (her) total payoff. Ten points can be transformed into NT$1 at the end of the experiment.

Since the marginal private return from the Private Account is greater than the marginal private return from the Public Account, the dominant-strategy Nash equilibrium in a single-shot game is $g_i = 0$ for all i. However, if everyone allocates all points to the Public Account, that will be a Pareto efficient outcome, because one point allocated to the Public Account yields a marginal group return of 2.5 points, which is greater than the marginal private return from the Private Account.

Table 2 summarizes seven experiments conducted in this study and reports the average earning for each experiment.[5] Twenty subjects were used in each of the seven experiments, which resulted in a total of 140 subjects participating in this research. One hundred subjects were undergraduate students at National Chengchi University and the other 40 subjects were undergraduate students at National Taiwan University. All subjects were inexperienced and the mean age of the 140 subjects was 20.29 years. Among them, 51 were males and 89 were females, and 136 subjects majored in economics or economics-related fields.

These seven experiments are divided into the Partners condition and the Strangers condition. Under the Partners condition, the subjects were randomly assigned to groups of five and played the public-good game with the same group members in all iterations. Under the Strangers condition, the subjects were also randomly assigned to groups of five, but were randomly rematched at the beginning of each new round.[6] We employed two different lengths of trials (twenty

[5]Each experiment lasted about one hour and fifty minutes, and on that basis average hourly earnings per subject were about NT$315, more than three times the part-time hourly wage rate for an undergraduate student in Taiwan. The exchange rate between the NT (New Taiwan) dollar and the US dollar was about 35:1 when these experiments were conducted. Data from the experiments are available from the author upon request.

[6]Each subject was assigned a subject number between 1 and 20. Random assignments were

Table 2: The Structure and Average Earnings of the Experiments

Condition	Experiment	Rounds	Information Provided	Average Earnings
Partners	PartnerGH20	20	group histories	NT$628
	PartnerIH20	20	group and individual histories	NT$605
	PartnerIH10	10+10	group and individual histories	NT$653
Strangers	Stranger20	20	no histories	NT$531
	StrangerGH20	20	group histories	NT$544
	StrangerIH20	20	group and individual histories	NT$604
	StrangerIH10	10+10	group and individual histories	NT$512

and ten rounds) and provided subjects with different information spaces regarding past actions.

Three experiments were conducted under the Partners condition. In two of them (PartnerGH20 and PartnerIH20) subjects were instructed to play the public-good game for twenty rounds. In PartnerGH20, subjects were aware of the histories of group contributions, while in PartnerIH20 subjects were aware of each of the other four group members' histories of individual contributions. The final experiment was PartnerIH10, in which the information space was the same as that of PartnerIH20, but subjects first played only ten rounds (the original game) and then they were asked to play the same game for another ten rounds after an unexpected restart (the restart game).[7] They remained in the same group after the restart.

The information regarding group history is actually redundant for Partners, since they can always figure it out from their own past payoffs. However, this information is not necessarily available for Strangers because they are rematched beginning with each new round. As a result, we conducted four corresponding experiments under the Strangers condition.

The first two experiments were Stranger20 and StrangerGH20. In Stranger20, neither group nor individual histories of past contributions were available. In StrangerGH20, subjects were provided with the information regarding current group members' total contributions in each of the preceding rounds. Notice

made by the computer program.

[7] Before we restarted the experiments, we made sure that all subjects had enough spare time to complete another ten rounds.

that PartnerGH20 and Stranger20 are analogous to the experiments used in Andreoni (1988) and Croson (1996), apart from the fact that they instructed subjects to play only ten rounds before the restart. Since group history is always available under the Partners condition, the information spaces of PartnerGH20 and Stranger20 are actually not equivalent. Therefore, we use PartnerGH20, Stranger20, and StrangerGH20 to serve as the basic treatments.

The third experiment under the Strangers condition was StrangerIH20, which corresponded to PartnerIH20. Subjects in StrangerIH20 were aware of each of the other four group members' histories of individual contributions and of course the history of group contributions.[8] The last experiment was StrangerIH10, which corresponded to PartnerIH10 and had the same information space as that of StrangerIH20. In Stranger20, StrangerGH20, and StrangerIH20, subjects played the public-good game for twenty rounds. In StrangerIH10, subjects first played only ten rounds and then another ten rounds after an unexpected restart. In the restart game of StrangerIH10, the individual histories generated in the original game were continuously provided to subjects.

Subjects' Instructions, Earnings Reports, and Decision Forms were handed out to each subject after he or she was seated.[9] Aside from the information illustrated in Table 2, total and average contributions made by the other four group members and all five group members in the current round were public information in all experiments.

3 Results of Experiments

Tables 3 through 5 report the experimental evidence. As can be seen from Table 3, several results observed here are quite consistent with related studies. First, contributions started high in all experiments. Partners began with an average contribution of more than 40% of their endowments. Though Strangers started with a lower average contribution, they still contributed more than 35% of their endowments. Second, contributions generally declined over time and reached

[8]Take subject #1 for instance. Subject #1 was randomly assigned with subjects #9, #13, #15, and #18 to the same group in round 10 of all Strangers treatments. Before the tenth round started, subject #1 in both StrangerIH20 and StrangerIH10 was provided with each of the other four subjects' contributions to the public good in each round of rounds 1 through 9. However, subject #1 in StrangerGH20 was only aware of the sum of the contributions made by subjects #9, #13, #15, and #18 in each round of rounds 1 through 9. Subjects in Stranger20 were unaware of any of the information above.

[9]Subjects' Instructions, Earnings Reports, and Decision Forms are available from the author upon request.

Table 3: The Percentage of Endowments Contributing to the Public Good

	Experiment						
Round	PartnerGH20	PartnerIH20	PartnerIH10	Stranger20	StrangerGH20	StrangerIH20	StrangerIH10
1	59.38	43.38	43.75	36.63	35.08	39.38	35.50
2	64.25	42.75	47.38	37.50	36.28	36.13	31.78
3	54.43	39.63	53.88	35.13	34.75	40.50	24.93
4	51.38	38.30	50.35	29.00	30.40	43.13	18.55
5	52.13	40.60	45.68	25.75	30.43	35.25	16.38
6	44.75	40.38	45.00	28.00	25.98	33.75	17.73
7	44.50	37.63	45.03	20.75	25.55	34.75	16.50
8	45.50	37.00	40.23	25.63	22.45	37.00	13.90
9	46.63	35.68	39.18	24.88	20.23	38.43	11.75
10	43.70	35.15	28.75	15.95	17.13	38.25	9.88
11	32.63	38.13	40.63	19.25	24.43	40.88	23.89
12	34.95	34.58	44.15	22.88	26.45	34.05	21.25
13	24.00	33.80	43.70	17.88	21.35	34.38	21.13
14	23.63	29.03	43.13	14.78	19.63	32.25	24.53
15	28.13	29.90	43.38	22.65	21.35	21.00	18.63
16	29.65	29.05	43.65	14.53	21.15	28.88	16.28
17	25.50	26.43	41.68	11.15	19.58	25.00	16.00
18	24.38	26.25	36.70	11.65	19.33	27.50	13.75
19	21.75	26.00	36.20	12.15	16.53	32.13	13.38
20	9.38	19.00	30.25	11.28	11.98	28.13	8.50
Mean	37.55	34.21	42.13	21.24	23.11	33.81	18.71

Note: Each subject's endowment was 200 points per round. Therefore, multiplying the percentage of endowments contributing to the public good by 200 results in the average contribution.

the lowest level in the final round of most experiments. Finally, though contributions declined over time, the single-shot-game Nash equilibrium, that is, zero contributions, never showed up in any of the experiments. The lowest rate of contributions was 8.5% in round 20 of StrangerIH10, which was far above zero.

The main findings from specifically testing the strategies hypothesis are summarized in a series of six observations. The results of first looking at the effect of the information regarding individual past contributions on subjects' behavior, and then comparing the contributions by Partners and Strangers in various treatments are summarized in Observations 1 through 5. The preliminary implications of these observations are discussed altogether after Observation 5. We then discuss more deeply the results from the four experiments using the information regarding individual past contributions (that is, PartnterIH20, PartnerIH10, StrangerIH20, and StrangerIH10) and summarize the findings in Observation 6. More discussions on the implications regarding all of these six observations and subjects' behavior are placed after Observation 6.

Observation 1. *The information on individual past contributions had no effect on Partners' behavior, but raised Strangers' contributions.*

From looking at Table 3, it can be seen that contributions in PartnerGH20 were higher than those in PartnerIH20 in the first ten rounds. After round 10, the contributions in these two experiments were quite close. We used a one-sided Mann-Whitney rank-sum test to examine such differences on a round-by-round basis. The results showed that contributions in PartnerGH20 were significantly higher than those in PartnerIH20 in only rounds 1, 2, 3, and 4.[10] Contributions in these two experiments did not differ significantly in the remaining rounds. As for the Strangers treatments, contributions per round in Stranger20 and StrangerGH20 seemed to differ only slightly, whereas contributions in StrangerIH20 were higher than those in Stranger20 and StrangerGH20 in almost every round. A one-sided Mann-Whitney rank-sum test showed that contributions in StrangerIH20 were significantly higher than those in StrangerGH20 in 8 out of 20 rounds,[11] and were significantly higher than those in Stranger20 in

[10] We use the significance level of 5% in all of the statistical tests. That is, in a two-sided Mann-Whitney rank-sum test, if $p \leq 0.05$, then we accept the alternative hypothesis that there is a difference between two populations. In a one-sided Mann-Whitney rank-sum test, if $p \leq 0.05$, then we accept the alternative hypothesis that one population is larger (or smaller) than the other population. The statistical results are available from the author upon request.

[11] These were rounds 4, 8, 9, 10, 11, 13, 14, and 19.

14 out of 20 rounds.[12]

Observation 2. *Contributions were significantly higher in PartnerGH20 than in Stranger20 in all twenty rounds except rounds 13, 14, 15 and 20.*

Observation 3. *Contributions were significantly higher in PartnerGH20 than in StrangerGH20 in all twenty rounds except rounds 11-20.*

As can be seen from Table 3, contributions in PartnerGH20 started much higher than those in Stranger20 and StrangerGH20. The differences declined over time, however, especially after round 10. A one-sided Mann-Whitney rank-sum test showed that contributions were significantly higher in PartnerGH20 than in Stranger20 in all twenty rounds except rounds 13–15 and round 20. A similar test showed that contributions in PartnerGH20 were significantly higher than contributions in StrangerGH20 in only the first ten rounds.

Observation 4. *There existed no significant difference between contributions in PartnerIH20 and contributions in StrangerIH20 in all twenty rounds.*

Looking again at Table 3, contributions started slightly higher in PartnerIH20 (43.4% of the endowment) than in StrangerIH20 (39.4% of the endowment). Though contributions in StrangerIH20 fluctuated over a wider range than those in PartnerIH20, the overall contributions of these two experiments were in fact quite close. A two-sided Mann-Whitney rank-sum test showed that contributions in PartnerIH20 and StrangerIH20 were the same in every round of the experiments.

Observation 5. *Contributions were significantly higher in PartnerIH10 than in StrangerIH10 in all rounds except rounds 1 and 10 in the original game and rounds 14, 18, 19, and 20 in the restart game.*

Table 3 shows that the contribution rates in these two experiments started quite close to each other, with that in PartnerIH10 being 8% higher than that in StrangerIH10. After that, the difference grew in the early stage of the original game and reached its highest level in the fourth round, where the contribution rate in PartnerIH10 was about 32% higher than that in StrangerIH10. This huge difference (though lower than that in round 4) was maintained until the start of the final round, and continuously appeared in the restart game. We also observed that contribution rates seemed to decline in both the original and restart

[12] These were rounds 4, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, and 20. Contributions in Stranger20 and StrangerGH20 did not differ significantly in all 20 rounds.

games of StrangerIH10, but the decline in the restart game of PartnerIH10 was less obvious. A one-sided Mann-Whitney rank-sum test showed that contributions were significantly higher in PartnerIH10 than in StrangerIH10 in all twenty rounds except rounds 1, 10, 14, and 18–20.

We now explore the implications of Observations 1 through 5. The strategies hypothesis implies that contributions by Partners will exceed contributions by Strangers, especially in the early stages of the game. As has been shown, the strategies hypothesis held between PartnerGH20 and Stranger20, where the histories of group contributions were made available to Partners only. The difference between contributions by Partners and contributions by Strangers declined when the histories of group contributions were also revealed to Strangers, and disappeared when more detailed information, namely, the histories of individual contributions, was made available to both Partners and Strangers. As was also shown in Observation 1, the information regarding individual past contributions also raised Strangers' contributions, but had no effect on Partners' contributions. Therefore, the differences between contributions by Partners and contributions by Strangers may also have been attributable to asymmetric information, and cannot be explained by the strategic incentive alone.

Nonetheless, the information regarding individual past contributions raised Strangers' contributions only when the time frame was long enough for them to adjust their behavior. If subjects were informed at the beginning of the games that they were about to play only ten rounds, then contributions by Partners still exceeded contributions by Strangers. As was shown in Observation 5, contributions in PartnerIH10 were significantly higher than contributions in StrangerIH10 in all rounds of the original game except in the first and the final rounds. Though Strangers had observed plenty of individual past contributions when the restart game began, Partners still contributed more than Strangers in the later rounds. This evidence suggests that the information regarding individual past contributions helped Strangers' contributions, but that sufficient repetitions were necessary for them to raise the contributions.

Observation 6. *When individual past contributions were observable, Strangers in the experiment involving twenty iterations were more cooperative than Strangers in the experiment involving ten iterations, whereas Partners were equally cooperative in both kinds of experiments.*

Observation 6 is a complementary result to Observations 4 and 5. Observations 4 and 5 indicate that the difference (if any) between contributions by Strangers and contributions by Partners depends crucially upon the duration

of the experiments. Therefore, it may be helpful to look at how Partners and Strangers behaved in relation to different lengths of experiments. We can see from Table 3 that, apart from round 10, contributions per round were higher in PartnerIH10 than in PartnerIH20, with the difference being moderate before round 10. By contrast, though contributions started close together in StrangerIH20 and StrangerIH10, the difference between these two grew during the first ten rounds. Even though after round 10 the contributions in these two experiments became quite close in round 15, the difference grew again after that.

The results of a two-sided Mann-Whitney rank-sum test confirmed the observations above. The test statistics showed that contributions in each round in PartnerIH20 did not differ significantly from those in PartnerIH10. By contrast, contributions in StrangerIH20 were significantly higher than contributions in StrangerIH10 in fourteen out of twenty rounds.[13]

An interesting question is why Strangers behaved so differently from Partners when the time frame changed from 20 rounds to 10 rounds. Before answering this question, we shall look at the standard deviations of individual contributions for various treatments, with particular attention being paid to the experiments that provided the histories of individual past contributions. As can be seen from Table 4, the standard deviations of individual contributions were higher in PartnerGH20 than in Stranger20 and StrangerGH20 in most rounds, with the differences generally being moderate. An F-test showed that the variances of individual contributions in PartnerGH20 differed significantly from those in Stranger20 in eight out of twenty rounds, and that the variances of individual contributions in PartnerGH20 differed significantly from those of StrangerGH20 in only two out of twenty rounds.

The variances of individual contributions were higher (but insignificant) in PartnerIH20 than in StrangerIH20 in only round 15. The opposite was the case in the other nineteen rounds, with significant differences existing in only five rounds. The most notable difference was that observed between PartnerIH10 and StrangerIH10. PartnerIH10 seemed to have the highest standard deviations of individual contributions, while StrangerIH10 had the lowest. An F-test showed that the variances of individual contributions were significantly higher in PartnerIH10 than in StrangerIH10 in eighteen out of twenty rounds. These observations suggest that contributions by Partners and Strangers seemed to vary according to a similar pattern in longer relationships, while in shorter ones contributions by Partners varied more severely than contributions by Strangers.

[13] These fourteen rounds were rounds 3, 4, 5, 6, 7, 8, 9, 10, 11, 13, 16, 18, 19, and 20.

Table 4: The Standard Deviations of Individual Contributions

	Experiment						
Round	PartnerGH20	PartnerIH20	PartnerIH10	Stranger20	StrangerGH20	StrangerIH20	StrangerIH10
1	39.3	33.9	55.0	58.8	45.5	51.2	30.7
2	34.3	31.6	54.6	48.7	43.8	55.1	33.3
3	45.9	32.1	57.5	47.1	48.8	41.7	38.5
4	44.8	28.9	64.7	44.0	36.3	48.1	27.2
5	60.5	37.8	67.1	53.6	48.9	50.2	27.3
6	50.1	29.0	66.3	51.8	44.0	49.4	45.3
7	56.8	31.6	65.6	29.2	48.3	48.9	26.8
8	56.0	42.9	70.5	49.6	43.8	52.0	28.3
9	52.5	46.6	73.6	41.9	36.0	51.4	24.8
10	58.2	41.5	75.3	28.5	30.8	41.7	24.9
11	45.3	35.6	64.4	42.9	42.7	51.2	47.4
12	42.1	35.8	70.7	44.6	36.1	52.4	36.7
13	43.5	37.4	73.1	29.7	37.4	41.8	31.0
14	44.1	40.5	76.0	26.0	33.3	41.4	48.3
15	56.4	39.6	74.9	47.4	36.9	34.8	32.9
16	47.6	42.0	72.5	28.3	36.5	48.7	34.9
17	45.1	38.0	75.4	14.9	32.9	43.0	26.3
18	36.0	40.0	76.9	18.6	30.5	46.8	27.6
19	44.1	42.0	79.1	44.4	31.8	53.8	32.9
20	29.7	44.2	80.8	48.5	29.9	59.8	23.9
All	53.2	39.7	69.4	44.3	39.6	48.2	35.1

Table 5: The Numbers and Percentages of Observations under Various Contribution Levels

	Contribute 0		Contribute no more than 1/3 of the endowment	
Experiment	Number of observations	Percentage	Number of observations	Percentage
PatnerGH20	57	14.25%	188	47.00%
PartnerIH20	29	7.25%	170	42.50%
PartnerIH10	73	18.25%	185	46.25%
Stranger20	68	17.00%	308	77.00%
StrangerGH20	75	18.75%	275	68.75%
StrangerIH20	37	9.25%	220	55.00%
StrangerIH10	81	20.25%	325	81.25%
	Contribute all of the endowment		Contribute at least 2/3 of the endowment	
Experiment	Number of observations	Percentage	Number of observations	Percentage
PatnerGH20	19	4.75%	69	17.25%
PartnerIH20	2	0.50%	12	3.00%
PartnerIH10	67	16.75%	89	22.25%
Stranger20	10	2.50%	22	5.50%
StrangerGH20	4	1.00%	13	3.25%
StrangerIH20	21	5.25%	35	8.75%
StrangerIH10	3	0.75%	5	1.25%

The total number of observations in each experiment is 400.

One may then ask in what directions did contributions made by Partners and Strangers vary. To answer this question, we looked specifically at the individual behavior in each experiment. Table 5 reports the numbers and percentages of observations under various contribution levels. Following Issac, Walker, and Thomas (1984) and Weimann (1994), we called the subjects who gave more than two-thirds of their endowments the cooperators, the subjects who gave less than one-third of their endowments the free riders, and those who gave in between the weak free riders.

The most remarkable phenomenon regarding free riding is that on average 81.3% of the subjects in StrangerIH10 behaved like free riders, of which 20.3% gave nothing to the public good. Free-riding was less severe in the other three-Strangers experiments, yet more than one half or even two-thirds of the subjects still behaved like free riders. Notice that a very high level of complete free-riding (18.3%) also appeared in PartnerIH10, but less than one half (46.3%) of the subjects behaved like free riders on average. Likewise, on average less than one half of the subjects in the other two Partners experiments were free riders.

In taking a look at how often the subjects behaved nicely, we observe first of all that there were strikingly more cooperators in PartnerIH10 than in the other experiments: 16.8% of the subjects in PartnerIH10 contributed all of their endowments on average, while about 22.3% of the subjects behaved like cooperators. By contrast, only 3% of the subjects in PartnerIH20 were cooperators. There were also very few cooperators in the Strangers sessions: on average only 1.3% of the subjects in StrangerIH10 were cooperators and less than 10% of the subjects in each of the other three Strangers experiments behaved cooperatively.

In sum, Strangers generally free rode more than Partners, especially those in StrangerIH10, in which an astonishing 81% of the subjects behaved as free riders on average. By contrast, Partners not only behaved differently across different experiments, but they also behaved diversely within the same experiments. By looking specifically at the original data of PartnerIH10 and PartnerIH20, we found that subjects that made an average contribution of more than one half of their endowments were always not the only ones to do so in the group, reflecting a certain degree of positive reciprocity or fairness.[14] However, other group members may at the same time have taken advantage of the subjects who behaved nicely. As a result, positive reciprocity and free riding were seen to co-exist within the same group.[15] We also observed in other groups that all five members gave

[14]Rabin (1993) develops fairness equilibria which are outcomes reflecting the motivations that people like to help those who are helping them, and to hurt those who are hurting them. Fehr and Gächter (2000) adopt a similar concept of reciprocity to that of Rabin. One may refer to surveys by Güth and Tietz (1990) and Fehr and Gächter (2000) for more studies related to fairness and reciprocity. Since the group size in our experiments is greater than two, it is complicated examining the reciprocal incentive explicitly. So here we would like to focus more on the effect of the information regarding individual past contributions on cooperation, and leave the reciprocal issue to another paper.

[15]For instance, in PartnerIH20 only subjects #11, #13, and #14 made an average contribution of more than one half of their endowments and these subjects happened to be assigned to the same group. In PartnerIH10, subjects #8, #11, #13, #14, and #16 were in the same group and each of them contributed at least one half of their endowments on average. Moreover, in another group

very low levels, suggesting that some Partners played the negatively reciprocal strategy.[16]

So now we can see why Strangers behaved so differently from Partners when the time frame changed from twenty rounds to ten rounds. Recall that the most salient difference between Partners and Strangers was that Partners stayed in the same group until the end of the game, while Strangers were reshuffled at the start of each new round. Therefore, Strangers were less likely to play the reciprocal strategy, either positively or negatively. Nevertheless, this does not necessarily imply that Strangers simply free rode all the time. They left traces that were associated with the way in which they had behaved. These traces provided Strangers with an incentive to play strategically and prevented them from being harsh, especially when more iterations were to be played. Furthermore, though Strangers were randomly rematched starting each new round, they might have met again. The possibility of meeting the same people again also gave Strangers an incentive to play strategically,[17] and this possibility was certainly higher if the game lasted longer. These reasons could explain why strangers in StrangerIH10 behaved more uniformly and were less cooperative than strangers in StrangerIH20.

As in the case of Strangers, the possibility of Partners playing strategically is greater in longer relationships than in shorter ones. Yet, as was mentioned previously, a primary difference between Partners and Strangers was that Partners stayed in the same group regardless of whether the game was longer or shorter, while Strangers did not. Therefore, besides playing strategically, Partners may have reciprocated or matched others' behavior. Positive reciprocity works in the same direction as strategic play, yet negative reciprocity may cancel or outweigh these two incentives. If reciprocity (and maybe other incentives) really counts for Partners, then intuitively we may observe higher standard deviations of Part-

subjects #12 and #17 gave all of their endowments (200 points) in all of the twenty rounds, and #7 made only an average contribution of 15 points.

[16] For instance, Subject #9 in PartnerIH10 gave nothing in the first round and continued to consistently give nothing for another sixteen rounds, making an average contribution of only 8 points. The sum of the average contributions made by Subject #9's partners (Subjects #1, #2, #4, and #20) was only 169 points, while the sums of the average contributions made by the other three groups were 561, 398, and 550 points, respectively.

[17] Sonnemans, Schram, and Offerman (1999) conjectured that Strangers could play strategically as well if they had a chance to meet with other Strangers more than once, though they were unaware when exactly they might meet. The reason was that if a Stranger defected this round, then other group members would be educated by this Stranger and defected later on, and therefore he (she) would be punished by his (her) own defection once he (she) met the original group member (s) again. If Strangers recognize this, then they will play strategically and behave nicely sometimes.

ners' contributions in some cases, but in other cases the standard deviations of Strangers' contributions will be higher. Moreover, since these incentives may not work in the same direction and may not appear for everyone, there should be no ex ante predictions as to whether Partners should contribute more or less in longer relationships than in shorter ones, or whether Partners should contribute more or less than Strangers. Furthermore, the difference between Partners' contributions and Strangers' contributions, if any, cannot be explained by the incentive of reputation building or strategic play alone.

The discussion above may explain why there were no significant differences between contributions in PartnerIH20 and contributions in StrangerIH20 or between contributions in PartnerIH20 and contributions in PartnerIH10. As to why contributions in PartnerIH10 were much higher than those in StrangerIH10, we presume that this was because Strangers in shorter relationships consistently tended to free ride, and because most Partners that were drawn in PartnerIH10 were relatively more prone to positive reciprocity as compared with subjects in other experiments.

It should be noted that Weimann (1994) also provided subjects with information regarding group and individual past contributions and instructed them to play ten rounds. He found that there was no difference between contributions by Partners and contributions by Strangers regardless of whether or not subjects were informed about the histories of group contributions or the histories of individual contributions. He also found that contributions by Strangers varied more than contributions by Partners. These findings are different from those in our experiments involving ten rounds. These differences may be the result of differences in experimental designs or because of the subjects' nationalities or for other reasons. In any case, Partners' and Strangers' behavior appeared to be more complicated than we had predicted. To reach a more solid conclusion, more sessions may need to be conducted.

4 Conclusions and Discussions

Studies examining the reputation effect in public goods experiments have given rise to inconsistent conclusions. In this paper, we have tried to test the reputation effect by providing subjects with different dimensions of histories of past actions and varying the duration of experiments. In experiments with twenty trials, we found that contributions by Partners exceeded contributions by Strangers in almost all rounds when the histories of group contributions were not revealed to Strangers, thus supporting the reputation effect. However, the difference de-

clined when the histories of group contributions were also made available to Strangers, and eventually disappeared completely when the histories of individual contributions were made publicly available to both Partners and Strangers. These observations are consistent with complementary evidence that showed that providing subjects with information on individual past contributions improved Strangers' contributions, but had no effect on Partners' behavior.

Evidence from experiments with two sequences of ten trials is quite different from evidence from experiments with twenty trials. Under the condition that the history of individual contributions was public information to all players, we ran additional experiments with two sequences of ten trials and found that Partners still contributed more than Strangers. Furthermore, Strangers in the experiment involving twenty trials were more cooperative than Strangers in the experiment involving two sequences of ten trials, whereas Partners were equally cooperative in both kinds of experiments. These observations suggest that information and the length of the game may also affect subjects' (especially Strangers') behavior, and therefore the difference between contributions by Partners and contributions by Strangers cannot be explained sufficiently by the reputation effect.

Besides the information aspect, our experimental evidence also suggests that Partners might have played with the incentives of reciprocity or fairness besides strategies. On the other hand, though Strangers were randomly rematched after each round, they might have met again. The chance of meeting again may allow Strangers to play strategically. In addition, how individual subjects in groups of more than two members respond to other members' aggregate or individual decisions makes the analysis of Partners versus Strangers even more complicated. As already noted by Andreoni and Croson (1998), "the evidence on Partners versus Strangers suggests that repeated play is quite different from repeated single-shot play, but it is unlikely that much of that difference is due to the game-theoretic reputation effect."

In addition to those possible reasons discussed above, the diversities in experimental designs that we mentioned previously have certainly contributed to the inconsistent conclusions in the Partners versus Strangers literature. Another problem that has been noticed in some studies, but has not yet been discussed explicitly, is concerned with the manipulation of data.[18] Take several studies illustrated in Table 1 for example. In analyzing the difference between contributions by Partners and contributions by Strangers, Andreoni (1988) used the

[18]The author is very grateful to a referee for bringing up this problem. Keser and van Winden (2000, p. 24) euphemistically pointed out that few experimental sessions conducted in studies in the Partners versus Strangers literature may be one reason for the mixed evidence.

mean contributions of each subject as observations. Weimann (1994) used round contributions of each subject and compared the difference between Partners and Strangers on a round-by-round basis. We also used this method to deal with the data in this paper. Burlando and Hey (1997) used round data as well in addition to overall data. Furthermore, Croson (1996), Keser and van Winden (2000), and Brandts and Schram (2001) used session averages as observations in the Strangers condition and group average contributions as observations in the Partners condition.[19] Admittedly, different presentations of the data may produce different results, even if the data are generated by the same subjects.

References

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

Andreoni, J. and R. Croson (1998), "Partners versus strangers: Random rematching in public goods experiments", to appear in C. R. Plott and V. L. Smith, eds., *Handbook of Experimental Economics Results*, Elsevier Science, the Netherlands.

Brandts, J., T. Saijo and A. Schram (2002), "How universal is behavior? A four country comparison of spite, cooperation and errors in voluntary contribution mechanisms", Tinbergen Institute Discussion Paper.

Brandts, J. and A. Schram (2001), "Cooperation and noise in public goods experiments: applying the contribution function approach", *Journal of Public Economics*, 79, 399–427.

Burlando, R. and J. D. Hey (1997), "Do anglo-saxons free-ride more?", *Journal of Public Economics*, 64, 41–60.

Croson, R. (1996), "Partners and strangers revisited", *Economics Letters*, 53, 25–32.

Davis, D. D. and C. A. Holt (1993), *Experimental Economics*, Princeton University Press, Princeton, New Jersey.

Fehr, E. and S. Gächter (2000), "Fairness and retaliation: The economics of reciprocity", *Journal of Economic Perspectives*, 14(2), 159–181.

Güth, W. and R. Tietz (1990), "Ultimatum bargaining behavior: A survey and comparison of experimental results", *Journal of Economic Psychology*, 11, 417–449.

[19] However, Croson (1996) inconsistently used overall data when analyzing the variances of individual contributions.

Issac, R. M. and J. M. Walker (1988), "Group size effects in public goods provision: the voluntary contributions mechanism", *Quarterly Journal of Economics*, 103, 179–200.

Issac, R. M., J. M. Walker and S. H. Thomas (1984), "Divergent evidence on free-riding: An experimental examination of possible explanations", *Public Choice*, 43, 113–150.

Keser, C. and F. van Winden (2000), "Conditional cooperation and voluntary contributions to public goods", *Scandinavian Journal of Economics*, 102, 23–39.

Kreps, D., P. Milgrom, J. Roberts and R. Wilson (1982), "Rational cooperation in the finitely repeated prisoners' dilemma", *Journal of Economic Theory*, 27, 245–252.

Ledyard, J. O. (1995), "Public goods: a survey of experimental research", in John H. Kagel and A. E. Roth eds., *Handbook of Experimental Economics*, Princeton University Press, Princeton, New Jersey, 111–194.

Marwell, G. and R. E. Ames (1981), "Economists free ride, does anyone else"? *Journal of Public Economics*, 15, 295–310.

Palfrey, T. R. and J. E. Prisbrey (1996), "Altruism, reputation and noise in linear public goods experiments", *Journal of Public Economics*, 61, 409–427.

Rabin, M. (1993), "Incorporating fairness into game theory and economics", *American Economic Review*, 83, 1281–1302.

Sonnemans, L., A. Schram and T. Offerman (1999), "Strategic behavior in public good games: when partners drift apart", *Economics Letters*, 62, 35–41.

Weimann, J. (1994), "Individual behavior in a free riding experiment", *Journal of Public Economics*, 54, 185–200.

過去行為可觀察下隨機配對對實驗對象合作行為的影響: 聲譽效果的實驗驗證

徐麗振
國立政治大學財政學系

我們對實驗對象提供不同層面的過去行爲的歷史, 以及採用不同長度的實驗, 來比較合夥人與陌生人在公共財實驗中對公共財捐獻的差異。我們的實驗數據顯示, 在二十回合的實驗中, 當群體捐獻的歷史不對陌生人揭露時, 在大多數回合中合夥人的捐獻都比陌生人多。若群體捐獻的歷史對陌生人揭露, 則此一差異縮小, 若我們再進一步對合夥人及陌生人都揭露個人捐獻的歷史, 則此一差異終將消失。然而, 在兩個序列的十回合實驗中, 即使個人捐獻的歷史爲已知, 合夥人的捐獻仍較陌生人多。這些觀察建議, 若過去行爲的資訊在合夥人與陌生人間較對稱也較詳細, 則合夥人與陌生人間捐獻的差異將縮小, 但實驗回合數必須夠長才能使此一差異明顯下降。

關鍵詞: 公共財實驗, 聲譽效果, 策略假說
JEL 分類代號: C91, H41, C71