

Non-Excludable Public Good Experiments[†]

by

Timothy N. Cason*

Tatsuyoshi Saijo**

Takehiko Yamato***

and

Konomu Yokotani****

May 1997; revised October 2003

[†] We thank Takenori Inoki, Mamoru Kaneko, Hajime Miyazaki, Toru Mori, Mancur Olson, Mitsuo Suzuki, two anonymous referees, and Economic Science Association conference participants for their helpful comments and discussions. This research was partially supported by the Zengin Foundation for the Studies on Economics and Finance, Grant in Aid for Scientific Research 08453001 and 15310023 of the Ministry of Education, Culture, Sports, Science and Technology in Japan, and the Tokyo Center for Economic Research Grant.

* Department of Economics, Krannert School of Management, Purdue University, 100 S. Grant Street, West Lafayette, IN 47907-2076, USA. E-mail address: cason@mgmt.purdue.edu

** Institute of Social and Economic Research, Osaka University, Ibaraki, Osaka 567-0047, Japan, E-mail address: saiyo@iser.osaka-u.ac.jp; Research Institute of Economy, Trade and Industry, 1-3-1 Kasumigaseki, Chiyoda, Tokyo 100-8901, Japan; and Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125, USA, E-mail address: saiyo@hss.caltech.edu

*** Department of Value and Decision Science, Graduate School of Decision Science and Technology, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan. E-mail address: yamato@valdes.titech.ac.jp

**** Research and Survey Department, The Government Housing Loan Corporation, 1-4-10 Koraku, Bunkyo, Tokyo 112-8570, Japan. E-mail address: QZT03105@niftyserve.or.jp

Correspondent: Tatsuyoshi Saijo, Institute of Social and Economic Research, Osaka University Ibaraki, Osaka 567-0047, Japan. Phone 81-6 (country & area codes) 879-8582 (office)/878-2766 (fax). E-mail: saiyo@iser.osaka-u.ac.jp

Abstract

We conduct a two-stage game experiment with a non-excludable public good. In the first stage, two subjects choose simultaneously whether or not they commit to contributing nothing to provide a pure public good. In the second stage, knowing the other subject's commitment decision, subjects who did not to commit in the first stage choose contributions to the public good. We found no support for the evolutionary stable strategy equilibrium, and the ratio of subjects who did not commit to contributing nothing increased as periods advanced; that is, the free-riding rate declined over time. Furthermore, this behavior did not arise due to altruism or kindness among subjects, but from *spiteful* behavior of subjects.

Keywords:

Laboratory, Fairness, Spite, Social Preferences, Voluntary Contribution Mechanism, Hawk-Dove Game

1. Introduction

Research on public goods has been one of the most important economic problems after Samuelson (1954). A *pure* public good is characterized by the following two properties: (1) non-excludability: no agent can be excluded from consuming the public good, and (2) non-rivalriness: consumption of the public good by one agent does not decrease the quantity available for consumption by any other agents (see Samuelson, 1954; Musgrave and Musgrave, 1973). Although both mechanism designers and experimentalists have been investigating efficient provision of public goods, their focus is frequently on *excludable* public goods. The reason is that free-riders are easily excluded from the benefit of public goods through an organization such as a club. However, the problems arising from non-excludability are increasingly important in many practical circumstances, such as for international treaties.

A recent example is the Kyoto Protocol to cope with global warming and climate change. It took years to agree on the basic framework, the United Nations Framework Convention on Climate Change (UNFCCC), to reduce the green house gases. UNFCCC was adopted in 1992 and entered into force in 1994. The parties of UNFCCC adopted the Kyoto Protocol in 1997. The Protocol is a mechanism in our terminology to attain the aim of UNFCCC. As of July 2003, 84 parties including the United States have signed the Protocol, and 111 parties have either ratified or acceded to it. In March 2001, however, President Bush announced that the U.S. would not ratify the Protocol because it is detrimental to the U.S. economy. This could limit the effectiveness of the Protocol.¹

Voluntary public goods provision by individuals -- such as for public broadcasting -- often faces a similar problem. For example, a part of public broadcasting in Japan is supported by the public broadcasting fee. Every family must pay the fee by law, but many choose not to and enjoy the benefit of the public broadcasting that is non-excludable, since enforcement is practically non-existent. A natural question to ask is what would happen if we allow agents to commit to contributing nothing before they play the voluntary contribution mechanism.

Mechanism designers such as Groves and Ledyard (1977), Hurwicz (1979), Walker (1981) and Dutta et al. (1995) and their followers constructed mechanisms achieving Pareto efficient outcomes with several other normative criteria, but all agents in these mechanisms must play a strategy specified in them. That is, agents do not have freedom to not play any strategy so that they can free-ride on the benefit from the pure public good provided by others.

Experimental research on the provision of public goods has focused on the voluntary contribution mechanism.² All subjects in the experiments must choose a number corresponding to their amount of contributions. Although zero contribution is often an option, subjects cannot refuse to choose a number and at the same time enjoy the benefit from the public goods.

Recently, Saijo and Yamato (1999, 2001) shed new light on this aspect of pure public good provision mechanisms. For example, in the voluntary contribution mechanism, agents may have a choice to commit to contributing nothing before they play the game, and hence some of them can commit to free-ride. Saijo and Yamato (2001) proved an impossibility theorem stating that it is impossible to design a mechanism where all agents choose to commit to play the game in very reasonable environments.³

This paper reports an experiment in which non-excludability is incorporated explicitly in the voluntary contribution mechanism. The following are the major features that distinguish it from other public goods experiments.

First, in order to introduce non-excludability, we model the voluntary contribution mechanism with pre-commitment in a two-stage game following Saijo and Yamato (1999). In the first stage, two subjects choose simultaneously whether or not they commit to contributing nothing to provide a pure public good. In the second stage, knowing the other subject's commitment decision, subjects who elected not to commit in the first stage choose contributions to the public good. Second, we employed a *non-linear* payoff function rather than the linear payoff function used in most of the previous experiments. Subjects receive payoffs based on a Cobb-Douglas transformation of their consumption of the public good and their private good. In our design, two subjects have the same non-linear payoff function. Therefore, the equilibrium outcome is interior rather than on the boundary of the strategy space. See Laury and Holt (2003) for a survey of the limited number of voluntary contribution game studies with interior Nash equilibria.

Third, we designed experiments in which the information is as complete as possible. We explained to subjects that everyone has the same payoff table and the same initial holdings. In addition, everyone knew the total number of repetitions. Moreover, we provided a payoff table called a *detailed table* that has complete payoff information and is qualitatively different from the *rough* payoff tables of previous experiments.⁴ Our detailed table has information of

two dimensions due to the non-linearity of the payoff function. Most of previous experiments used rough tables that have information of one dimension. We did not provide population feedback, however, regarding choices of other subjects outside the players own current pairing. For example, information on commitment decisions in the first stage is common knowledge between paired subjects, but no pair learns the commitment or contribution decisions of other pairs.

Fourth, we used only two subjects in each group. The purpose of this design feature is twofold. First, we wanted to study the strategic behavior of subjects in a most simple environment. Second, we wanted an environment that fit well into basic evolutionary game theory: each treatment had twenty subjects and each subject was randomly paired with each other subject one at a time – a so-called “strangers” design. The same game was repeated 19 periods, 4 for practice and 15 for monetary reward, so as not to pair the same two subjects more than once.⁵

The above features of the experiment allow us to classify all strategies into three categories in the second stage: own payoff-maximizing, altruistic, and spiteful. If subject 2 commits to contributing nothing in the first stage and subject 1 does not, subject 1 who can choose a public good investment number has three possible strategies. Subject 1’s best response (that maximizes her own payoff) to the zero investment of subject 2 is called an own payoff-maximizing strategy.

On the other hand, subject 1 could invest more than the own payoff-maximizing strategy investment so that both subjects could enjoy an even higher level of the public good. Since subject 1’s investment level exceeds the own payoff-maximizing level, subject 1 suffers payoff loss comparing with the own payoff-maximizing strategy. We call this strategy an *altruistic* strategy. In our payoff setting, every investment that exceeds the own payoff-maximizing strategy falls into this category.

Another type of strategy is to invest less than the own payoff-maximizing strategy. Although subject 1 suffers payoff loss when she reduces the level of public good relative to her own payoff-maximizing investment, subject 2 suffers an even greater payoff loss than subject 1 does. This is because some reduction of investment from the own payoff-maximizing strategy does not hurt subject 1 much due to the first order condition at the own payoff-maximizing

strategy.⁶ We call this strategy a *spiteful* strategy. In our payoff setting, every investment less than the own payoff-maximizing strategy satisfies this condition.

When neither subject commits to the zero investment, all strategies can be classified into these three categories in a similar manner, although it is necessary to make assumptions about subjects' beliefs regarding the other subject's investment.

The Prisoners' Dilemma game represents the typical linear voluntary contribution mechanism without pre-commitment. However, in our two stage game setting, the normal form game representation of the first stage commitment decision is a Hawk-Dove game rather than the Prisoners' Dilemma game. As usual, this Hawk-Dove game has two pure strategy Nash equilibria and one mixed strategy Nash equilibrium that is the unique evolutionarily stable strategy (ESS) equilibrium.

We observed the following results. In our setting, subjects can easily recognize that free-riding is an option (i.e., subjects can commit to investing nothing in the first stage). Hence, one might expect more free-riding than in the usual voluntary contribution mechanism experiments in which a typical contribution pattern is early-period cooperation with eventual decay toward the free-riding outcome. However, we observed that the non-commitment rate, which is the ratio of the number of non-committing subjects to the total number of subjects, increased as periods advanced. In other words, the free-riding rate declined over time. Consequently, in the final two-thirds of our experiment, subjects' non-commitment rates nearly always exceeded the ESS equilibrium non-commitment rate.

Why did this happen? A typical subject, say subject 1, behaved as follows. In the early periods subject 1 committed to investing zero, expecting a high payoff with free-riding. However, when her opponents did not commit to investing zero, they often did not choose the own payoff-maximizing investment strategy. Instead, these opponents often played a *spiteful* strategy by investing a smaller amount so as to reduce subject 1's payoff more than their own payoff reduction. In fact, 68.4% of investment strategies chosen when one player committed to investing zero were spiteful, 30.9% were own payoff-maximizing, and 0.7% were altruistic. This spiteful behavior occurred even though spiteful subjects knew that they would not play the same subjects again in our experimental design. That is, the "punishment" through spitefulness could not have *direct* influence on payoffs in subsequent periods. Rather, punishment could have had only an *indirect* effect.

Nevertheless, the committing subject learned that commitment to investing zero was not beneficial to her because of the spiteful response by non-committing subjects, and hence she began regularly not committing. That is, it seemed that the reason why the ratio of non-committing subjects increases is not altruism or kindness, but instead is a strategic response to the spiteful behavior of other subjects.

The remainder of the paper is organized as follows. In Section 2 we explain the voluntary contribution mechanism with pre-commitment to contributing zero. Section 3 describes the experimental design. We present the results of the experiment in Section 4, and Section 5 concludes.

2. The Voluntary Contribution Mechanism

2.1. The Basic Model

There are two subjects, a and b , and subject i ($=a,b$) has w_i units of initial endowment of a private good. Each subject faces a decision of splitting w_i between her own consumption of the private good (x_i) and investment (y_i). The level of the public good each subject receives from the investments is $y = y_a + y_b + w_y$, where w_y is the initial level of the public good. Therefore, each subject's decision problem is to maximize her payoff $u_i(x_i, y)$ subject to the constraint $x_i + y_i = w_i$. We assume that all subjects have the same payoff function that is a monotonic transformation of a Cobb-Douglas type function: $u_i(x_i, y) = \{x_i^\alpha y^{1-\alpha}\}^\beta / 50 + 500$. We set $(w_a, w_b, w_y) = (24, 24, 3)$, $\alpha = 0.47$, and $\beta = 4.45$. With these parameters the Nash equilibrium investment pair of the voluntary contribution mechanism is $(\hat{y}_a, \hat{y}_b) = (7.69, 7.69)$ and the equilibrium level of the public good is $\hat{y} = \hat{y}_a + \hat{y}_b + w_y = 18.38$. The Pareto efficient level of the public good is $y^* = y_a^* + y_b^* + w_y = 12.02 + 12.02 + 3 = 27.04$, which is determined uniquely by the Samuelson condition and the feasibility condition. Clearly, the level of the public good with the voluntary contribution mechanism \hat{y} is less than the Pareto efficient level of the public good y^* . In our experiment, subjects choose integer investment numbers only. Hence the Nash equilibrium of this game is for each subject to contribute 8. No other Nash equilibria sneak into our model due to the discrete strategy choice set.

2.2. A Two-Stage Game with Pre-commitment to Contributing Nothing

In the above basic model, we have assumed implicitly that no subjects are allowed to pre-commit to contributing zero in the voluntary contribution mechanism. However, Saijo and Yamato (1999) show that there is a wide class of public goods provision mechanisms where subjects have incentives to pre-commit to investing nothing for public goods. The voluntary contribution mechanism is one of them. Consider now a two-stage game (see Figure 1). In the first stage, each subject simultaneously decides whether or not she should commit to investing zero in the voluntary contribution mechanism without knowing the other subject's decision. In the second stage, each subject decides how many units of her initial endowment she should invest after learning the other subject's commitment decision.

 Figure 1 is around here

Notice that commitment to investing zero is different from zero investment without commitment. Once subject *a* decides not to commit to investing zero, subject *b* must take account of this fact when she chooses her investment number *without* knowing subject *a*'s investment number. On the other hand, if subject *a* chooses zero commitment, then subject *b* knows that subject *a* invests nothing.

If neither subject decides to commit to investing zero, then the Nash equilibrium of that subgame is for each subject to contribute 8 and obtain a payoff of 7345 (see cell *c* of Table 1 in which the rows are for the subject's own investment numbers and the columns are for the other subject's investment numbers). If one subject commits to investing zero and the other does not, then the subject who chose non-commitment maximizes her payoff at $y_i = 11$ and obtains a payoff of 2658 (cell *ℓ* of Table 1), and the subject who chose commitment clearly invests nothing and obtains a payoff of 8278 (cell *m* of Table 1). If both subjects choose to commit to investing zero, both end up with a payoff of 706. These subgame equilibrium payoffs are incorporated into the normal form game payoff table shown in Table 2.

 Tables 1 and 2 are around here

The game in Table 2 is a well-known Hawk-Dove game. Although the usual simplification of the public good problem is a Prisoners' Dilemma game, we find that the proper simplification is a Hawk-Dove game when we allow subjects to commit to investing zero. There are two pure strategy Nash equilibria: either one of subjects commits to investing

zero. One more Nash equilibrium is a mixed strategy equilibrium: each subject i chooses 0.68 as her non-commitment probability p_i . Among these three equilibria, the mixed strategy equilibrium is a unique ESS equilibrium.⁷

2.3. Classification of Strategies: Own Payoff-Maximizing, Spiteful, and Altruistic Strategies.

The subgame perfect equilibrium analysis in the above subsection is based on the assumption of own payoff-maximizing behavior: each subject chooses a strategy maximizing her own payoff. However, we observed in our experiment that subjects selected strategies that appeared altruistic and spiteful. Therefore, it is useful to introduce formal definitions of altruistic strategies and spiteful strategies. We regard own payoff-maximizing behavior as a standard of comparison, and then classify all possible strategies into three regions.

Definitions: 1) A subject is said to choose an *own payoff-maximizing strategy* if she selects a strategy maximizing her own payoff, given an expected strategy of the other subject.

2) A subject is said to choose an *altruistic strategy* if she selects a strategy reducing her own payoff, but increasing the other subject's payoff in comparison to the own payoff-maximizing payoffs, given an expected strategy of the other subject.

3) A subject is said to choose a *spiteful strategy* if she selects a strategy reducing both her own payoff and the other subject's payoff in comparison to the payoffs when she takes an own payoff-maximizing strategy, given an expected strategy of the other subject. It is also useful to distinguish spiteful strategies into two subcategories in our two-stage game. A spiteful strategy is called "punishably spiteful" if the other subject pre-commits to contributing nothing, while it is called "rivalistically spiteful" otherwise.

Table 1 illustrates how strategies are classified in our setting. Now suppose that you decide not to commit to investing zero, but your opponent does. Then investing less than your own payoff-maximizing strategy level, 11, is a punishably spiteful strategy. For instance, by investing 7 instead of 11, your own payoff is reduced from 2658 (cell ℓ) to 2210 (cell n), and your opponent's payoff is also reduced from 8278 (cell m) to 4018 (cell p). On the other hand, investing more than 11 is an altruistic strategy. For instance, by investing 17 instead of 11, your own payoff is reduced from 2658 to 1871 (cell q), while your opponent's payoff increases from 8278 to 18539 (cell r).

Next suppose that neither you nor your opponent pre-commits to contribute nothing, and that you expect your opponent to choose 8. Then investing less than your own payoff-

maximizing strategy level, 8, is a rivalistically spiteful strategy. For instance, if you invest 6 instead of 8, your own payoff is reduced from 7345 (cell *c*) to 7237 (cell *d*), and the payoff of your opponent is also reduced from 7345 to 5766 (cell *e*). On the other hand, investing more than 8 is an altruistic strategy. For example, by investing 16 instead of 8, your own payoff is reduced from 7345 to 4179 (cell *f*), while the payoff of your opponent increases from 7345 to 16179 (cell *g*).

For any given investment number of the other subject, investing less than an own payoff-maximizing number is a spiteful strategy, and investing more than it is an altruistic strategy. In Table 1 the own payoff-maximizing strategy region represents 4.16% (=26/625) of the cells, the spiteful strategy region is 22.40% (=140/625), and the altruistic strategy region is 73.4% (=459/625). In the case of neither committing to invest zero, whether an investment choice of a subject is own payoff-maximizing, rivalistically spiteful, or altruistic depends on how much she expects the other subject to invest. We will consider several possible expectations in the data analysis when classifying behavior into these three categories.

3. Experimental Design

Our experiment consisted of three sessions, each with 20 different subjects. In each session, the twenty subjects were seated at desks in a relatively large room and had random identification numbers. These identification numbers were not publicly displayed, however, so subjects could not determine who had which number. In each period we made ten pairs out of twenty subjects, and these ten pairs played the two-stage game with the commitment decision as described in the previous section. The pairings were anonymous and were determined in advance by experimenters so as not to pair the same two subjects more than once – a so-called “strangers” design. The first four periods were for practice and the remaining fifteen determined the subjects' monetary payoffs. Instructions were given by tape recorder to minimize the interaction between subjects and experimenters.

First, subjects decided whether or not they would commit to investing zero. In the experiment, we used the term “participate (resp. not participate) in investment” instead of “not commit (resp. commit) to investing zero.” These decisions were collected by experimenters and then redistributed only to their paired subjects. No information (such as the total number of subjects who chose commitment) or decisions were publicly announced. After the redistribution of the commitment decisions, subjects who decided not to commit to investing

zero chose their investment numbers on investment sheets by circling an integer between 0 and 24. In order to not reveal the number of subjects who chose commitment and to obscure the identity of those subjects, even those who had chosen to commit to zero also filled out this investment sheet, circling the phrase "Not Participate" (i.e., commit to investing zero). Experimenters collected these investment sheets and then redistributed them only to the paired subjects. No information (such as the average investment number) or decisions were publicly announced. During the redistribution, subjects were asked to fill out the reasons why they chose these numbers. After the redistribution of investment cards, subjects calculated their payoffs from the payoff tables. Then the next period started.

One of the three sessions differed slightly from the other two by the term used to describe the person that each subject is matched with at each period. In two sessions the term "your opponent" was employed in the instructions, record sheets and payoff tables. In order to investigate the influence of framing effects, the phrase "the person you are paired with" replaced "your opponent" in all materials of the third session. One might expect that the term "opponent" forces subjects to think in relative terms. We will discuss differences between the results for the two *Opponent* sessions and those for the *No Opponent* session below.

Every subject had the same payoff function and every subject knew this fact. We distributed a detailed payoff table, which is a plain table deleting the tags and shading from Table 1. All subjects were able to readily calculate their payoffs following the instructions and practice periods.⁸ We allowed subjects three minutes to examine the three payoff tables before the practice periods and ten minutes to examine the three new payoff tables before the real periods. The tables used for the practice and real periods were different.

We conducted one *Opponent* session at the University of Tsukuba, and one *Opponent* and one *No Opponent* session at the Tokyo Metropolitan University (TMU). We recruited the student subjects by campus-wide advertisement. These students were told that there would be an opportunity to earn money in a research experiment. None of them had prior experience in a public good provision experiment. No subject attended in more than one session. Each session required approximately two hours to complete. The mean payoff per subject was \$28.38 (\$1=100 yen). The maximum payoff among the sixty subjects was \$45.53, and the minimum payoff was \$17.28.

4. Experimental Results

Figure 2 shows that the frequency distribution of investment pairs for all three sessions. Each period had 10 pairs and 15 periods were conducted in each session, so our data consist of 450 choice pairs. The order of investment numbers does not matter, so we rearranged each pair (x,y) with $x \geq y$. The maximum frequency pair was $(8,7)$ with 57 pairs, the second most common was $(11,0)$ with 44 pairs, followed by $(8,8)$ and $(0,0)$ with 37 pairs each.

Figure 2 is around here

4.1 Commitment Data

We begin by examining whether the commitment data were compatible with the ESS equilibrium of the Hawk-Dove game described in Table 2. The non-commitment rate data from the three sessions are statistically indistinguishable in virtually all periods after period 2.⁹ This provides evidence that neither the experiment site (Tokyo versus Tsukuba) nor the experiment wording (“your opponent” versus “the person you are paired with”) affect commitment choices. Therefore, we pool the commitment data across the three sessions.

The null hypothesis is the ESS non-commitment probability of 0.68. We first conducted a binomial test separately by period in order to avoid pooling commitment decisions made by the same subject in the same test. Under the ESS null hypothesis, the probability of observing 10 commitment decisions or less out of 60 is less than one percent, and the probability of observing 12 commitment decisions or less out of 60 is less than five percent. As Figure 3 shows, the non-commitment rate rose as periods advanced (although a brief decline was observed in periods 8 and 9). The smooth curve is a simple log-linear regression. More to the point for this binomial test, the low commitment rate permits the binomial test to reject the ESS null hypothesis in 9 periods of the 15 total periods (in periods 5, 6, 7, 10, 11, 12, 13, 14, and 15) at the five percent (usually one percent) significance level.

Figure 3 is around here

We also examined the overall non-commitment rates for each of the 60 subjects separately. The mean non-commitment rate was 80% (12 of 15 decisions), and the median non-commitment rate was 86.7% (13 of 15 decisions). Note that the ESS rate of 0.68 implies on average slightly more than 10 non-commitment decisions. Only 14 of the 60 subjects (23.3%) did not commit 10 times or less, while the other 46 subjects (76.6%) did not commit 11 times or more. Fifteen of the 60 subjects (25%) were apparently using a pure strategy, or randomized

once at the beginning of the session and played the same realization in every period, as they did not commit in 15 out of 15 periods. Using the 60 separate subject observations, the data reject the ESS prediction of 0.68 at better than the 0.0001 significance level using the non-parametric Wilcoxon signed-rank test.¹⁰

4.2 Why were Non-commitment Rates High?

In order to understand these high non-commitment rates, consider the case in which only one subject in a pair did not commit to contributing zero. When a subject committed to investing zero, the other could obtain her maximum payoff by investing 11 (see Table 1). There were 136 observations with exactly one subject who did not commit. In these cases the person choosing non-commitment invested eleven 43 times, invested less than eleven 92 times,¹¹ and invested more than eleven (twenty-four) only 1 time. The mean investment was 6.90.

As shown in Table 1, by investing 11 in response to the other committing to investing zero, the committing subject earns 8278 while the non-committing subject earns 2658. On the other hand, by investing 7 in this situation, the committing subject earns 4018 while the non-committing subject earns 2210. That is, the payoff reduction of the non-committing subject (448=2658-2210) was relatively small, while the payoff reduction of the committing subject (4269=8278-4018) was relatively large. In fact, for this subgame we observed that over two-thirds of the data fall into the punishably spiteful region, and less than one-percent in the altruistic region in each of the three sessions, as shown in Table 3. By pooling the investment data for the case of only one committing subject across the three sessions and using a random effects model, we soundly reject the hypothesis that mean investment equals 11 ($t=8.14$).¹²

Table 3 is around here

As Table 2 shows, the payoff should be 7345 when neither subject committed to investing zero according to the Nash equilibrium prediction. However, the average payoffs of subjects in this case were less than 7000 units in all 15 periods. When only one subject committed to investing zero, the payoff for the committing player should be 8278 according to the Nash equilibrium, but the actual average payoffs for this player were less than 6000 in 12 of 15 periods. Even more striking, according to Table 2, the average payoff for a single

committing player should be above that in the case of neither committing, but the former was above the latter in only one period. These average payoffs are significantly different at the 5-percent level according to a nonparametric two-sample Wilcoxon test in periods 4, 7, 8, 10 and 12 (two-tailed tests).¹³ Table 4 illustrates the realized payoff matrix for commitment decisions, which is based on the average values of payoffs subjects actually obtained up to period 5, rather than the subgame equilibrium payoffs as shown in Table 2. As this table illustrates, non-commitment became a dominant strategy even in early periods.

 Table 4 is around here

4.3 Learning Processes for Non-commitment

Our interpretation of “as if subjects play a dominant strategy game” is based on subject learning and proceeds roughly as follows. After inspection of the payoff tables, some subjects initially commit to investing zero hoping the other will not commit and invest 11. They therefore expect (perhaps with the ESS probability of 0.68) to receive a payoff of 8278. However, since their non-committing opponent invested less than 11, the subject realized that her earnings in this subgame fell below 6000 on average. After learning this, she chose not to commit to investing zero in later periods and frequently earned more than 6000.¹⁴

Essentially, subjects learn (contrary to the ESS equilibrium) that expected payoffs from non-commitment tend to exceed those from commitment. This subsection summarizes a simple model to document this learning process. Many alternative approaches to learning have been advanced recently in the literature, including reinforcement learning (e.g., Erev and Roth, 1998), belief-based learning (Cheung and Friedman, 1997), and creative hybrid approaches (e.g., Camerer and Ho, 1999). Rather than provide an exhaustive evaluation of the various learning models using our data, we just consider a simple adaptive learning model in which the commitment choice is reinforced based on previous earnings. In particular, we estimate a probit model in which the probability of non-commitment depends on the ratio of expected non-commitment earnings (ENCE) to expected commitment earnings (ECE): $\text{Probability}(\text{Non-commitment}) = f(\text{ENCE}/\text{ECE})$, where f is the normal density function $(e^{-x^2/2} / \sqrt{2\pi})$ since this is estimated as a probit model.

The next step is to specify the process underlying subjects' expectations. In terms of Camerer and Ho's (1999) experienced-weighted attraction (EWA) model – which nests basic

choice reinforcement and belief-based learning models as special cases – we assume that hypothetical payoffs to strategies not chosen are not updated (i.e., EWA parameter $\delta = 0$), as in many reinforcement learning models. A reinforcement learning approach seems most plausible for our design with randomly re-paired subjects, since this design feature prevents subjects from developing beliefs about individual players. We also imposed the $\delta = 0$ restriction because hypothetical payoffs are unknown in this extensive form game for the subgame not played.¹⁵ We update payoff expectations following each commitment or non-commitment choice, and we evaluate two polar cases in this simple model: (1) Cournot (or myopic) expectations and (2) Fictitious Play expectations (e.g., Cheung and Friedman, 1997; Cox and Walker, 1998).

According to Cournot, ENCE are simply the realized earnings the last time the subject did not commit; and ECE are simply the realized earnings the last time the subject committed. In other words, subjects maintain a very short (myopic) memory length--of one observation for each (commit or not) decision. This is consistent with EWA parameters $\rho = \phi = 0$, since previous experience completely depreciates. By contrast, according to Fictitious Play, subjects have a long memory, and each observation updates the expectation with a declining weight. For example, if a subject has not committed N times up to this period, and they did not commit in this period, they update ENCE as follows: $ENCE = ((N * \text{previous ENCE}) + \text{current non-commitment earnings}) / (N + 1)$. In other words, as subjects accumulate evidence they simply include it in their running average of the payoffs from non-commitment for ENCE. ECE is, of course, analogous. This fictitious play alternative is consistent with EWA parameters $\rho = \phi = 1$ since all observations count equally.¹⁶

We estimate this probit model with a random subject effect, separately and pooled for the three sessions. For Fictitious Play, the expected payoff ratio is not significantly different from zero except in the TMU *No Opponent* session, where the coefficient estimate has the wrong sign. For the Cournot specification, the ratio is significantly positive except in the TMU *No Opponent* session. The positive coefficient on the ratio implies that as the relative profitability of non-commitment increases, the likelihood of non-commitment increases. So, we can conclude that (1) subjects' non-commitment decisions respond to their experience, and (2) subjects appear to update their expectations in this environment using a short (Cournot) memory length.

Summarizing the above observations, we have the following:

Observation:

(a) *The ESS prediction regarding the non-commitment rate is rejected.*

(b) *The non-commitment rate rises as periods advance.*

(c) *It seems that the reason why the rate of non-committing subjects increases is not altruism or kindness but is spiteful behavior of subjects. Subjects learn that commitment to investing zero will invoke a spiteful response, which reduces the payoff of commitment below the payoff of non-commitment.*

Our observation is consistent with what one would expect in light of the results observed in ultimatum game experiments (e.g., Ochs and Roth, 1989; Prasnikar and Roth, 1992). In these experiments "proposing too little" invoked a spiteful response of "rejecting the proposal". Subjects quickly learned this and ceased proposing too little at the first stage. The outcome was not a subgame perfect equilibrium, like our result.

4.4 An Evaluation of Our Observations using Recent Models of Social Preferences

The choices observed in our experiment could be consistent with notions of "inequality aversion" or reciprocal altruism advanced recently by some researchers (e.g., see Rabin, 1993; Levine, 1998; Dufwenberg and Kirchsteiger, 2003; Falk and Fischbacher, 1998; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Charness and Rabin, 2002; Costa-Gomes and Zauner, 2001). Our experiment was not designed to differentiate between these alternative models, so we do not wish to overstate what our experiment can say about them. Nevertheless, a brief evaluation of our results in the context of these new models is worthwhile.

In Fehr and Schmidt's (1999) model subjects have "inequity aversion," so they prefer higher but more equal earnings among participants in their group. Utility payoffs are equal to monetary payoffs less inequity costs that rise as the difference between a subject's own and other's monetary payoff increases.¹⁷ In the model some subjects suffer both from earning more as well as earning less than their counterparts, but the cost of advantageous inequality is assumed to be no more than the cost of disadvantageous inequality. Fehr and Schmidt demonstrate that their model can describe many outcomes in ultimatum games, market games with both proposer and responder competition, as well linear voluntary contribution mechanism games with and without punishment opportunities. They even derive parameter distributions of the relative tradeoff of monetary gains and inequity aversion that describes

behavior across games, which we can use to assess conveniently the effectiveness of this approach in describing the new data reported here. Applying their distribution of preferences to our subjects, it is straightforward to show that when only one subject does not commit to investing zero, the optimal contribution is 11 for 30 percent of the subjects (these 30 percent are standard “money-maximizers”), is 6 for 30 percent of the subjects, is 4 for another 30 percent of the subjects, and is 1 for the remaining 10 percent.¹⁸ The mean of this distribution is 6.4.

The distribution of contributions in our data is remarkably close to this predicted distribution if one makes allowances for a bit of choice error for the lower contributions. When they are the only non-committing subject, 32 percent of our subjects contributed 11, 26 percent contributed 6 or 7, 11 percent contributed 3 or 4, and 13 percent contributed 0 or 1. The close correspondence between our Japanese data and the Fehr and Schmidt model predictions may at first seem surprising, because the model parameters were calibrated using data collected in Europe and North America. But the Fehr and Schmidt model – and all the related models just cited – are culturally neutral and so they should apply to non-Western cultures as well.

A well recognized drawback of the approach taken by Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) is that they model players’ utilities as depending only on the final payoff allocations and not on players’ intentions. Rabin’s (1993) fairness equilibrium, Dufwenberg and Kirchsteiger’s (2003) model of sequential reciprocity, and Charness and Rabin’s (2002) model of social preferences allow for (positive or negative) reciprocal behavior based on how kind or unkind a player believes his opponent is treating him. Our experiment provides fairly clear evidence that non-committing subjects act spitefully to punish committing subjects, or in other words they exhibit negative reciprocal behavior (see Table 3).

4.5 The Framing Effect of the “Opponent” Wording on Investment Data in the Case of Neither Committing to Investing Zero

The commitment rates and investments when only one subject pre-commits to investing zero exhibit virtually no significant differences in the data based on the experiment site or the experiment wording, i.e., “your opponent” versus “the person you are paired with”. However, there is a modest effect of the “opponent” framing for the case in which neither pre-commits to contributing nothing: relatively higher investments and more altruistic strategies are chosen in the session without the “opponent” framing than in the sessions with the opponent framing.

First, we compare investments in the case of neither committing across the three sessions. According to period by period nonparametric Wilcoxon rank sum tests, the average investment in the TMU *No Opponent* session is significantly higher than in the TMU *Opponent* session in 3 of the 15 periods (periods 2, 3, and 12); and it is significantly higher than in the Tsukuba (*Opponent*) session in 2 of the 15 periods (periods 2 and 14).¹⁹ That is, the average investments are different in some periods, although the rate that we reject the null hypothesis of no framing treatment difference is still relatively low.

Next we compare how frequently own payoff-maximizing, rivalistically spiteful, or altruistic behavior (defined in Section 2.2) occurs across the three sessions. There are several ways in which subjects could construct expectations for the investment number of the other subject, and the frequency classification of the behavior may depend on types of expectations. We consider the following types: 1) *Nash equilibrium expectation*: each subject expects the other subject to choose the Nash equilibrium investment, 8; 2) *Cournot expectation*: each subject expects the investment number of the other subject to be the same as the observed investment of the other subject in the previous period; and 3) *An average expectation with a declining weight (fictitious play)*: each subject has a long memory when she forms expectations for the other's investment number, in contrast to the Cournot expectation case in which each subject has a short memory length. This is simply a recursive version of fictitious play expectations.²⁰

Table 3 shows the rates of rivalistically spiteful, own payoff-maximizing, and altruistic strategies chosen by subjects in each of the three sessions for each type of expectation. In the two sessions with the *Opponent* wording, most data fall into the rivalistically spiteful strategy region and the own payoff-maximizing strategy region, and only about 10% of the data fall in the altruistic strategy region for all types of expectations. On the other hand, in the session without the opponent wording, 20 to 30% of the data are in the altruistic strategy region depending on the type of expectations.²¹ Nevertheless, in spite of these modest framing effect differences when neither subject commits, even in this subgame the overall characteristics of the data are robust: the ratio of altruistic strategies is still smaller than the ratios of rivalistically spiteful and own-payoff maximizing strategies; the ratio of rivalistically spiteful strategies is the largest for almost all sessions and types of expectations; and subjects on average chose investments below the payoff-maximizing level.

This result is similar to the findings in Andreoni (1993) and Chan et al. (2002), who studied the crowding out hypothesis in public good experiments with non-linear payoff functions and minimum contribution “taxes” in some treatments, though they did not refer to rivalistically spiteful behavior. They also used Cobb-Douglas type payoff functions and presented complete payoff matrices that showed how payoffs depend on both own and other’s contributions to the public good, as in our experiments. In both of these previous experiments contributions were close to but slightly below the Nash equilibrium prediction at most periods when no tax was imposed. The contribution levels did not get closer to the interior Nash equilibrium over time, similar to our data. Laury and Holt (2003) also note that complete, detailed payoff information in this type of non-linear environment seems to reduce contribution levels.

These observations are different from those in standard public goods experiments with linear payoff functions, where contributions usually exceed the Nash equilibrium level. However, it is easy to see that contributing zero is an own payoff-maximizing strategy and contributing any positive amount is an altruistic strategy, that is, only own payoff-maximizing and altruistic strategies can be chosen and no rivalistically spiteful strategy is available with linear payoff functions (see Saijo et al. (2002) for details).

6. Concluding Remarks

Our data reject the ESS prediction for this two-stage (Hawk-Dove) voluntary contribution game with an initial opportunity to commit to a contribution of zero. The number of non-committing subjects exceeds the ESS prediction and increased across time. Furthermore, this increase did not arise due to altruism or kindness among subjects, but from their *spiteful* behavior. Acting spitefully in this way is costly, and this kind of spiteful or negative reciprocal behavior has also been observed recently by independent research on public goods (Fehr and Gächter, 2000) and in the ultimatum game.

In neo-classical economic theory, it is assumed that each agent cares only about himself and maximizes his own payoff subject to some constraints. If people care about how they are doing relative to others (for example, see Hume, 1739),²² however, then it is natural to think that they might often take *spiteful* actions in an attempt to decrease the happiness of others. One would think that such spiteful behavior might result in outcomes that are socially inferior to outcomes arising from the interaction of purely selfish individuals. We find that the

opposite may occur: spitefulness leads to cooperation in the sense that the number of subjects who do not commit to contributing nothing to the provision of a public good increases. This finding suggests a need to rethink our fundamental assumptions of human nature and the implications of alternative assumptions for our models.

REFERENCES

- Akerlof, G. A., Yellen, J. L., 1985. Can Small Deviations from Rationality Make Significant Difference to Economic Equilibria? *Amer. Econ. Rev.* 75, 708-720.
- Andreoni, J., 1993. An Experimental Test of the Public-Goods Crowding-Out Hypothesis. *Amer. Econ. Rev.* 83, 1317-1327.
- Bolton, G., Ockenfels, A., 2000. ERC: A Theory of Equity, Reciprocity and Competition. *Amer. Econ. Rev.* 90, 166-193.
- Camerer, C., Ho, T.-H., 1999 Experienced Weighted Attraction Learning in Normal-Form Games. *Econometrica* 67, 827-874.
- Chan, K., Godby, R., Mestelman, S., Muller, R. A., 2002. Crowding Out Voluntary Contributions to Public Goods. *J. Econ. Behav. Organ.* 48, 305-317.
- Charness, G., Rabin, M., 2002. Understanding Social Preferences with Simple Tests. *Quart. J. Econ.* 117, 817-869.
- Chen, Y., Plott, C. R., 1996. The Groves-Ledyard Mechanism: An Experimental Study of Institutional Design. *J. Public Econ.* 59, 335-364.
- Cheung, Y.-W., Friedman, D., 1997. Individual Learning in Normal Form Games: Some Laboratory Results. *Games Econ. Behav.* 19, 46-76.
- Costa-Gomes, M., Zauner, K., 2001. Ultimatum Bargaining Behavior in Israel, Japan, Slovenia, and the United States: A Social Utility Analysis. *Games Econ. Behav.* 34, 238-269.
- Cox, J., and Walker, M., 1998. Learning to Play Cournot Duopoly Strategies. *J. Econ. Behav. Organ.* 36, 141-161.
- Dixit A., Olson M., 2000. Does Voluntary Participation Undermine the Coase Theorem? *J. Public Econ.* 76, 307-335.
- Dufwenberg, M., Kirchsteiger, G., 2003. A Theory of Sequential Reciprocity. *Games Econ. Behav.* forthcoming.
- Dutta, B., Sen, A., Vohra, R., 1995. Nash Implementation through Elementary Mechanisms in Economic Environments. *Econ. Design*, 1, 173-204.
- Erev, I., Roth, A., 1998. Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *Amer. Econ. Rev.* 88, 848-881.
- Falk, A., Fischbacher, U., 1998. A Theory of Reciprocity. University of Zurich, Working Paper No. 6.
- Fehr, E., Gächter, S., 2000. Cooperation and Punishment in Public Goods Experiments. *Amer. Econ. Rev.* 90, 980-994.

- Fehr, E., Schmidt, K. M., 1999. A Theory of Fairness, Competition, and Cooperation. *Quart. J. Econ.* 114, 817-868.
- Groves, T., Ledyard, J., 1977. Optimal Allocation of Public Goods: A Solution to the 'Free Rider' Problem. *Econometrica* 45, 783-811.
- Ho, T.-H., Camerer, C., Wang, X., 2002. Individual Differences in EWA Learning with Partial Payoff Information. Caltech Working Paper.
- Hume, D. 1739. *A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects, Book II Of the Passions*, Prometheus Books, (1992).
- Hurwicz, L., 1979. Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Points. *Rev. Econ. Stud.* 46, 217-225.
- Issac, R.M., Walker, J., 1998. Nash as an Organizing Principle in the Voluntary Provision of Public Goods: Experimental Evidence. *Experimental Econ.* 1, 191-206.
- Laury, S., Holt, C., 2003. Voluntary Provision of Public Goods: Experimental Results with Interior Nash Equilibria. In: Plott, C., Smith, V. (Eds.), *The Handbook of Experimental Economic Results*, Elsevier Press, Amsterdam, forthcoming.
- Levine, D., 1998. Modeling Altruism and Spitefulness in Experiments. *Rev. Econ. Dynamics* 1, 593-622.
- Maynard Smith, J., 1982. *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.
- Moulin, H., 1986. Characterizations of the Pivotal Mechanism. *J. Public Econ.* 31, 53-78.
- Musgrave, R. A., Musgrave, P., 1973. *Public Finance in Theory and Practice*, McGraw-Hill, New York.
- Ochs, J., Roth, A. E., 1989. An Experimental Study of Sequential Bargaining. *Amer. Econ. Rev.* 79, 355-84.
- Palfrey, T., Rosenthal, H., 1984. Participation and the Provision of Discrete Public Goods: a Strategic Analysis. *J. Public Econ.* 24, 171-193.
- Prasnikar, V., Roth, A. E., 1992. Considerations of Fairness and Strategy: Experimental Data from Sequential Games *Quart. J. Econ.* 107, 866-888.
- Rabin, M., 1993. Incorporating Fairness into Game Theory and Economics. *Amer. Econ. Rev.* 83, 1281-1302.
- Saijo, T., Nakamura, H., 1995. The "Spite" Dilemma in Voluntary Contribution Mechanism Experiments. *J. Conflict Resolution* 39, 535-560.
- Saijo, T., Yamato, T., 1999. A Voluntary Participation Game with a Non-Excludable Public Good. *J. Econ. Theory* 84, 227-242.
- Saijo, T., Yamato, T., 2001. Voluntary Participation in the Design of Non-excludable Public Goods Provision Mechanisms. ISER Discussion Paper, Osaka University.
- Saijo, T., Yamato, T., Yokotani, K., Cason, T., 2002. Non-excludable Public Good Experiments. Social Science Working Paper No. 1154, California Institute of Technology.
- Samuelson, P. A., 1954 The Pure Theory of Public Expenditure. *Rev. Econ. Stat.* 36, 387-389.

Walker. M., 1981. A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations.
Econometrica 49, 65-71.

Footnotes

¹ Another contemporary example is the chemical weapons convention (CWC). Several countries suspected of developing chemical weapons, such as Afghanistan, Iraq, Libya, North Korea, and Syria, have not yet ratified the CWC. This could limit the effectiveness of the CWC. An important historical example is the League of Nations. Following World War I President Woodrow Wilson strongly supported the League, but the U.S. Congress never ratified the Treaty of Versailles and so the U.S. never joined the League.

² Another line of experimental research in public goods investigates performance of mechanisms achieving Pareto efficient outcomes such as the Groves-Ledyard mechanism. As Chen and Plott (1996) show, the Groves-Ledyard mechanism works very well under some suitable punishment parameters. But all agents must play strategies specified in the mechanism, consistent with the mechanism that they evaluate.

³ See also Dixit and Olson (2000), Moulin (1986), and Palfrey and Rosenthal (1984).

⁴ Saijo and Nakamura (1995) compared the effects of detailed payoff tables with rough payoff tables for a public goods environment with linear payoff functions.

⁵ Our design also differed from most earlier experiments because subjects responded to questions such as, what is “your reason for your decision on your investment number?” in each period rather than only after the end of a session. This method has advantages and disadvantages. Subjects might be able to justify what they did after the end of a session, but they might not be able to do so at each period. Of course, these types of questions in each period might distract subjects from their decision process or encourage them to focus on self-justification or rationalization. But the questions might also promote more thoughtful decisions.

⁶ For the maximizing agent (subject 1), the difference between the payoff at the maximum and the payoff at a strategy close to the maximum is small since the payoff function is approximately flat at the maximum. Akerlof and Yellen (1985) observed similar phenomena in

Keynesian business cycles and industrial organization theory: small deviations from maximizing behavior may cause changes in the equilibrium that are larger in magnitude.

⁷ See Maynard Smith (1982). Although expected payoffs from Commit and Not Commit are of course equal in the mixed strategy equilibrium, as an anonymous referee points out Commit is considerably riskier (payoff standard deviation of 3542) than Not Commit (payoff standard deviation of 2193). This could discourage risk averse subjects from committing to investing zero in this two-stage game.

⁸ In order to minimize the likelihood of any possible misunderstanding, we also presented a payoff table summarizing average payoffs for sets of 9 or 12 payoff cells as well as an iso-payoff map. Most subjects indicated in their post-experiment questionnaire that they understood all three kinds of payoff tables and used the detailed payoff table only.

⁹ See Saijo et al. (2002) for detailed comparisons of the three sessions.

¹⁰ We also conducted a binomial test of the mixed strategy of $p = 0.68$ separately for each subject. Under the null hypothesis that subjects play this mixed strategy and randomize each period, commitment decisions--even for an individual subject--are statistically independent. At the five percent significance threshold, 23 of 60 subjects (38.3%) did not commit too much (either 14 or 15 times), rejecting $p = 0.68$; 6 of 60 subjects (10%) did not commit too little (7 times or less), rejecting $p = 0.68$ in the other direction. At the ten percent significance threshold, 32 of 60 subjects (53.3%) did not commit too much (13, 14 or 15 times), rejecting $p = 0.68$; 9 of 60 subjects (15%) did not commit too little (8 times or less), rejecting $p = 0.68$ in the other direction.

¹¹ Specifically, the non-committing subject invested zero 14 times, invested one 3 times, invested two 7 times, invested three 8 times, invested four 7 times, invested five 1 times, invested six 16 times, invested seven 19 times, invested eight 9 times, invested nine 6 times, and invested ten 2 times.

¹² The nonparametric Wilcoxon test rejects the null hypothesis that average investment equals 11 at the five percent level in 9 of the 15 periods, even though the average sample size per period is only about 9 because of the high non-commitment rate.

¹³ In this test we only employ one observation from each pair of subjects for each period (their average payoff if neither commits to zero) since the two payoffs in a pair are not independent.

¹⁴ One might say that in our two-stage game, the low investment by the single non-committing subject could be interpreted as belonging to a tit-for-tat strategy to “teach” others to cooperate. Because subjects were re-paired with a new opponent each period and never interacted with the same subject in more than one period, however, such a strategy is not subgame perfect. Furthermore, although there is no need to choose a tit-for-tat strategy at the final period of the experiment, the ratio of such a strategy at the final period is equal to $6/9 = 67$ percent.

¹⁵ An anonymous referee, however, has brought to our attention a recent extension of the original EWA model to incorporate unobserved payoffs, which makes EWA applicable for extensive form games (Ho, Camerer and Wang, 2002). When (foregone) payoffs from unchosen strategies are unknown, Ho et al. approximate the unobserved payoffs in the learning model with either (1) the last payoff observed when previously choosing that strategy or (2) the actual foregone payoff as if it were known. In their initial application the learning model predicts better with this latter approximation, which they refer to as “payoff clairvoyance.”

¹⁶ For both Cournot and Fictitious Play, we need the expectations to start somewhere when no evidence has yet accumulated. For these initial expectations we employ the ESS expected payoffs, which are 5829 for both ENCE and ECE. Therefore, the ENCE/ECE ratio is 1 in period 1.

¹⁷ In particular, for a two-person game player i 's utility is

$U_i(x) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}$, $i \neq j$, where x_k denotes monetary earnings ($k = i, j$), $\alpha_i \geq \beta_j$, and $1 > \beta_i \geq 0$.

¹⁸ For this calculation one only needs the distribution of α , because the non-committing subject's earnings are always lower than the committing subject's earnings. β is only used for cases of advantageous inequality. We use Fehr and Schmidt's distribution of $\alpha = \{0, 0.5, 1, 4\}$ in proportions of $\{0.3, 0.3, 0.3, 0.1\}$.

¹⁹ On the other hand, in 14 of the 15 periods, there is no statistical difference in the average investment between the Tsukuba and TMU *Opponent* sessions at the five percent significance level, indicating that investments in the case of neither committing differed very little across the two sites when the framing was identical.

²⁰ Suppose that a subject faces the case of neither committing to zero in period t , and the last time she experienced such a case is period $s (< t)$. Let $EI(t)$ (resp. $EI(s)$) be the average expected investment number of the other subject in period t (resp. s). Then $EI(t)$ is given by $EI(t) = [I(s) + (N(s) - 1) \times EI(s)] / N(s)$, where $I(s)$ is the investment number of the other subject in period s and $N(s)$ is the number of the cases of neither committing to zero that the subject has experienced between period 1 and period s .

²¹ Mean investment when neither subject commits is 7.05 pooled across the two *Opponent* sessions, and using a random-effects panel data model we find that investments are significantly below the Nash equilibrium investment (8) with this framing ($t=3.93$). By contrast, mean investment when neither subject commits is 7.47 for the *No Opponent* session, which is not significantly different from the Nash equilibrium of 8 according to a random-effects model ($t=1.34$). Nonparametric Wilcoxon tests provide similar conclusions.

²² "Now as we seldom judge of objects from their intrinsic value, but form our notions of them from a comparison with other objects; it follows, that according as we observe a greater or less

share of happiness or misery in others, we must make an estimate of our own, and feel a consequent pain or pleasure. The misery of another gives us a more lively idea of our happiness, and his happiness of our misery. The former, therefore, produces delight; and the latter uneasiness."

Own Payoff-Maximizing Strategy Region

Your Investment Number

Your Opponent's Investment Number

Your Payoff	0	1	2	3	4	5	6	7	8	9	10	12	13	14	15	16	17	18	19	20	21	22	23	24	
0	706	871	1072	1297	1536	1775	2003	2210	2386	2523	2615	2658	2648	2585	2470	2309	2106	1871	1614	1349	1091	858	669	543	500
1	905	Spiteful Strategy Region					2427	2641	2816	<i>n</i>	3019	3039	3001	<i>l</i>	2755	<i>q</i>	2313	2038	1743	<i>i</i>	1484	894	685	548	500
2	1186						2913	3129	3297	3411	3465	3456	3385	3252	3061	2819	2534	2217	1881	1543	1220	933	703	552	500
3	1554						3463	3675	3831	3925	3952	3911	3801	3626	3391	3102	2770	2406	2027	<i>h</i>	1690	973	721	556	500
4	2017						4078	4281	4420	4488	4483	4403	4250	4028	3743	3404	3020	2608	2181	1759	1363	1015	740	561	500
5	2578						4762	4950	5064	5101	5057	4934	4733	4459	4119	3725	3287	2821	2344	1877	1441	1060	760	566	500
6	3244	3718	4171	4590	4960	5272	5515	5681	5766	<i>e</i>	5706	5504	5249	4918	4519	4065	3568	3045	2516	2000	1522	1106	781	571	500
7	4018	4529	5008	5440	5812	6115	6339	6478	6526	6481	6343	6114	5800	5406	4944	4425	3866	3282	2696	2129	1607	1155	802	576	500
8	4904	5447	5944	6383	6751	7038	7237	7340	7345	7250	7056	6765	6385	5924	5393	4806	4179	3532	2886	2265	1696	1206	825	582	500
9	5907	<i>p</i>	6984	7422	<i>d</i>	8043	8209	8271	8225	8073	<i>c</i>	7458	7007	6472	5867	5207	4508	3793	<i>f</i>	2407	1789	1259	849	588	500
10	7031	7616	8130	8561	8897	9132	9257	9270	9168	8951	8624	8193	7664	7051	6367	5628	4854	4067	3292	2555	1886	1315	874	594	500
11	8278	<i>m</i>	9800	10109	10306	10384	10339	10173	9886	9482	8970	8359	Altruistic Strategy Region					3509	2710	1987	1372	899	600	500	
12	9653	10250	10750	11142	11416	11567	11589	11480	11242	10877	10390	9791						9090	3736	2871	2092	1432	926	606	500
13	11158	11749	12229	12589	12820	12916	12875	12694	12376	11925	11349	10656						9860	3972	3039	2201	1494	953	613	500
14	12796	13372	13824	14144	14323	14356	14243	13982	13576	13033	12358	11565						10667	4217	3213	2315	1559	982	620	500
15	14570	15123	15538	15808	15925	15888	15694	15344	14844	14199	13420	12520						11514	4473	3394	2433	1626	1012	627	500
16	16484	17003	17372	17583	17630	17513	17229	16783	16179	<i>g</i>	13521	12399	11191	9918	8606	7285	5984	4738	3582	2555	1695	1042	635	500	
17	18539	19016	19328	19471	19439	19232	18850	18299	17583	16714	15704	14568	13324	11995	10605	9180	7751	6350	5013	3777	2681	1767	1074	642	500
18	20739	21163	21409	21474	<i>k</i>	20559	19893	19057	18064	16926	15661	14290	12834	11320	9776	8235	6730	5298	3978	2812	1841	1107	650	500	
19	23086	23447	23617	23594	23374	22960	22355	21566	20602	19476	18203	16803	15296	13706	12063	10395	8737	7123	5593	4187	2947	1917	1141	659	500
20	25583	25870	25954	25832	25504	24972	24241	23319	22218	20951	19536	17992	16342	14614	12835	11038	9257	7531	5899	4403	3087	1996	1176	667	500
21	28231	28433	28420	28190	27743	27083	26217	25154	23907	22491	20924	19230	17431	15556	13636	11704	9796	7953	6214	4625	3231	2078	1212	676	500
22	31034	31141	31020	30670	30094	29296	28285	27071	25669	24095	22370	20516	18561	16533	14465	12393	10354	8388	6540	4855	3380	2162	1249	685	500
23	33993	33993	33753	33273	32557	31611	30445	29071	27505	25764	23872	21852	19733	17546	15325	13106	10930	8838	6877	5092	3533	2248	1287	694	500
24	37111	36993	36622	36001	35135	34030	32699	31155	29416	27500	25432	23239	20949	18595	16214	13843	11525	9303	7224	5337	3691	2337	1326	703	500

Table 1: Own payoff-maximizing, spiteful, and altruistic strategies.

Subjects used a plain payoff table without tags or shades.

		2	
		p_2	$1-p_2$
		<i>Not commit</i>	<i>Commit</i>
1	<i>Not commit</i> p_1	7345 7345	8278 2658
	<i>Commit</i> $1-p_1$	8278 2658	706 706

Notes: The payoffs (7345, 7345) are based on the Nash equilibrium second stage investments of (8, 8) when neither commits to investing zero at the first stage (see Table 2). The payoffs (2658, 8278) (resp. (8278, 2658)) are based on the equilibrium investments of (0, 11) (resp. (11, 0)) when only player 1 (resp. player 2) commits. The payoffs are (706, 706) when both commit. The subgame perfect Nash equilibria of this initial stage commitment game are non-commitment probabilities for players 1 and 2 (p_1, p_2) of (1, 0), (0, 1), and (0.68, 0.68). The unique Evolutionary Stable Strategy (ESS) is $(p_1, p_2) = (0.68, 0.68)$.

Table 2. The payoff table becomes a Hawk-Dove game.

	Expectation	Punishably or Rivalistically Spiteful strategies (%)	Own payoff-maximizing strategies (%)	Altruistic strategies (%)
Only one commits: random choice (ratios in Table 1)		44.00 (=11/25)	4.00 (=1/25)	52.00 (=13/25)
Only one commits: Tsukuba with the “opponent” wording		69.23 (=27/39)	30.77 (=12/39)	0 (=0/39)
Only one commits: TMU with the “opponent” wording		68.09 (=32/47)	31.91 (=15/47)	0 (=0/47)
Only one commits: TMU without the “opponent” wording		66.00 (=33/50)	32.00 (=16/50)	2.00 (=1/50)
Only one commits: all sessions		67.65 (=92/136)	31.62 (=43/136)	0.74 (=1/136)
Neither commits: random choice (ratios in Table 1)		22.40 (=140/625)	4.16 (=26/625)	73.40 (=459/625)
Neither commits: Tsukuba with the “opponent” wording	Nash	50.00 (=96/192)	42.71 (=82/192)	7.29 (=14/192)
	Cournot	44.47 (=77/172)	45.35 (=78/172)	9.88 (=17/172)
	Average (Fictitious Play)	46.51 (=80/172)	45.35 (=78/172)	8.14 (=14/172)
Neither commits: TMU with the “opponent” wording	Nash	64.50 (=129/200)	23.50 (=47/200)	12.00 (=24/200)
	Cournot	60.56 (=109/180)	29.44 (=53/180)	10.00 (=18/180)
	Average (Fictitious Play)	63.33 (=114/180)	26.11 (=47/180)	10.56 (=19/180)
Neither commits: TMU without the “opponent” wording	Nash	41.67 (=80/192)	38.02 (=73/192)	20.31 (=39/129)
	Cournot	38.95 (=67/172)	31.98 (=55/172)	29.07 (=50/172)
	Average (Fictitious Play)	34.30 (=59/172)	34.88 (=60/172)	30.81 (=53/172)
Neither commits: all sessions	Nash	52.23 (=305/584)	34.59 (=202/583)	13.18 (=77/584)
	Cournot	48.28 (=253/524)	35.50 (=186/524)	16.22 (=85/524)
	Average (Fictitious Play)	48.28 (=253/524)	35.31 (=185/524)	16.41 (=86/524)

For Cournot and average (fictitious play) expectations, we exclude the first investment choice of each subject when neither commits.

Table 3. Ratios of Spiteful, Own Payoff-Maximizing, and Altruistic Strategies Chosen by Subjects.

		2	
		p_2 Not commit	$1-p_2$ Commit
1	Not commit p_1	6570 >	4795
	Commit $1-p_1$	2049 >	706

The table displays average payoff values for two players, 1 and 2, based on their commitment decisions. Player 1's strategies are 'Not commit' (p_1) and 'Commit' ($1-p_1$). Player 2's strategies are 'Not commit' (p_2) and 'Commit' ($1-p_2$). The values are color-coded: blue for Player 2's payoffs and red for Player 1's payoffs. Arrows indicate that Player 1's payoff is higher when they commit than when they do not, regardless of Player 2's choice.

Table 4. The average values of payoff data up to period 5 for the three sessions.

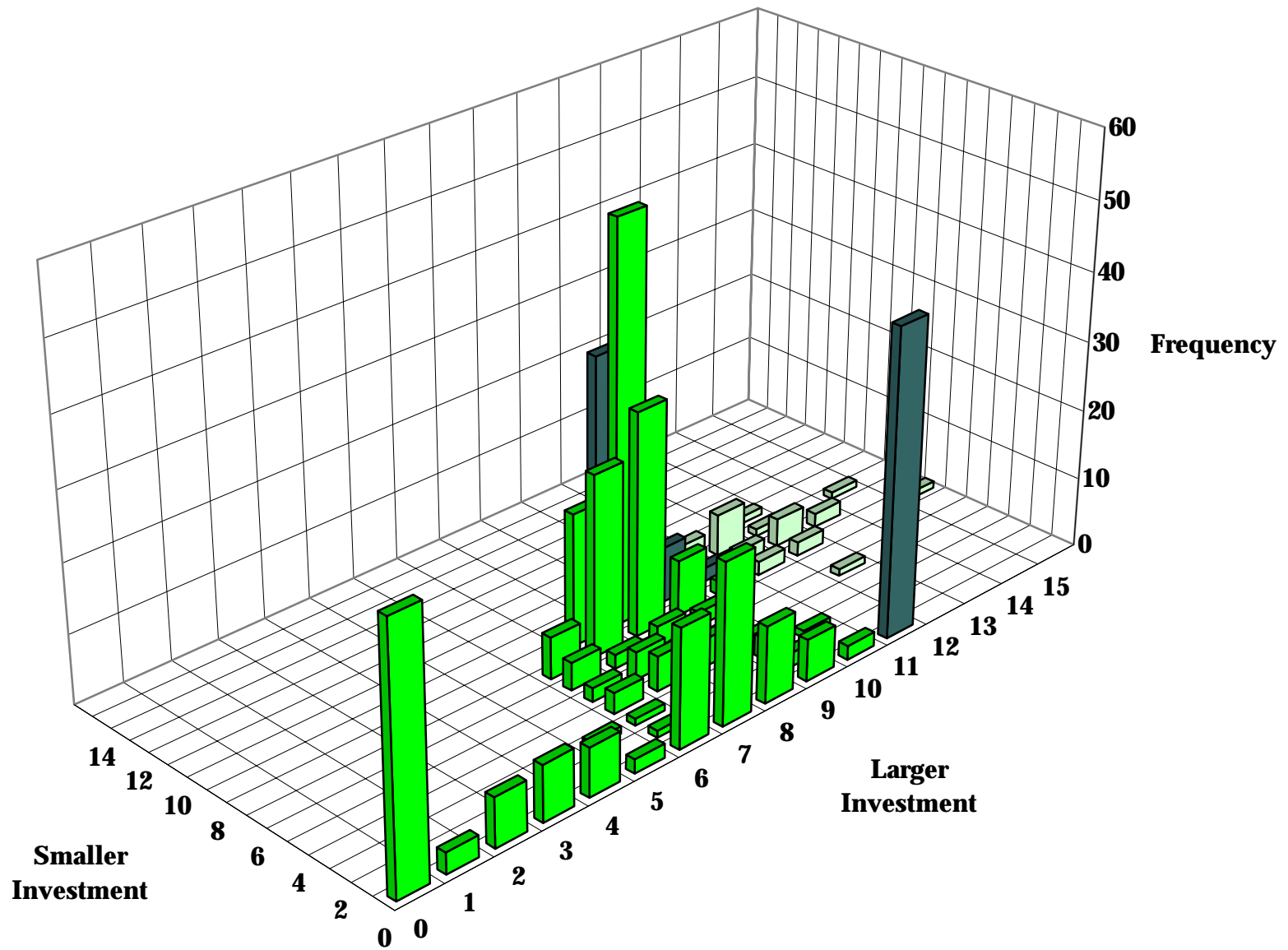


Figure 2. Investment Pattern in the Three Sessions.

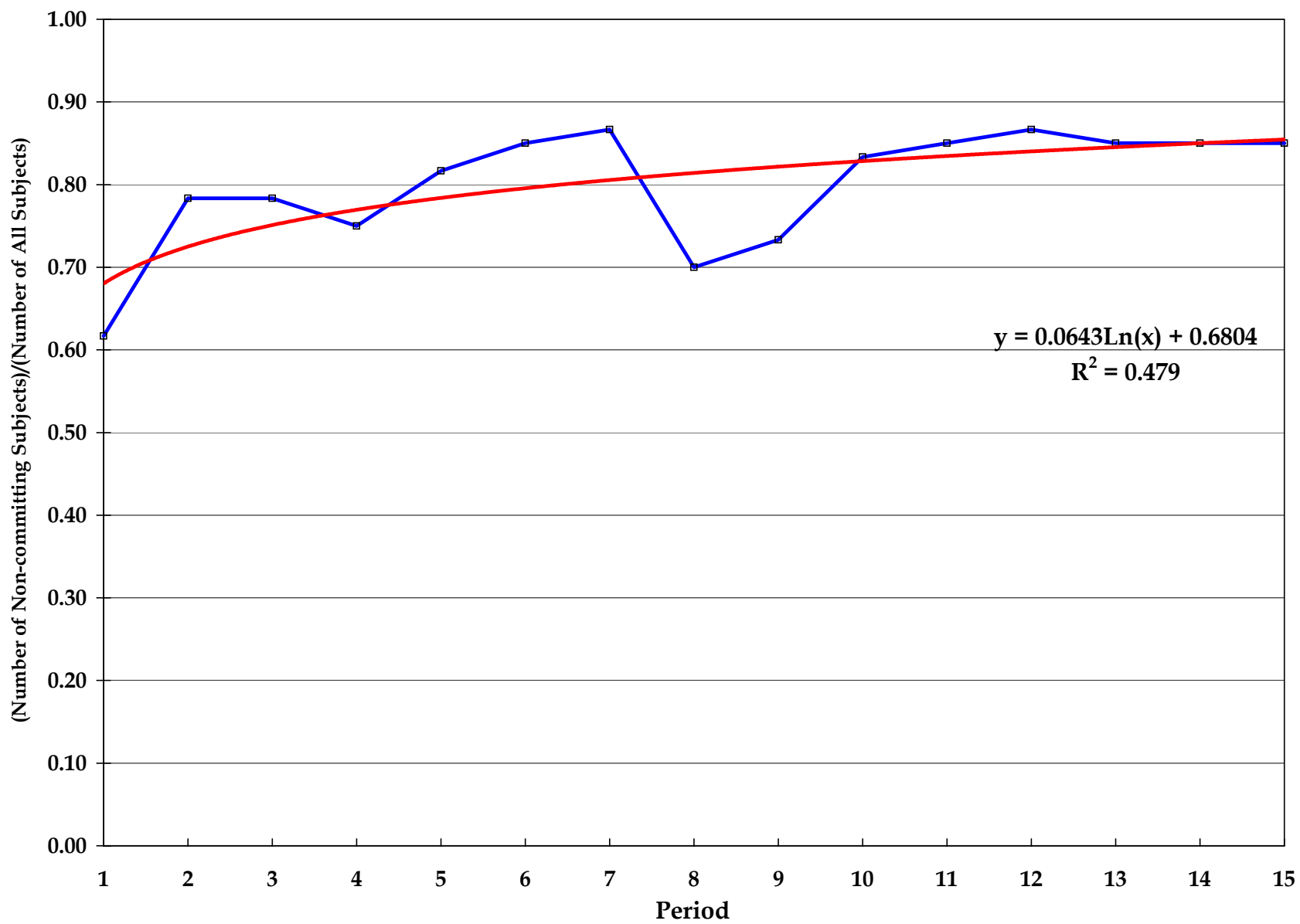


Figure 3. Non-commitment Rate Pattern.