

Non-Excludable Public Good Experiments¹

by

Tatsuyoshi Saijo*

Takehiko Yamato**

Konomu Yokotani***

and

Timothy N. Cason****

May 1997; revised October 2002

¹ We thank Takenori Inoki, Mamoru Kaneko, Hajime Miyazaki, Toru Mori, Mancur Olson, Mitsuo Suzuki, two anonymous referees, and participants at the Fall 1996 Economic Science Association conference 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 of the Ministry of Education, Science and Culture in Japan, and the Tokyo Center for Economic Research Grant.

* Institute of Social and Economic Research, Osaka University, Ibaraki, Osaka 567-0047, Japan, E-mail address: saijo@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: saijo@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

****Department of Economics, Krannert School of Management, Purdue University, West Lafayette, IN 47907-1310, USA. E-mail address: cason@mgmt.purdue.edu

Abstract

We conducted 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 selected 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 rounds 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.

Correspondent:

Until July 31, 2003

Tatsuyoshi Saijo
228-77 Humanities and Social Sciences
California Institute of Technology
Pasadena, CA 91125
USA
Phone 1-626-395-4216 (phone and fax)
E-mail: saijo@hss.caltech.edu

From August 1, 2003

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: saijo@iser.osaka-u.ac.jp

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) and 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 September 2002, 84 parties including the U.S.A. have signed the Protocol, and 93 parties have either ratified or acceded to it. In March 2001, however, President Bush announced not to 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 -- also 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.

¹ Another contemporary example is the chemical weapons convention (CWC). It took 24 years to reach agreement on the disposition of chemical weapons, and the number of countries signing the CWC was more than 160 by the end of 1995. The CWC entered into force in April 1997, following ratification by 65 signatories. The CWC is a mechanism creating a public good, i.e., a greater likelihood of world peace. However, 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. The League of Nations is also an example. 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.

Mechanism designers such as Groves and Ledyard (1977), Hurwicz (1979), Walker (1981) and Dutta, Sen and Vohra (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. Experimentalists also conducted experiments on the provision of public goods, and their main focus is on the voluntary contribution mechanism.² All subjects in the experiments must choose a number corresponding to the 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 a 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 explicitly incorporated 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 set up a model of 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 selected not to commit in the first stage choose contributions to the public good. In our experiments, we conducted two treatments. In the control Treatment A both subjects could *not* commit to contributing nothing. This treatment corresponds to the usual public good experiments that do not have the first stage,

² Another line of experimental research in public goods is to investigate 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).

and consist of the second stage only. In Treatment B each subject could choose her commitment decision before her contribution decision, Treatment B included both stages.

Second, we employed a *non-linear* payoff function rather than the linear payoff function used in most of the previous experiments.⁴ 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.

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 payoff tables of previous experiments, which we call *rough tables*.⁵ 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 in Treatment B is common knowledge between paired subjects, but no pair learns the commitment or contribution decisions of other pairs.

Fourth, we used two subjects rather than many subjects. 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 to the 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 rounds, 4 for practice and 15 for monetary reward, so as not to pair the same two subjects more than once.

Fifth, we asked subjects several questions such as, what is “your reason for your decision on your investment number?” in each round rather than only after the end of a

⁴ Andreoni (1993) used a Cobb-Douglas payoff function and a detailed payoff table that we will explain later. Isaac and Walker (1998) and Sefton and Steinberg (1996) used quasi-linear payoff functions with rough payoff tables. See Laury and Holt (2002) for a survey of these and other voluntary contribution game studies with interior Nash equilibria.

⁵ Saijo and Nakamura (1995) compared the effects of detailed payoff tables with rough payoff tables for a public goods environment with linear payoff functions.

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 round. Of course, these types of questions in each round 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.

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 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 strategy 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, almost every investment strategy less than the own payoff-maximizing strategy satisfies this condition.

When neither subject commits to the zero investment, all strategies can be basically classified into these three categories in the same manner.

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

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.

Nash equilibria and one mixed strategy Nash equilibrium which is the unique evolutionarily stable strategy (ESS) equilibrium.

We observed the following results. In Treatment A, subjects on average contributed (or "invested") close to but slightly below the interior Nash equilibrium level, because some subjects selected apparently spiteful investment strategies so as to reduce their opponent's payoff more than their own payoff reduction.⁷ Based upon different models of subject expectations regarding the contribution choices of the other subject, we found that 44-50% of choices are spiteful, 42-49% are own payoff-maximizing, and 8-12% are altruistic. For comparison, if subjects choose strategies randomly, 22.4% of choices are spiteful, 4.2% are own payoff-maximizing, and 73.4% are altruistic. That is, spiteful behavior prevailed among subjects who were not own payoff-maximizing. Moreover, the investment levels did not get closer to the Nash equilibrium level over time.

Our results are different from those in standard public goods experiments with linear payoff functions, where investments usually exceed the Nash equilibrium level in early periods but decline and approach the Nash equilibrium level over time. However, since investing zero is a dominant strategy with linear payoff functions, subjects can only contribute greater than or equal to the Nash equilibrium level of no investment. In other words, every strategy is either own payoff-maximizing or altruistic and no spiteful strategy is available with linear payoff functions. On the other hand, with Cobb-Douglas type payoff functions, own payoff-maximizing, spiteful, and altruistic strategies can all be selected. This is one reason why our results with non-linear payoff functions differ from those in standard public goods experiments with linear payoff functions.

In Treatment B, it is made explicit to subjects 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 in Treatment B that the non-commitment rate, which is the ratio of the number of non-committing subjects to the total number of subjects, rose as rounds

⁷ Saijo and Nakamura (1995) observed spiteful behavior in a traditional voluntary contribution mechanism experiment. See also Ito, Saijo, and Une (1995). Spiteful behavior is also observed in biology. Iwasa, Nakamaru, and Levin (1998) set up a model explaining spiteful behavior of colicin-producing bacteria against colicin-sensitive bacteria.

advanced. In other words, the free-riding rate declined over time. Moreover, the average investment in the final rounds in Treatment B was very close to that in Treatment A. Consequently, in the final two-thirds of Treatment B, 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 high payoff with free-riding. However, her opponent, subject 2 who decided not to commit to investing zero, did not choose the own payoff-maximizing investment strategy. Instead, subject 2 often took a *spiteful* strategy by investing a smaller amount so as to reduce subject 1's payoff more than subject 2's payoff reduction. In fact, 68.4% of investment strategies chosen when the opponents commit to investing zero were spiteful, 30.9% are own payoff-maximizing, and 0.7% are altruistic. This spiteful behavior is unusual since 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 the following rounds. Rather, it could have had only an *indirect* effect.

The committing subject learned and expected that commitment to investing zero was not beneficial to her because of the spiteful behavior 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 mechanisms with and without pre-commitment to contributing zero. Section 3 describes the experimental design. We present the results of the experiment on the voluntary contribution mechanism without pre-commitment in Section 4 and those of the experiment on the voluntary contribution mechanism with pre-commitment in Section 5. Section 6 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). From the investment, each subject receives $y = y_a + y_b + w_y$, where w_y is the initial level of the public good. That is, the level of the public good is the sum of the investments of two subjects and the initial level of the public good. Therefore, each subject's decision problem is

$$\max u_i(x_i, y) \text{ subject to } x_i + y_i = w_i,$$

where $u_i(x_i, y)$ is subject i 's payoff function. We use a Cobb-Douglas type payoff function to transform contributions and the private good into subject payoffs, and all subjects have the same payoff function. That is, $u_i(x_i, y) = x_i^\alpha y^{1-\alpha}$, where $\alpha \in (0, 1)$. Using a monotonic transformation, we specify the payoff function as follows:

$$(1) \quad u_i(x_i, y) = \frac{\{x_i^\alpha y^{1-\alpha}\}^\beta}{50} + 500.$$

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 payoff is $u_i(\hat{x}_i, \hat{y}) = 7089$, where $\hat{x}_i = 24 - 7.69 = 16.31$ and $\hat{y} = \hat{y}_a + \hat{y}_b + w_y = 18.38$. The Pareto efficient level of the public good is determined uniquely by the Samuelson condition and the feasibility condition. Its symmetric contribution level is 12.02. Therefore, the Pareto efficient level of the public good is $27.04 = 12.02 + 12.02 + 3$. Clearly, the level of the public good with the voluntary contribution mechanism is less than the Pareto efficient level of the public good.

In our experiment, subjects choose integer investment numbers only. The detailed payoff table used in the experiment is Table 1: the rows are for the subject's own investment numbers and the columns are for the other subject's investment numbers. If your opponent's investment number is 8, your own payoff is maximized by investing 8. The same thing holds for your opponent who has the same payoff table. In other words, the Nash equilibrium of this game is for each subject to contribute 8 and obtain a payoff of 7345. No other Nash equilibria sneak into our model due to the discrete strategy choice set.

 Table 1 is around here

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

The Nash 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, it has been observed in many experiments on the voluntary contribution mechanism that subjects select different types of strategies such as altruistic and spiteful ones. Therefore, it is useful to introduce formal definitions of spiteful strategies as well as altruistic strategies in the voluntary contribution mechanism. We regard own payoff-maximizing behavior as a standard of comparison, and then classify all possible strategies into three regions.

(i) 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.

(ii) 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.

(iii) 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.

Given any expected value of the other subject's investment number, every possible investment strategy that a subject can choose is regarded as one of three strategies mentioned above. Table 2 illustrates how strategies are classified in our setting. For example, suppose that a subject expects her opponent to choose 8. Then:

(i) Investing 8 is an own payoff-maximizing strategy. In this case, both subjects get 7345 (see the number in cell *c* of Table 2 when the choice of the subject is 8 and the choice of her opponent is 8).

(ii) Investing less than 8 is a spiteful strategy. For instance, if the subject invests 6 instead of 8, her own payoff is reduced from 7345 to 7237 (the number in cell *d* of Table 2 when the choice of the subject is 6 and the choice of her opponent is 8), and the payoff of her opponent is also reduced from 7345 to 5766 (the number in cell *e*).

(iii) Investing more than 8 is an altruistic strategy. For example, by investing 16 instead of 8, her own payoff is reduced from 7345 to 4179 (the number in cell *f*), while the payoff of the other subject increases from 7345 to 16179 (the number in 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 2 the own payoff-maximizing strategy region represents 4.16% (=26/625) of the cells, the spiteful strategy region is 22.40% (=140/625) of the cells, and the altruistic strategy region is 73.4% (=459/625) of the cells.

 Table 2 is around here

Moreover, spiteful strategies can be divided in two categories. A spiteful strategy that causes a reduction of your opponent's payoff that is smaller than one's own payoff reduction is called a *weakly spiteful strategy*. For example, when the expected investment number of your opponent is 18, by investing 1 instead of 3, you can reduce the payoff of your opponent from 2027 (cell *h* of Table 2) to 1743 (cell *i*). But the reduction of your opponent's payoff, 284=2027-1743, is smaller than your own payoff reduction, 311=21474-21163 (compare cell *j* and cell *k*). In Table 2, however, the number of such weakly spiteful strategies is only 12. For the other 128 spiteful strategies, the reduction of the opponent's payoff is larger than the own payoff reduction.

Whether an investment choice of a subject is own payoff-maximizing, spiteful, or altruistic depends on how much she expects the other subject to invest. We will consider several possible expectations in Sections 4 and 5 when classifying behavior into spiteful, altruistic and own payoff-maximizing categories.

2.3. Linear versus Non-Linear Payoff Functions

Both spiteful and altruistic strategies can be selected in our model with non-linear Cobb-Douglas type payoff functions. However, this is not the case when payoff functions are linear: only own payoff-maximizing and altruistic strategies can be chosen and no spiteful strategy is available.

Suppose that two subjects have the same linear payoff function:

$$u_i = x_i + \beta y = w_i - y_i + \beta(y_i + y_j) = w_i - (1 - \beta)y_i + \beta y_j, \text{ where } 0 < \beta < 1 \text{ and } i, j = a, b; i \neq j.$$

Figure 1-(a) illustrates how strategies are classified with linear indifference curves, where the horizontal axis is subject *i*'s consumption of the private good, $x_i = w_i - y_i$, and the vertical axis

is the level of the public good, $y = y_i + y_j$. It is well-known that investing zero is a dominant strategy, that is, given any investment level of the other subject, $y_j \in [0, w_j]$, $y_i = 0$ ($x_i = w_i$) is subject i 's own payoff-maximizing strategy subject to the constraint $y = w_i - x_i + y_j$. Also, it is easy to check that contributing any positive amount, $y_i \in (0, w_i]$, is an altruistic strategy. Note that there is no spiteful strategy available in this case.

On the other hand, both spiteful and altruistic strategies are available when payoff functions are non-linear as assumed in our model. Figure 1-(b) illustrates the classification of strategies with Cobb-Douglas type indifference curves, which is a diagrammatic representation of Table 2. Investing less than an own payoff-maximizing strategy is a spiteful strategy, and investing more than an own payoff-maximizing strategy is an altruistic strategy, given any investment level of the other subject.

 Figure 1-(a) and 1-(b) are around here

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

So far 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) found 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 2). 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 knowing the other subject's commitment decision.

 Figure 2 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.

In our experiment, subjects choose integer investment numbers only. If neither of subjects decide 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. 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, and the subject who chose commitment clearly invests nothing and obtains a payoff of 8278. If both subjects choose to commit to investing zero, both end up with a payoff of 706. These subgame perfect payoffs are incorporated into the normal form game payoff table shown in Table 3.

Table 3 is around here

The game in Table 3 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.⁸

2.5. Own Payoff-Maximizing, Spiteful, and Altruistic Strategies

For the two-stage game with pre-commitment to contributing nothing, we define own payoff-maximizing, spiteful, and altruistic investment strategies at the second stage in the same way as in Section 2.2 for the usual voluntary contribution mechanism.

When one subject decides not to commit to investing zero, but the other subject decides to commit to investing zero, the strategies of the former non-committing subject are classified as follows:

⁸ 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.

(i) Investing 11 is an own payoff-maximizing strategy. In this case, the non-committing subject gets 2658 (see the number in cell ℓ of Table 2), while the other subject earns 8278 (the number in cell m).

(ii) Investing less than 11 is a spiteful strategy. For example, if the non-committing subject invests 7 instead of 11, her own payoff is reduced from 2658 to 2210 (the number in cell n), and the payoff of the other subject is also reduced from 8278 to 4018 (the number in cell p).

(iii) Investing more than 11 is an altruistic strategy. For instance, by investing 17 instead of 11, her own payoff is reduced from 2658 to 1871 (the number in cell q), while the payoff of the other subject increases from 8278 to 18539 (the number in cell r).

When no subject pre-commits to contribute nothing, the second stage game corresponds to the usual voluntary contribution mechanism game, and thus the definitions of spiteful and altruistic strategies are the same as those defined in Section 2.2.

The subgame perfect equilibrium analysis in the previous subsection 2.4 is based on the assumption that each subject chooses an own payoff-maximizing investment strategy. However, we observed different types of strategies in our experiments, in particular spiteful strategies. As we discuss in the results, this dramatically changes the nature of the first stage commitment decision game.

3. Experimental Design

Our experiment consisted of three treatments, A, B, and B'. Treatment A corresponded to the voluntary contribution mechanism without pre-commitment to investing zero and Treatments B and B' corresponded to the voluntary contribution mechanism with a pre-commitment decision. We will explain the difference between Treatment B and Treatment B' below.

We conducted one session in Treatment A and one session in Treatment B at the University of Tsukuba during December of 1995, one session in Treatment B and one session in Treatment B' at the Tokyo Metropolitan University during December of 1997, and one session in Treatment A at Tokyo Institute of Technology during December of 2001. All sessions employed 20 subjects, so the five sessions reported here employed a total of 100 different student subjects. We recruited these students for each session 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. The Treatment A session required approximately 90 minutes and the Treatment B and B' sessions required approximately two hours to complete. The mean payoff per subject was \$28.38 (\$1=100 yen). The maximum payoff among the one hundred subjects was \$45.53, and the minimum payoff was \$17.28.⁹

Let us describe Treatment A first. Twenty subjects seated at desks in a relatively large room had random identification numbers. These identification numbers were not publicly displayed, however, so subjects could not determine who had which number. We made ten pairs out of twenty subjects. In each round, ten pairs played the game without 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 rounds were for practice and the remaining fifteen determined the subjects' monetary payoffs. Each subject received an experimental procedure sheet, a record sheet, payoff tables, 15 investment sheets, and 4 practice investment sheets.¹⁰ Instructions were given by tape recorder to minimize the interaction between subjects and experimenters.

Each subject determined her investment by selecting an integer number between 0 and 24 by using payoff tables and then marked this number on an investment sheet. Experimenters collected these investment sheets and then redistributed them only to the paired subjects. No information (such as the sum of investment numbers) 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 round started.

Treatment B had one additional step. Before choosing investment numbers, 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

⁹ We paid higher earnings to the subjects at Tokyo Institute of Technology than at the other two universities, because Tokyo Institute of Technology is located at the center of Tokyo, while the others at local towns.

¹⁰ The use of practice periods permits subjects to send signals costlessly prior to the paid rounds. However, subjects interacted with each other in only one period and there was no public information after the instructions were completed and data collection began. Therefore, the impact of these signals is somewhat limited. Nevertheless, see footnote 18 for some minor evidence of spillover from the practice periods.

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.

Treatment B' is exactly the same as Treatment B except for the term that was used to describe the person that each subject is matched with at each period. The term "your opponent" was employed in the instructions, record sheets and payoff tables for both Treatment A and Treatment B. The phrase "the person you are paired with" replaced "your opponent" in all materials of Treatment B'. One might say that the term "opponent" forces subjects to think in relative terms. However, there was no essential difference between the results for Treatment B and those for Treatment B', as we document below.

Every subject had the same payoff function and every subject knew this fact. We distributed three kinds of payoff tables to minimize the likelihood of any possible misunderstanding, and all of the one hundred subjects were able to readily calculate their payoffs following the instructions and practice periods. The payoff tables used in Treatments A and B are Tables 1, 4, and 5. Table 1 is a detailed payoff table. We also presented a payoff table summarizing average payoffs for sets of 9 or 12 payoff cells shown in Table 4, as well as an iso-payoff map of Table 5. The payoff tables distributed in Treatment B' are the same as Tables 1, 4, and 5 except that the term "your opponent" is replaced by the term "the person you are paired with". Most subjects indicated in their post-experiment questionnaire that they understood all three kinds of payoff tables and used the detailed payoff table (Table 1) only. We allowed subjects three minutes to examine the three payoff tables before the practice rounds and ten minutes to examine the three new payoff tables before the real rounds. The tables used for practice and real rounds were different.

Tables 4 and 5 are around here

4. Experimental Results for Treatment A

The session in Treatment A conducted at the University of Tsukuba and the session in Treatment A at Tokyo Institute of Technology (TIT) are called the Tsukuba A session and TIT A session, respectively. The data from these two sessions are statistically indistinguishable in virtually all rounds (see the Appendix). This provides strong evidence that the experiment site (Tsukuba versus Tokyo) does not affect choices. Thus, in the subsequent analysis we pool the data across the two Treatment A sessions.

Treatment A was a control treatment intended to permit a comparison between without and with pre-commitment. The Nash equilibrium investment pair was (8, 8), and no other Nash equilibrium exists. Since each round had 10 pairs and 15 rounds were conducted at each of the two treatment A sessions, we have 300 pairs of data. The order of investment numbers does not matter, so we rearranged each pair (x,y) with $x \geq y$. Figure 3 shows the frequency distribution of investment data. The maximum frequency pair was (8, 7) with 73 pairs, the second was (8, 8) with 69 pairs, the third was (8,6) with 26 pairs, and the fourth was (7, 7) with 18 pairs. The average investment across all forty subjects was 7.31.

Figure 3 is around here

Figure 4 shows the average investment pattern from round 1 to 15. The average investment was less than the Nash equilibrium level of investment (8) in all rounds. In order to understand why this happened, we checked the record sheets and questionnaire sheets of subjects. The record sheet was not only for keeping the investment record but also for specifying the reasons why a subject chose her investment numbers. We found eleven subjects who explicitly stated the following reasoning: they estimated that their opponents would chose 8, and then they chose 6 or 7 because this would make their opponents' payoffs much lower than their own payoffs. For example, consider the case that subject a chooses 7 and subject b chooses 8. Then subject a obtains 7340, and subject b obtains 6526. At the Nash equilibrium (8, 8), both subjects obtain 7345. From the viewpoint of subject a , the reduction of 5 units of payoff is minor, but subject a can make the reduction of subject b 's payoff (819) much greater than his

own reduction.¹¹ In other words, these subjects took spiteful strategies according to our definition in Section 2.2. We label these subjects as *spiteful* subjects.¹²

Figure 4 is around here

We now examine how frequently own payoff-maximizing, spiteful, or altruistic behavior (defined in Section 2.2) occurs in the data. 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 realized number in the previous period.

3) *An average expectation with a declining weight (fictitious play)*: each subject has a long memory when she decides an investment number, in contrast to the Cournot expectation case in which each subject has a short memory length. We consider the following expectation with a declining weight. Let $EI(t)$ be the expected investment number of the other subject at period t

¹¹ Akerlof and Yellen (1985a, 1985b) observed similar phenomena in Keynesian business cycles and industrial organization theory. A small amount of nonmaximizing behavior may cause some amount of changes at equilibrium that is larger in its magnitude than the losses due to nonmaximizing agents.

¹² Those spiteful subjects are subjects #1, #7, #14, and #19 at the University of Tsukuba and subjects #1, #3, #9, #11, #16, #17, and #19 at Tokyo Institute of Technology. They chose spiteful strategies at most periods. Their specific investment patterns were as follows:

Tsukuba: a) Subject #1 invested 6 or 7 (less than 8) to spite their opponents in eleven periods, and invested 8 in four periods; b) Subject #7 invested 3 in all fifteen periods; c) Subject #14 invested 4 or 7 (less than 8) in nine periods, and invested 8 in six periods; and d) Subject #19 invested 3, 4, 5, or 6 (less than 8) in six periods, invested 8 in five periods, and invested 9, 10, or 12 (more than 8) in four periods. The first three subjects selected spiteful strategies in most periods. The last subject tried large investment numbers until period 7. After period 8, however, he adopted spiteful strategies in most rounds: investing less than 8 six times, but investing 8 only two times.

Tokyo Institute of Technology: a) Subject #1 invested 6 or 7 (less than 8) in thirteen periods, and invested 9 or 24 (more than 8) in two periods; b) Subject #3 invested 6 or 7 (less than 8) in ten periods, and invested 8 in five periods; c) Subject #9 invested 3, 6 or 7 (less than 8) in all fifteen periods; Subject #11 invested 4, 5, 6, or 7 (less than 8) in eleven periods, invested 8 in three periods, and invested 10 (more than 8) in one period; e) Subject #16 invested 0, 2, 4, or 5 (less than 8) in thirteen periods, invested 8 in one period, and invested 24 (more than 8) in one period; f) Subject #17 invested 6 or 7 (less than 8) in ten periods, invested 8 in three periods, and invested 9 or 11 (more than 8) in two periods; and g) Subject #19 invested 6 or 7 (less than 8) in eleven periods, invested 8 in three periods, and invested 9 (more than 8) in one period. All subjects played spiteful strategies in at least ten periods.

and $I(t)$ be the investment number the other subject chose at period t . Then the next period's expected investment is given by $EI(t+1) = [I(t) + t \times EI(t)] / (t+1)$. This is simply a recursive version of fictitious play expectations.

Table 6 shows the rates of spiteful, own payoff-maximizing, and altruistic strategies chosen by subjects at Treatment A for each type of expectation. For all three types of expectations, most data fall into the spiteful strategy region and the own payoff-maximizing strategy region, and only a few fall in the altruistic strategy region. In particular, if subjects chose strategies randomly, then the rate of altruistic strategies would be 73.40% (the fraction of altruistic choices shown in payoff table 2). However, the actual rates in Table 6 were less than 12%.

 Table 6 is around here

These spiteful strategy choices tended to reduce average investments below the equilibrium level. We tested the hypothesis that mean investment equals the Nash equilibrium investment (8) first by pooling investments across rounds. Because each subject made 15 investment choices the data are clearly not independent. Therefore, we accounted for the panel nature of the data using a random effects error specification $v_{it} = e_i + \varepsilon_{it}$, where e_i is a subject-specific error term and ε_{it} is an iid error. The pooled data strongly reject the Nash equilibrium ($t=4.47$). We then focused on round by round tests of the Nash equilibrium. A nonparametric Wilcoxon signed rank test rejects the Nash equilibrium hypothesis at the five percent significance level in fourteen out of fifteen rounds (all except round 6). Nevertheless, the fact that many subjects chose the Nash equilibrium investment and the mean investments are only slightly below 8 indicates that at least economically the Nash equilibrium has attracting power for the data.

These Treatment A results lead to the following observation.

Observation 1:

The mean investment across all rounds is significantly less than the Nash equilibrium investment. Overall, mean investments were approximately ten percent below the Nash equilibrium.

The above 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 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 observation. Laury and Holt (2002) 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, only own payoff-maximizing and altruistic strategies can be chosen and no spiteful strategy is available with linear payoff functions, as pointed out in Section 2.3. On the other hand, with Cobb-Douglas type payoff functions, all three kinds of strategies, own payoff-maximizing, spiteful, and altruistic strategies, can be selected.

5. Experimental Results for Treatments B and B’

The session in Treatment B conducted at the University of Tsukuba, the session in Treatment B at the Tokyo Metropolitan University (TMU), and the session in Treatment B’ at the Tokyo Metropolitan University are called the Tsukuba B session, TMU B session, and TMU B’ session, respectively. As shown in the appendix, the data from these three sessions are statistically indistinguishable in virtually all rounds. This provides strong evidence that neither the experiment site (Tokyo versus Tsukuba) nor the experiment wording (“your opponent” versus “the person you are paired with”) affect choices. Therefore, in the subsequent analysis we pool the data across the three Treatment B sessions, i.e., Tsukuba B, TMU B, and TMU B’ sessions.

5.1 Investment Data

Figure 5 shows that the distribution of investment pairs in Treatment B is very different from that of Treatment A (Figure 3). 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 5 is around here

Also unlike Treatment A, the average investment (including the investments of zero by subjects who committed to investing zero) changed across rounds in Treatment B. Figure 6 illustrates that the average investment in Treatment B was clearly less than in Treatment A. However, the average investment in Treatment B ascended across rounds, and in the late rounds the average investments in the two Treatments were similar; for example, in the final round the average investments in Treatments A and B were 7.03 and 6.47 respectively. According to the nonparametric two-sample Wilcoxon rank sum test, the average investments are significantly higher in Treatment A than in Treatment B at the 5-percent level in periods 1, 3, 4, 5, 8, 9, 10 and 11. Although this is a majority of the periods (8 of 15), note that the investments are not significantly different in the A and B treatments in the later periods 12 through 15.

Figure 6 is around here

Using a random-effects panel data model that is the same as that for Treatment A, we tested the hypothesis that mean investment equals the Nash equilibrium investment (8) in the case of neither committing to investing zero (the number of observations for this case is 583). The mean investment in this case was too low (7.26), and we rejected the hypothesis ($t=4.21$).¹³ This is because several subjects exhibited spiteful behavior as in Treatment A. The rates of spiteful, own payoff-maximizing, and altruistic strategies for the case of neither committing at Treatment B are shown in Table 6. Again, most data fall into the spiteful region and the own

¹³ For those who prefer nonparametric tests, we also conducted Wilcoxon signed rank tests separately for each of the 15 periods. This test rejects the null hypothesis that average investment equals 8 at the five percent level in 12 of the 15 periods.

payoff-maximizing point, and few in the altruistic region for each of three types of expectations.¹⁴

However, like the result for Treatment A, a substantial number of subjects selected the Nash equilibrium, so the Nash equilibrium is meaningful at least economically in the case of neither committing.

5.2 Commitment Data

We next examined whether the commitment data in Treatment B were compatible with the mixed strategy equilibrium of the Hawk-Dove game described in Table 3. The null hypothesis is the ESS non-commitment probability of 0.68. We first conducted a binomial test separately by round in order to avoid pooling dependent 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 7 shows, the non-commitment rate rose as rounds advanced (although a brief decline was observed in rounds 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 rounds of the 15 total rounds (in rounds 5, 6, 7, 10, 11, 12, 13, 14, and 15) at the five percent (usually one percent) significance level.

Figure 7 is around here

We also examined the overall non-commitment rates for each of 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

¹⁴ The average expectation with a declining weight for Treatment B is defined as follows. 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 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 the other subject chose 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 .

more. 15 of the 60 subjects (25%) were apparently using a pure strategy, 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.

Furthermore, we 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, 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.

5.3 Why were Non-commitment Rates High?

In order to understand these rather 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 2). There were 139 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 95 times (zero 17 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), and invested more than eleven (twenty-four) only 1 time. In fact, the mean investment (6.90) was lower than when neither committed! Using a random effects model, we soundly reject the hypothesis that mean investment equals 11 in the case of one subject choosing non-commitment ($t=8.14$).¹⁵

As shown in Table 3, 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

¹⁵ 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.

(448=2658-2210) was relatively small, while the payoff reduction of the committing subject (4269=8278-4018) was relatively large. This appears to be stronger evidence of spiteful behavior than we observed in Treatment A. In fact, for this subgame we observed that over two-thirds of the data fall into the spiteful region, and less than one-percent in the altruistic region, as shown on the bottom of Table 6.

One might say that in the two-stage game of Treatment B, 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 spite rate of the final period is equal to $6/9 = 67$ percent.¹⁶

5.4 Commitment versus Non-commitment in Actual Payoffs

Figure 8 shows average payoffs by round depending on commitment. Curve α denotes the average payoffs when neither subject committed to investing zero. As Table 3 shows, the payoff should be 7345 according to the Nash equilibrium prediction, but actual data showed that average payoffs were less than 7000 units in all 15 rounds. Curve β denotes the average payoffs when only the other did not commit to investing zero. According to the Nash equilibrium, this should lead to 8278 units of payoff, but actual payoffs were less than 6000 in 12 of 15 rounds. Even more striking it that according to Table 3, curve β should be above curve α , but Figure 8 shows that curve α was above curve β in all but round 9. These average payoff curves 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).¹⁷ Table 7 illustrates the payoff matrix for commitment decisions, which is based on the average values of payoffs subjects actually obtained up to round 5, rather than the subgame equilibrium payoffs as shown in Table 3. As this table illustrates, non-commitment became a dominant strategy even in early rounds.

¹⁶ The full distribution of the 9 final period observations with exactly one non-committing subject is as follows: one subject invested two, one subject invested seven, three subjects invested eight, one subject invested nine, and three subjects invested eleven.

¹⁷ 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.

Figure 8 and Table 7 are around here

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 rounds and frequently earned more than 6000.¹⁸

5.5 Learning Processes for Non-commitment

Essentially, subjects learn (contrary to the ESS equilibrium) that expected profits from non-commitment tend to exceed those from commitment. In this subsection we provide a simple model of 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 report results of 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):

¹⁸ As Figure 11 in the appendix shows, the average non-commitment rates at early rounds in the TMU B session were relatively high. To see why this happened, we looked at the record sheets for the four practice periods, since the experiences of subjects in the practice periods might affect their commitment decisions in the actual experiment. We found that spiteful behavior during the practice periods would lead to high non-commitment rates in early periods. First of all, we counted the number of subjects who explicitly wrote that they realized that commitment was not beneficial because the non-committing subject would act spitefully in the practice periods. This number in the TMU B session (8) was much larger than both that in the TMU B' session (1) and that in the Tsukuba B session (1). We also calculated the number of subjects who had no chance to experience spiteful behavior in the practice periods. Consider a subject who experienced the following: (a) no cases in which only one subject did not commit, or (b) if only one subject did not commit (either himself or his opponent), the payoff-maximizing (non-spiteful) investment was selected. Call such a subject a "spite-free" subject. The number of spite-free subjects in the TMU B session was 5, the number in the TMU B' session was 8, and at the number in the Tsukuba B session was 11. As the number of spite-free subjects was larger, the non-commitment rate at period 1 was smaller.

$$(2) \quad \text{Probability(Non-commitment)} = f(\text{ENCE/ECE}),$$

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, and so they do not affect expectations. 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 update profit expectations following a commitment or a non-commitment choice, and we use two polar cases for 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 round, and they did not commit in this round, they update ENCE as follows:

$$(3) \quad \text{ENCE} = ((N * \text{previous ENCE}) + \text{current non-commitment earnings}) / (N + 1).$$

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.¹⁹

¹⁹ 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 round 1.

The empirical model is probit maximum likelihood, with a random subject effect. Table 8 presents the estimation results separately and pooled for the three Treatment B sessions. For Fictitious Play, the expected payoff ratio is insignificantly different from zero except in the TMU B session, where the coefficient estimate has the wrong sign. For the Cournot specification, the ratio is significantly positive except in the TMU B 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.

 Table 8 is around here

Summarizing the above observations, we have the following.

Observation 2:

- (a) The ESS prediction regarding the non-commitment rate is rejected.*
- (b) The non-commitment rate rises as rounds advanced and the average investment in Treatment B is not significantly different from that in Treatment A after period 12.*
- (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.*

5.6 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 (for example, see Rabin (1993), Levine (1998), Dufwenberg and Kirchsteiger (2002), Falk and Fischbacher (1998), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002) and 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.²⁰ 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 go as far as to 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.²¹ 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. However, 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 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

²⁰ 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 $1 > \beta_i \geq 0$.

²¹ 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 $\{0.3, 0.3, 0.3, 0.1\}$.

allocations and not on players' intentions. Rabin's (1993) fairness equilibrium, Dufwenberg and Kirchsteiger's (2002) 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.²² For example, note from Table 6 that the rate of spiteful strategy choices increases from about 50 percent to about 70 percent when subjects must make an investment that benefits a subject who has committed to an investment of zero.

6. Concluding Remarks

We found that the ESS prediction in a Hawk-Dove game was rejected and the number of non-committing subjects increased across time. Furthermore, this increase did not arise due to altruism or kindness among subjects, but from *spiteful* behavior of subjects. 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 (Ochs and Roth, 1989 and Prasnikar and Roth, 1992).

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),²³ 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

²² This is different from simple rivalistic behavior in which a subject seeks merely to earn more than his opponent earns. The data are inconsistent with such rivalistic motivations, because such motivations are more likely to lead to *less* non-commitment than the ESS prediction since commitment guarantees payoffs greater than or equal to those of the opponent. It is also unlikely that rivalistic choices would be so close to the Nash equilibrium when neither subject commits because rivalistic subjects have a strong incentive to reduce their public good investment.

²³ "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." (David Hume, 1739)

finding suggests a need to rethink our fundamental assumptions of human nature underlying our models.

Isaac and Walker (1998) also conducted public good provision experiments with interior Nash equilibria and examined how the location of the aggregate Nash equilibrium, relative to the group's aggregate endowment, influences contributions to a public good. They observed that (i) when the equilibrium was low relative to the endowment (19.4%), contributions were greater than the equilibrium; and (ii) when the equilibrium was in the middle of the endowment space (50%), contributions tracked the equilibrium at initial and middle rounds, but declined below it by final rounds. For our Treatment A and Treatment B when neither commits, the ratio of the individual Nash equilibrium level to the individual endowment was 33.3% ($=8/24$). This is between Isaac and Walker's "low Nash" treatment (19.4%) and "middle Nash" treatment (50%), but neither of their treatment results was similar to ours. One of the reasons for this difference might be how payoff information was given to subjects. Isaac and Walker provided each subject a rough payoff table, from which it was not easy to see the payoffs of the other subjects. On the other hand, we gave each subject a detailed payoff table to understand the payoff structure of the other subject, which might trigger spiteful behavior. In future experiments we plan to conduct a systematic exploration of this information effect.

It is also possible that spiteful behavior is more likely in two-person games such as this because it encourages relative payoff comparison and because it permits punishments that can be directed to specific individuals. Future experiments can test this conjecture using larger groups. Finally, the ESS non-commitment rate in the present experiment was 0.68. We are currently conducting experiments in which the rate is less than 0.5.

APPENDIX:

1. Comparisons of Tsukuba A and TIT A Sessions

Figure 9 illustrates how the average investment, which equals the sum of investment numbers divided by the number of subjects, changed from round 1 to 15 for each of the two A sessions. According to the nonparametric two-sample Wilcoxon rank sum test, the investments in the two A sessions are statistically different in only one of the 15 periods at the five percent significance level.²⁴

Figure 9 is around here

2. Comparisons of Tsukuba B, TMU B, and TMU B' Sessions

First of all, we compare investments conditional on non-commitment across the three B and B' sessions. Figure 10 illustrates how the average investment conditional on non-commitment, which equals the sum of investment numbers for non-committing subjects divided by the number of non-committing subjects, changed from round 1 to 15 for each of the three B and B' sessions. The average investments conditional on non-commitment were different in only a handful of periods according to the nonparametric Wilcoxon rank sum test. In 14 of the 15 periods there is no statistical difference between the TMU B and TMU B' sessions at the five percent significance level, which indicates that the "opponent" wording did not affect investments conditional on non-commitment.²⁵ Likewise, the TMU B and Tsukuba B sessions are not statistically different in 12 of the 15 periods, indicating that investments conditional on non-commitment differed very little across sites.²⁶

Figure 10 is around here

Figure 11 illustrates the average non-commitment rate patterns from round 1 to 15 in Tsukuba B session, TMU B session, and TMU B' session. There are some statistically significant differences in the non-commitment rates between the Tsukuba B and TMU B

²⁴ In period 11 the average investment in the TIT A session was significantly higher than in the Tsukuba A session.

²⁵ In period 8 the average investment in the TMU B' session was significantly higher than in the TMU B session.

²⁶ The average investment was significantly higher in the Tsukuba B session in periods 1, 9 and 13.

sessions in periods 1 and 2.²⁷ After period 2, there are basically no systematic differences among the non-commitment rates across sessions.²⁸

Figure 11 is around here

Overall, we conclude that there are virtually no statistically significant differences in the data based on the experiment site or the experiment wording.

²⁷ Fisher's exact test p -values for the TMU B versus Tsukuba B comparison are 0.022 in period 1 and 0.020 in period 2. We discuss the reasons for these differences in footnote 18. Fisher's exact test uses the hypergeometric probability distribution to calculate the exact probability of observing the distribution of non-commitment rates (and those more unequally distributed) under the null hypothesis of no differences in rates across treatments.

²⁸ In period 5 the non-commitment rate is significantly higher in the TMU B' session than in the TMU B session, but in period 10 the non-commitment rate is significantly lower in the TMU B' session than in the TMU B session.

REFERENCES

- AKERLOF, G. A., AND J. L. YELLEN (1985a): "A Near-Rational Model of the Business Cycle, with Wage and Price Inertia," *Quarterly Journal of Economics* 100 (Supplement), 823-838.
- AKERLOF, G. A., AND J. L. YELLEN (1985b): "Can Small Deviations from Rationality Make Significant Difference to Economic Equilibria?" *American Economic Review* 75(4), 708-720.
- ANDREONI, J. (1993): "An Experimental Test of the Public-Goods Crowding-Out Hypothesis" *American Economic Review* 83(5), 1317-1327.
- BOLTON, G. AND A. OCKENFELS (2000): "ERC: A Theory of Equity, Reciprocity and Competition," *American Economic Review* 90, 166-193.
- BENOIT, J.-P., AND V. KRISHNA (1985): "Finitely Repeated Games," *Econometrica*, 53 (4), 905-922.
- CAMERER, C., AND T.-H. HO (1999): "Experienced Weighted Attraction Learning in Normal-Form Games," *Econometrica* 67, 827-874.
- CHAN, K., R. GODBY, S. MESTELMAN AND R. A. MULLER (2002): "Crowding Out Voluntary Contributions to Public Goods," *Journal of Economic Behavior and Organization* 48, 305-317..
- CHARNESS, G., AND M. RABIN (2002): "Understanding Social Preferences with Simple Tests," *Quarterly Journal of Economics* 117, 817-869.
- CHEN, Y., AND C. R. PLOTT (1996): "The Groves-Ledyard Mechanism: An Experimental Study of Institutional Design," *Journal of Public Economics* 59, 335-364.
- CHEUNG, Y.-W., AND D. FRIEDMAN (1997): "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior* 19(1), 46-76.
- CONLON, J. R. (1996): "Cooperation for Pennies: A Note on e-Equilibria," *Journal of Economic Theory* 70, 489-500.
- COSTA-GOMES, M., AND K. ZAUNER (2001): "Ultimatum Bargaining Behavior in Israel, Japan, Slovenia, and the United States: A Social Utility Analysis," *Games and Economic Behavior* 34, 238-269.
- COX, J., AND M. WALKER (1998): "Learning to Play Cournot Duopoly Strategies," *Journal of Economic Behavior and Organization* 36(2), 141-161.
- DAVIS D. D., AND C. A. HOLT (1992): *Experimental Economics*. Princeton University Press.
- DIXIT A., AND M. OLSON (2000): "Does Voluntary Participation Undermine the Coase Theorem?" *Journal of Public Economics*, 76, 307-335.
- DUFWENBERG, M., AND G. KIRCHSTEIGER (2002): "A Theory of Sequential Reciprocity," University of Vienna working paper.

- DUTTA, B, A. SEN, AND R. VOHRA (1995): "Nash Implementation through Elementary Mechanisms in Economic Environments," *Economic Design*, 1, 173-204.
- EREV, I., AND A. ROTH (1998): "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review* 88, 848-881.
- FALK, A., AND U. FISCHBACHER (1998): "A Theory of Reciprocity," University of Zurich working paper.
- FEHR, E., AND S. GACHTER (2000): "Cooperation and Punishment in Public Goods Experiments," *American Economic Review* 90, 980-994.
- FEHR, E., AND K. M. SCHMIDT (1999): "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics* 114, 817-868.
- GROVES, T., AND J. LEDYARD (1977): "Optimal Allocation of Public Goods: A Solution to the 'Free Rider' Problem," *Econometrica* 45, 783-811.
- 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," *Review of Economic Studies* 46 (1979), 217-225.
- ISSAC, R.M. AND J. WALKER (1998) "Nash as an Organizing Principle in the Voluntary Provision of Public Goods: Experimental Evidence," *Experimental Economics* 1(3), 191-206.
- ITO, M., T. SAIJO, AND M. UNE (1995): "The Tragedy of the Commons Revisited: Identifying Behavioral Principles," *Journal of Economic Behavior and Organization* 28(3), 311-335.
- IWASA, Y., M. NAKAMARU, AND S. A. LEVIN (1998): "Allelopathy of Bacteria in a Lattice Population: Competition between Colicin-Sensitive and Colicin-Producing Strains," *Evolutionary Ecology* 12, 785-802.
- KREPS D. M., P. MILGROM, J. ROBERTS, AND R. WILSON (1982): "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory* 27, 245-52.
- LAURY, S. AND C. HOLT (2002): "Voluntary Provision of Public Goods: Experimental Results with Interior Nash Equilibria," in *The Handbook of Experimental Economic Results*, eds., C. Plott and V. Smith (Amsterdam, Elsevier Press), forthcoming.
- LEDYARD, J. O. (1995): "Public Goods: A Survey of Experimental Research," in *The Handbook of Experimental Economics*, eds., J. H. Kagel and A. E. Roth, (Princeton, Princeton University Press), 111-194.
- LEVINE, D. (1998): "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics* 1, 593-622.
- MAYNARD SMITH, J. (1982): *Evolution and the Theory of Games*, Cambridge, Cambridge University Press.

- MOULIN, H. (1986): "Characterizations of the Pivotal Mechanism," *Journal of Public Economics* 31, 53-78.
- MUSGRAVE, R. A. AND P. MUSGRAVE (1973): *Public Finance in Theory and Practice*, New York, McGraw-Hill.
- OCHS, J. AND A. E. ROTH (1989): "An Experimental Study of Sequential Bargaining," *American Economic Review* 79(3), 355-84.
- PALFREY, T. AND H. ROSENTHAL (1984): "Participation and the Provision of Discrete Public Goods: a Strategic Analysis," *Journal of Public Economics* 24, 171-193.
- PRASNIKAR, V. AND A. E. ROTH (1992): "Considerations of Fairness and Strategy: Experimental Data from Sequential Games," *Quarterly Journal of Economics* 107(3), 866-888.
- RABIN, M. (1993): "Incorporating Fairness into Game Theory and Economics," *American Economic Review* 83, 1281-1302.
- RADNER, R. (1980): "Collusive Behavior in Non-cooperative Epsilon-Equilibria in Oligopolies with Long but Finite Lives," *Journal of Economic Theory* 22(2), 289-303.
- SAIJO, T., AND H. NAKAMURA (1995): "The "Spite" Dilemma in Voluntary Contribution Mechanism Experiments," *Journal of Conflict Resolution* 39(3), 535-560.
- SAIJO, T., AND T. YAMATO (1999): "A Voluntary Participation Game with a Non-Excludable Public Good," *Journal of Economic Theory* 84 (2), 227-242.
- SAIJO, T., AND T. YAMATO (2001): "Voluntary Participation in the Design of Non-excludable Public Goods Provision Mechanisms," ISER Discussion Paper, Osaka University.
- SAMUELSON, P. A. (1954): "The Pure Theory of Public Expenditure," *Review of Economics and Statistics* 36, 387-389.
- WALKER, M. (1981): "A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations," *Econometrica* 49, 65-71.

Your Investment Number

**Your
Opponent's
Investment
Number**

Your Payoff	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
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	1127	1379	1647	1919	2183	2427	2641	2816	2944	3019	3039	3001	2905	2755	2555	2313	2038	1743	1443	1154	894	685	548	500
2	1186	1465	1764	2072	2374	2658	2913	3129	3297	3411	3465	3456	3385	3252	3061	2819	2534	2217	1881	1543	1220	933	703	552	500
3	1554	1888	2232	2575	2902	3202	3463	3675	3831	3925	3952	3911	3801	3626	3391	3102	2770	2406	2027	1648	1290	973	721	556	500
4	2017	2401	2787	3160	3508	3817	4078	4281	4420	4488	4483	4403	4250	4028	3743	3404	3020	2608	2181	1759	1363	1015	740	561	500
5	2578	3010	3432	3831	4193	4507	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	5765	5677	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	6475	6984	7422	7779	8043	8209	8271	8225	8073	7816	7458	7007	6472	5867	5207	4508	3793	3084	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	8873	9384	9800	10109	10306	10384	10339	10173	9886	9482	8970	8359	7661	6892	6070	5217	4354	3509	2710	1987	1372	899	600	500
12	9653	10250	10750	11142	11416	11567	11589	11480	11242	10877	10390	9791	9090	8302	7444	6534	5596	4654	3736	2871	2092	1432	926	606	500
13	11158	11749	12229	12589	12820	12916	12875	12694	12376	11925	11349	10656	9860	8976	8022	7019	5992	4967	3972	3039	2201	1494	953	613	500
14	12796	13372	13824	14144	14323	14356	14243	13982	13576	13033	12358	11565	10667	9681	8627	7526	6406	5292	4217	3213	2315	1559	982	620	500
15	14570	15123	15538	15808	15925	15888	15694	15344	14844	14199	13420	12520	11514	10419	9258	8055	6836	5631	4473	3394	2433	1626	1012	627	500
16	16484	17003	17372	17583	17630	17513	17229	16783	16179	15426	14535	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	21353	21047	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: Detailed Payoff Table

Own Payoff-Maximizing Strategy Region

Your Investment Number

Your Opponent's Investment Number

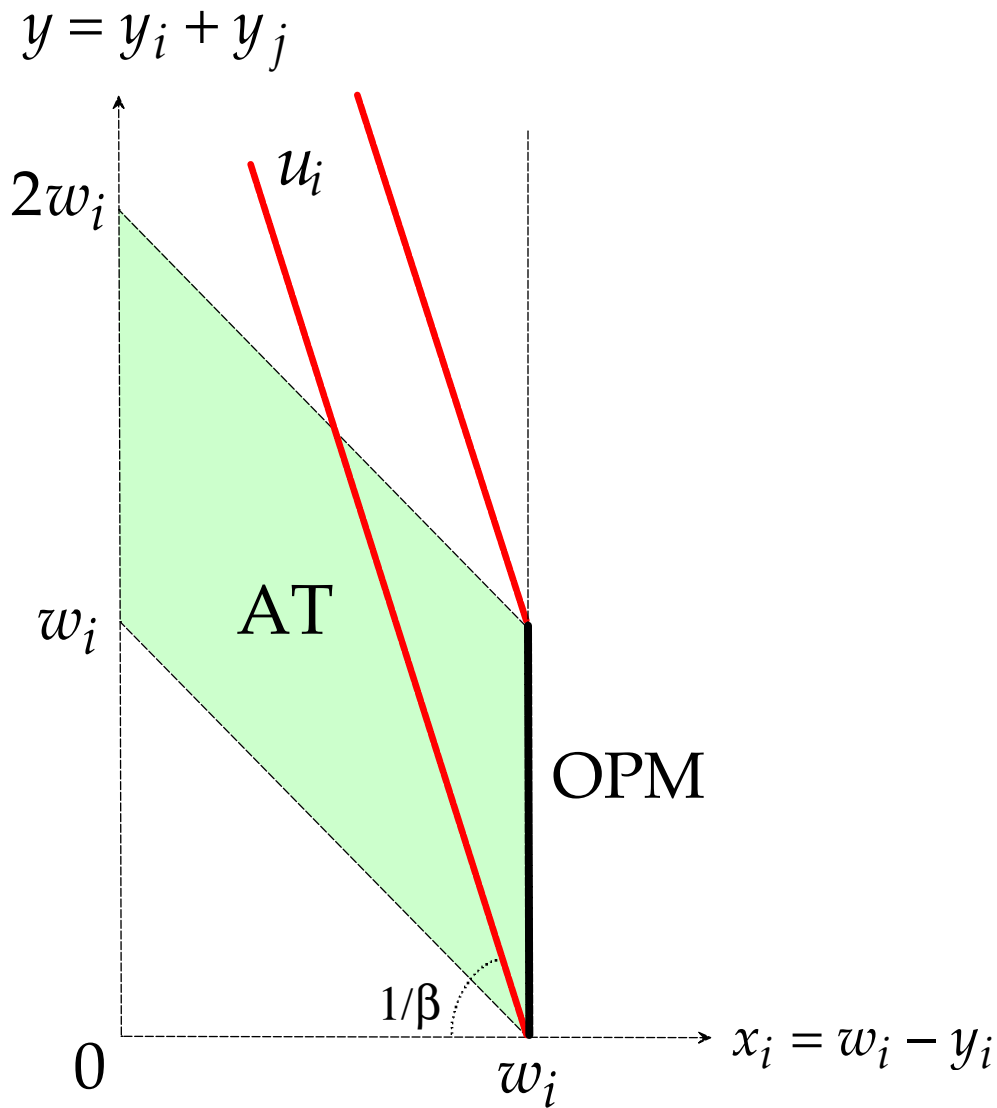
Your Payoff	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
0	706	871	1072	1297	1536	1775	2003	2210	2400	2580	2658	2735	2810	2885	2960	2039	2106	1871	1614	1349	1091	858	669	543	500
1	905	1186	1554	2017	2578	3244	4018	4904	5907	7031	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500
2	1186	1554	2017	2578	3244	4018	4904	5907	7031	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500
3	1554	2017	2578	3244	4018	4904	5907	7031	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500
4	2017	2578	3244	4018	4904	5907	7031	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500
5	2578	3244	4018	4904	5907	7031	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500
6	3244	4018	4904	5907	7031	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500
7	4018	4904	5907	7031	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500
8	4904	5907	7031	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500
9	5907	7031	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500
10	7031	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500
11	8278	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500
12	9653	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500
13	11158	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500
14	12796	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500
15	14570	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
16	16484	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
17	18539	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
18	20739	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
19	23086	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
20	25583	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
21	28231	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
22	31034	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
23	33993	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500
24	37111	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500	500

Spiteful Strategy Region

Altruistic Strategy Region

Weakly Spiteful Strategy Region

Table 2: Own payoff-maximizing, spiteful, and altruistic strategies



AT: Altruistic Strategy Region, OPM: Own Payoff-Maximizing Strategy Region
 SF: Spiteful Strategy Region

Figure 1-(a)

The case of a linear payoff function:
 no spiteful strategy exists.

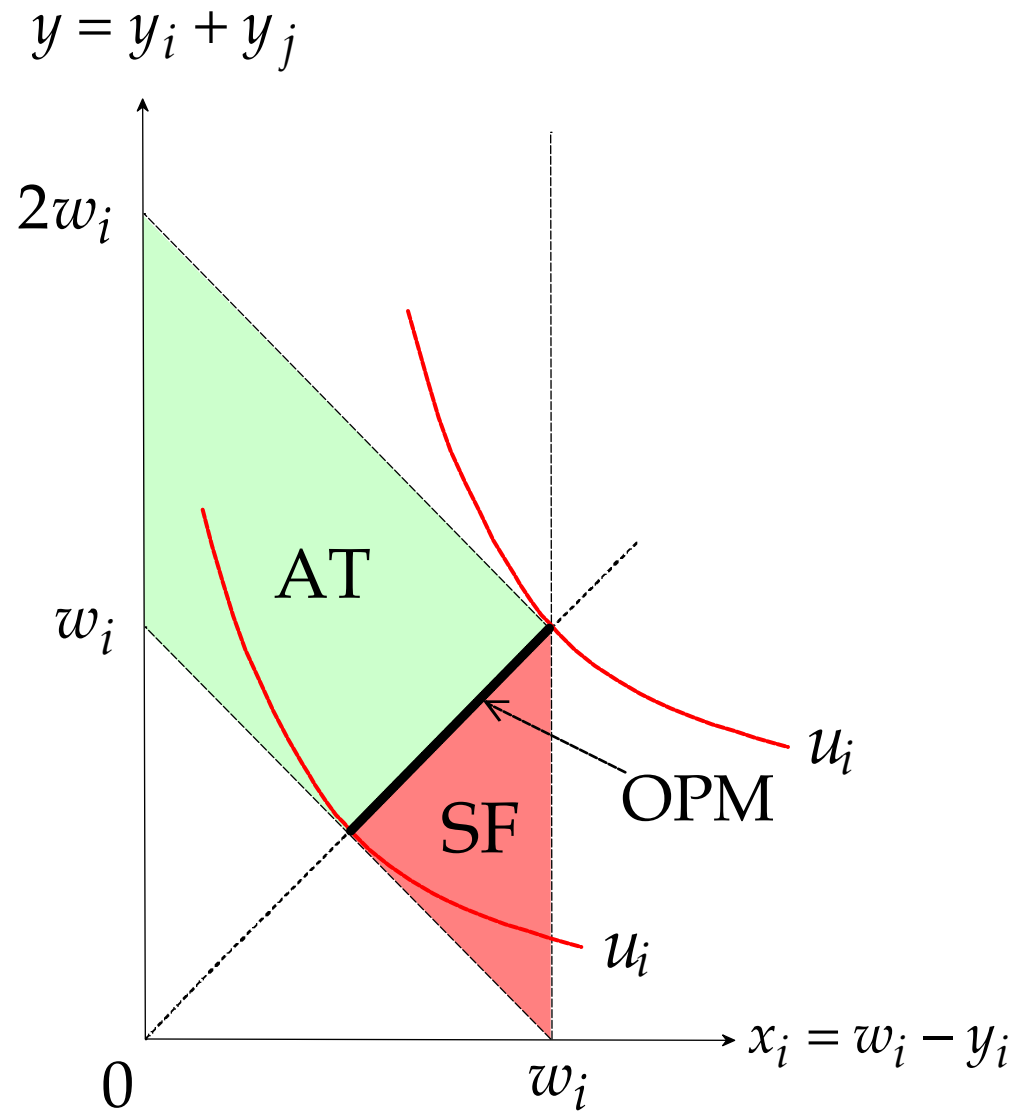


Figure 1-(b)

The case of a non-linear Cobb-Douglas type
 payoff function.

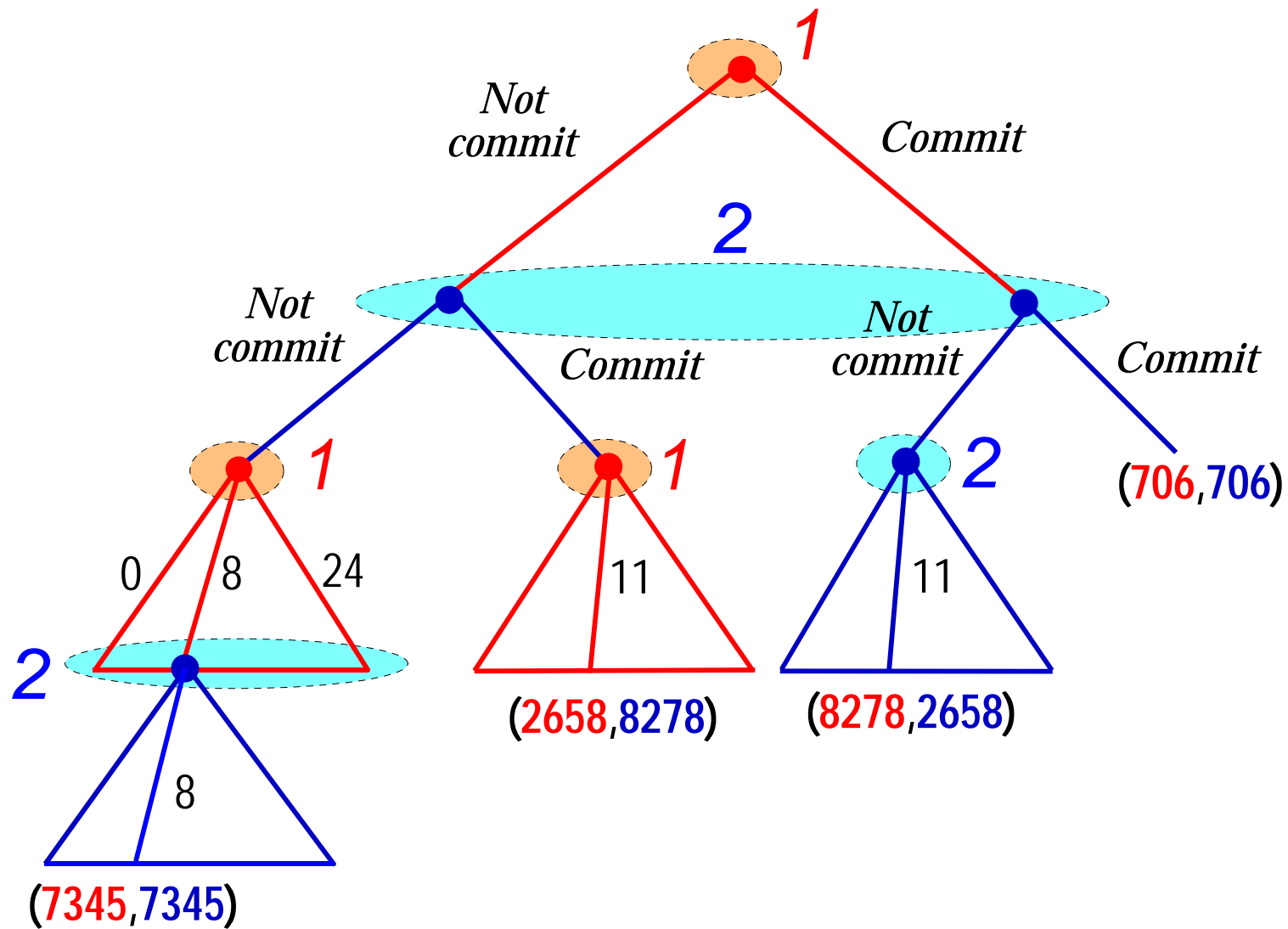


Figure 2. The game tree when subjects can choose whether or not they commit to contributing zero in the voluntary provision of a non-excludable public good.

		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 3. The payoff table becomes a Hawk-Dove game.

Your Investment Number

Your Payoff	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
0																									
1																									
2			1484	<		2446	<		3066	<		3246	>		2902	>		2127	>		1200	>		581	
3			^			^			^			^			^			^			^			^	
4																									
5			3245	<		4512	<		5057	>		4921	>		4109	>		2821	>		1460	>		609	
6			^			^			^			^			^			^			^			^	
7																									
8			5705	<		7036	<		7332	>		6749	>		5383	>		3536	>		1724	>		636	
9			^			^			^			^			^			^			^			^	
10																									
11			9122	<		10295	>		10154	>		8952	>		6883	>		4364	>		2024	>		667	
12			^			^			^			^			^			^			^			^	
13																									
14			13575	<		14338	>		13553	>		11545	>		8620	>		5310	>		2364	>		701	
15			^			^			^			^			^			^			^			^	
16																									
17			19132	<		19206	>		17555	>		14548	>		10603	>		6376	>		2743	>		739	
18			^			^			^			^			^			^			^			^	
19																									
20			25855	>		24939	>		22187	>		17973	>		12838	>		7567	>		3163	>		781	
21			^			^			^			^			^			^			^			^	
22																									
23			33800	>		31573	>		27472	>		21836	>		15336	>		8887	>		3626	>		827	
24																									

Your
Opponent's
Investment
Number

Table 4. Summary Payoff Table

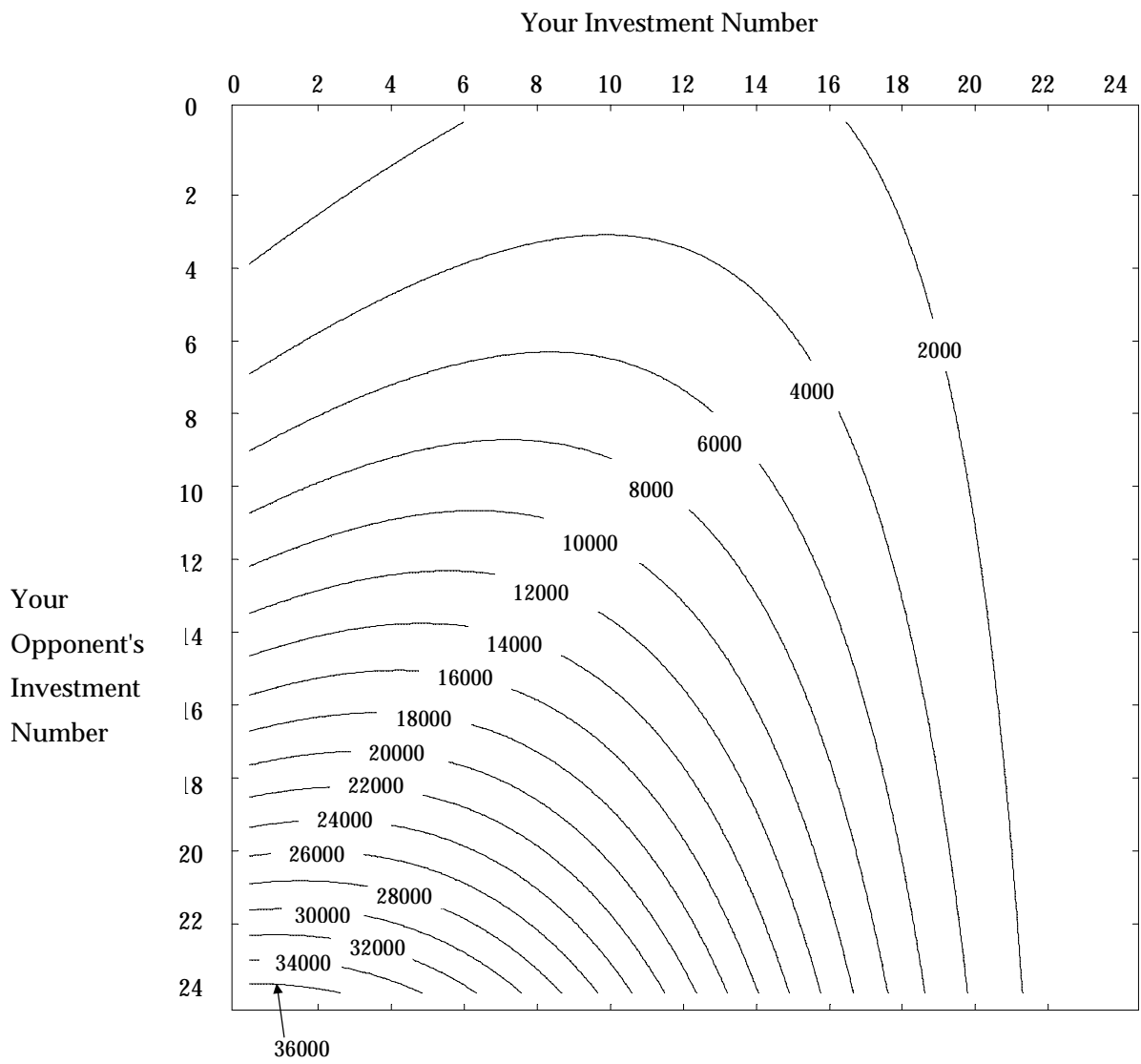


Table 5. Iso-Payoff Map

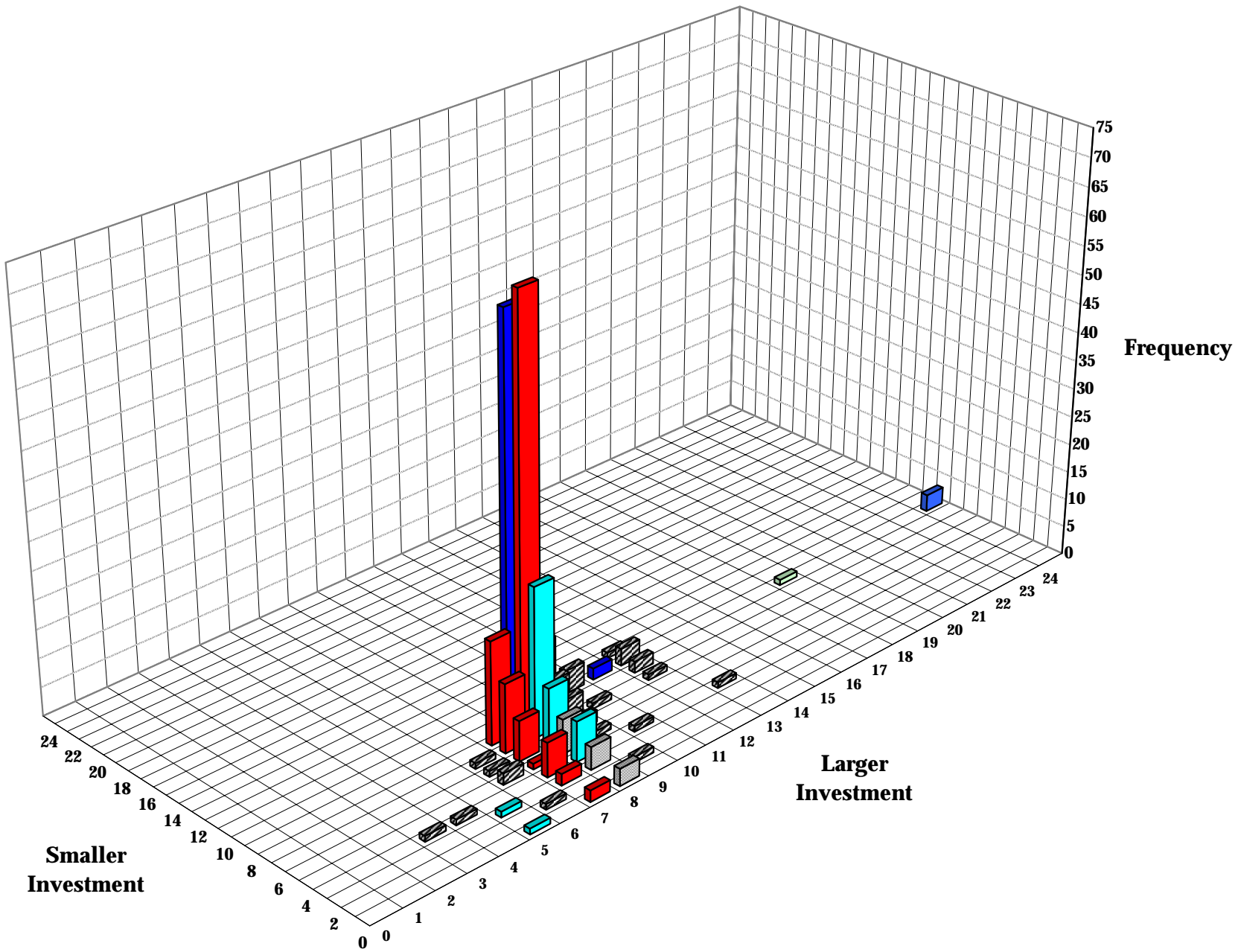


Figure 3. Investment Pattern for Treatment A.

Average Investment Pattern

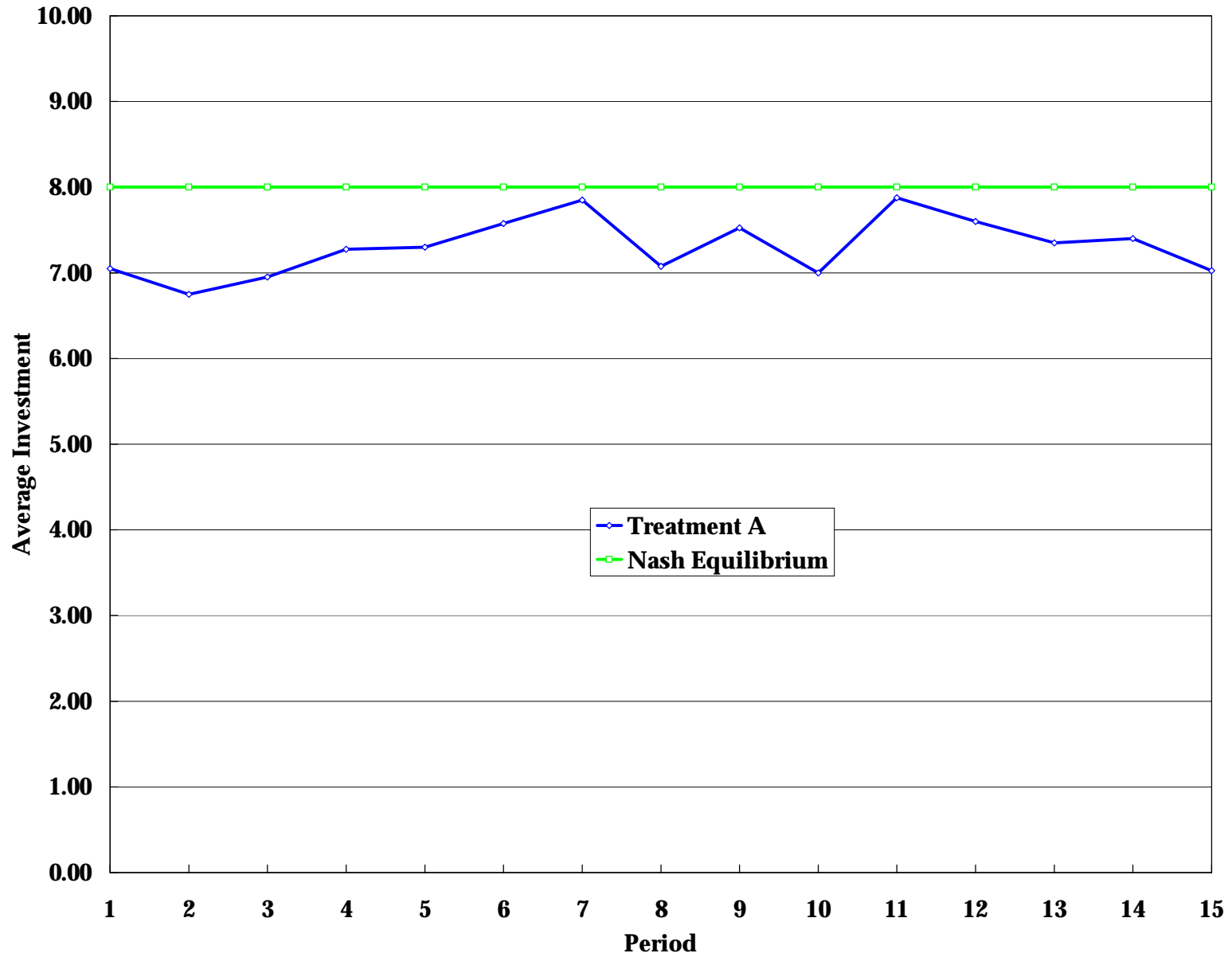


Figure 4. Average Investment Pattern for Treatment A.

	Expectation	Spiteful strategies (%)	Own payoff-maximizing strategies (%)	Altruistic strategies (%)
Random choice: ratios in Table 2		22.40 (=140/625)	4.16 (=26/625)	73.40 (=459/625)
Treatment A	Nash	43.83 (=263/600)	48.50 (=291/600)	7.67 (=46/600)
	Cournot	49.96 (=263/560)	41.79 (=234/560)	11.25 (=63/560)
	Average (Fictitious Play)	45.18 (=253/560)	43.21 (=242/560)	11.61 (=65/560)
Treatment B: Neither commits.	Nash	52.32 (=305/583)	34.48 (=201/583)	13.21 (=78/583)
	Cournot	48.57 (=254/523)	35.18 (=184/523)	16.25 (=85/523)
	Average (Fictitious Play)	48.76 (=255/523)	34.80 (=182/523)	16.44 (=86/523)
Treatment B: Only one commits.		68.35 (=95/139)	30.93 (=43/139)	0.72 (=1/139)

* For Cournot and average (fictitious play) expectations, we exclude the first investment choice of each subject for Treatment A and Treatment B when neither commits. Thus the number of observations to count for Treatment A is 600-40=560, and that for Treatment B when neither commits is 583-60=523.

Table 6. Ratios of Spiteful, Own Payoff-Maximizing, and Altruistic Strategies Chosen by Subjects.

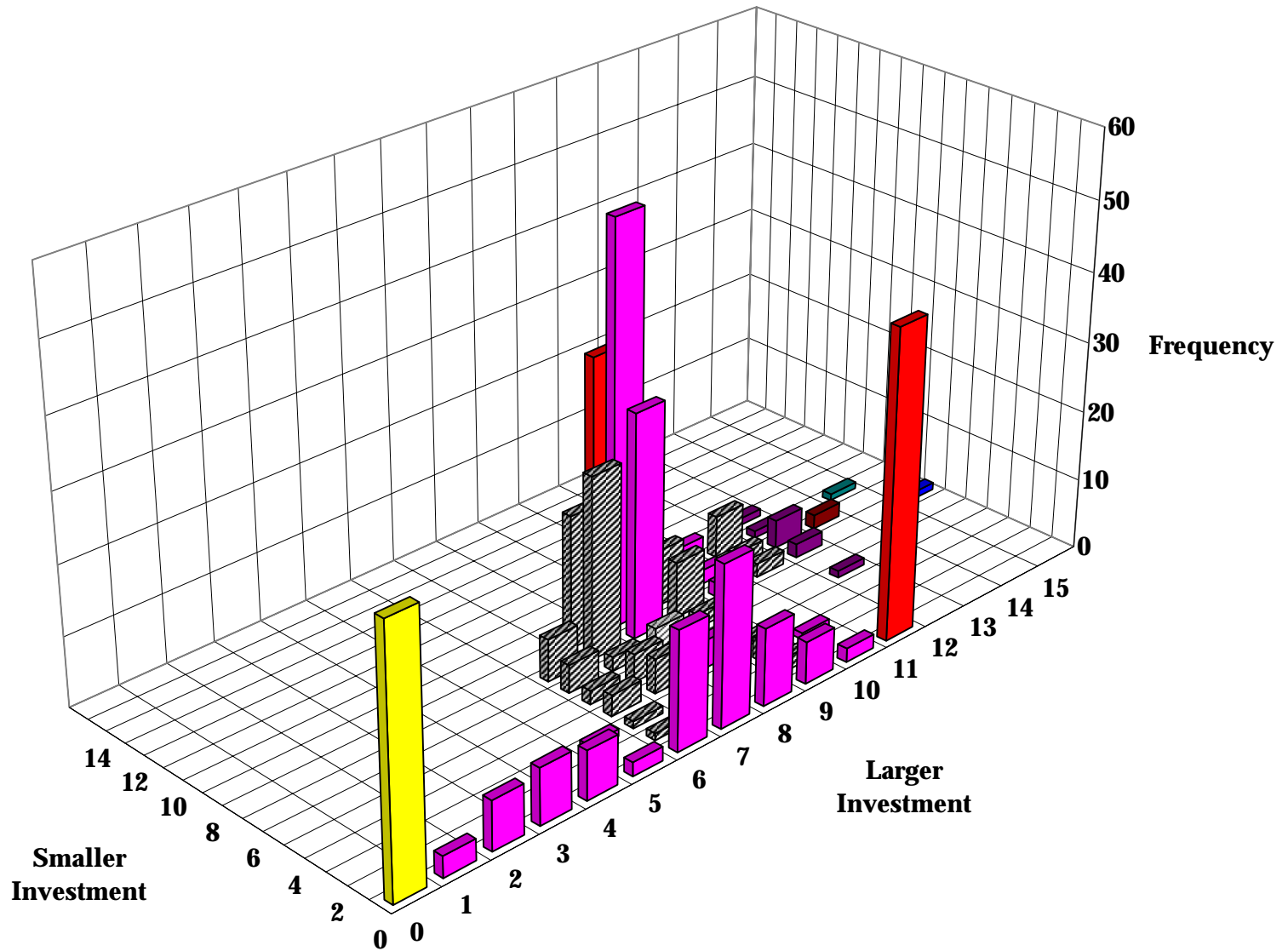


Figure 5. Investment Pattern in the Three Treatment B Sessions.

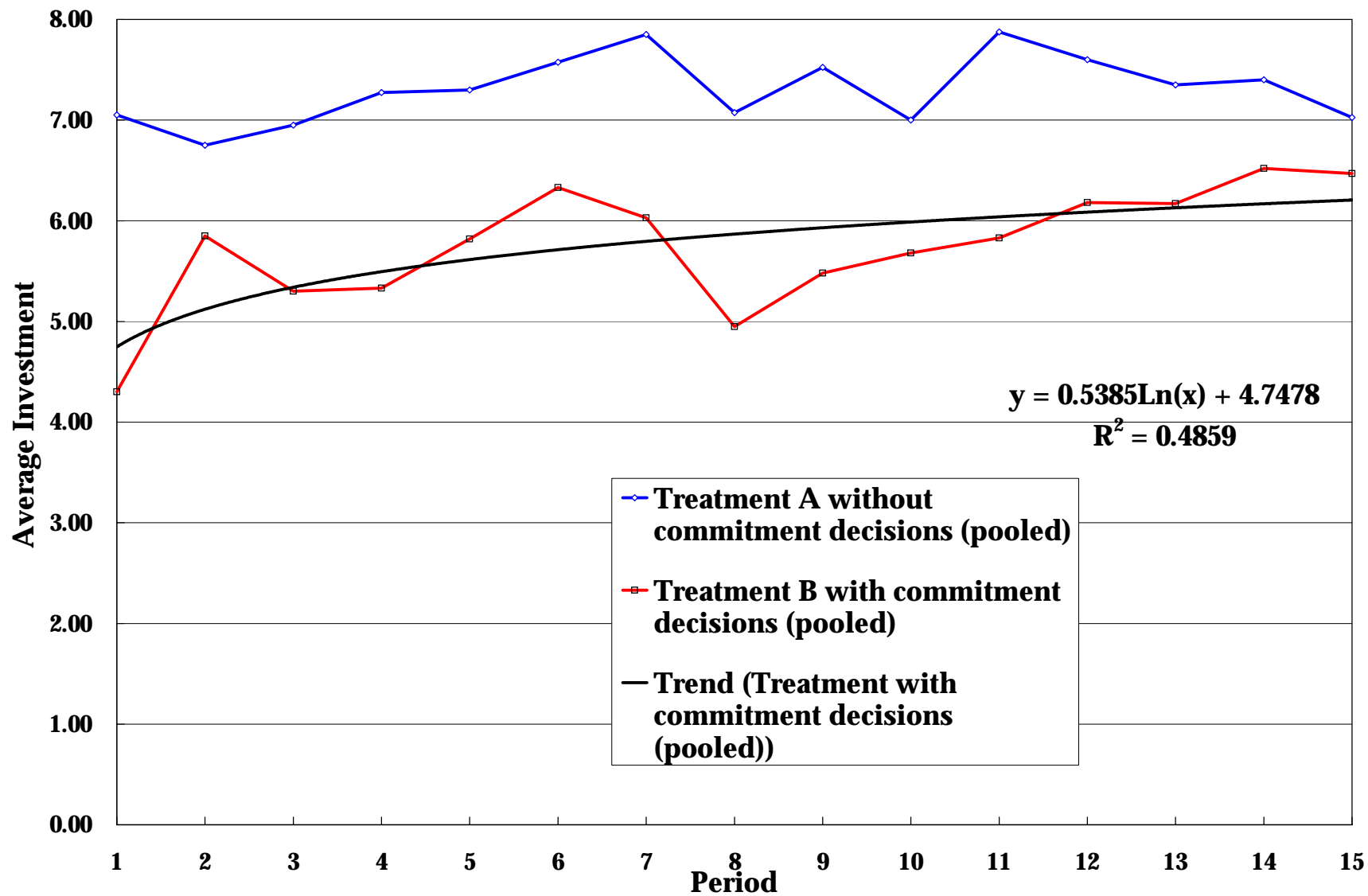


Figure 6. Comparison of Average Investment Patterns:
Treatment A versus Treatment B.

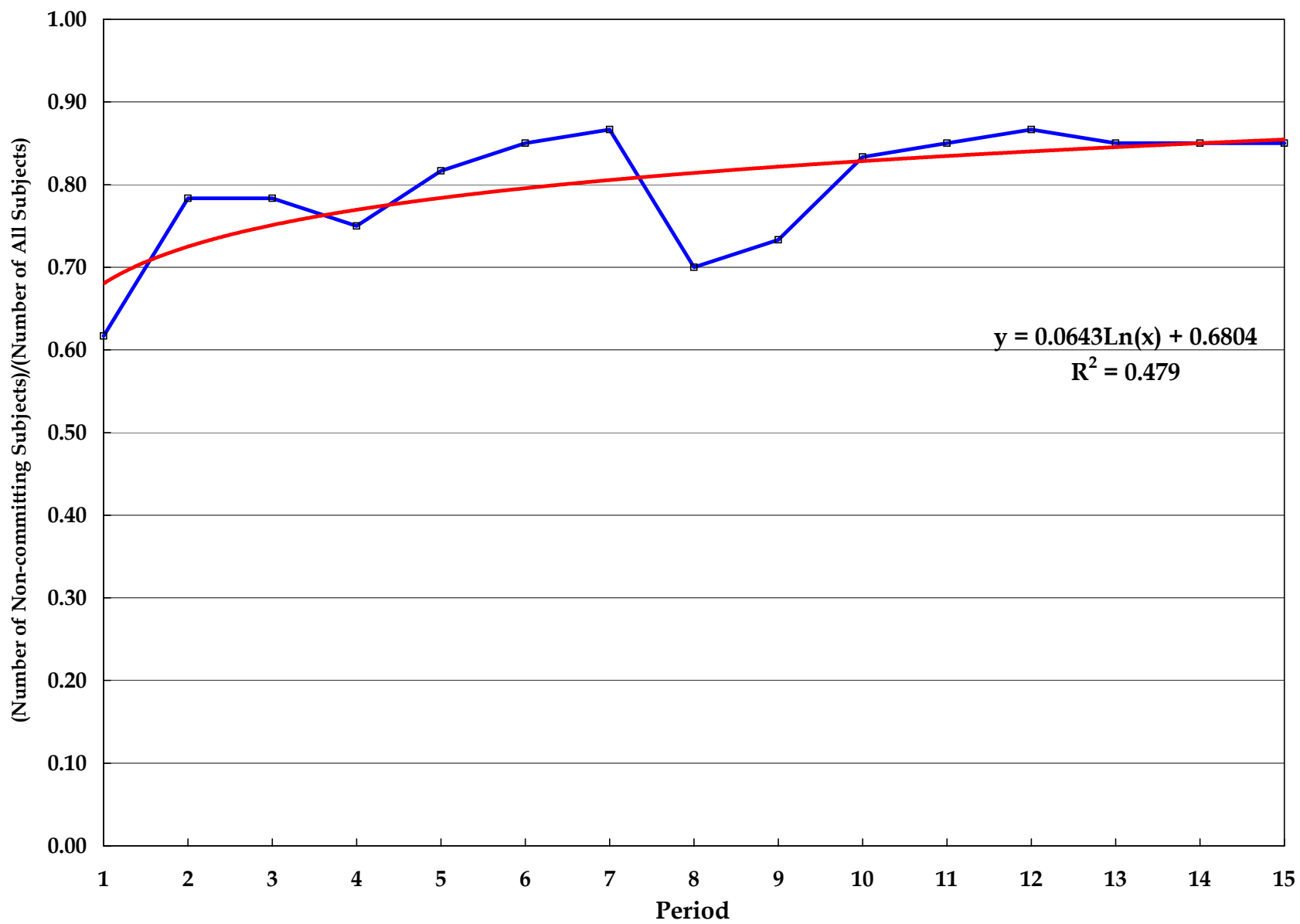


Figure 7. Non-commitment Rate Pattern.

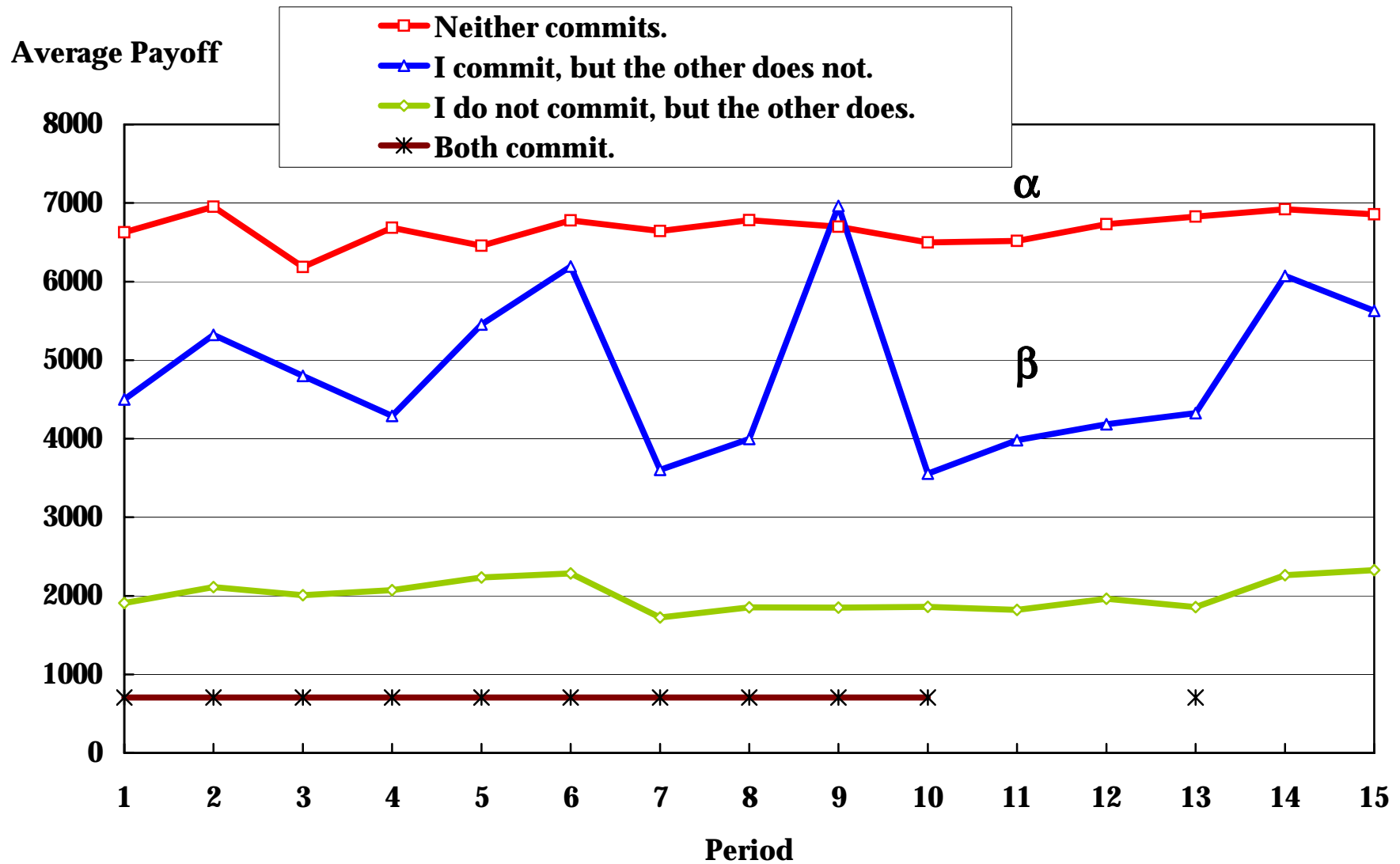


Figure 8. Average Payoffs by Commitment Decision.

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

Table 7. The average values of payoff data up to round 5 for the three treatment B sessions.

Dependent Variable = 1 if Not commit
 0 if Commit

Probit Model with random subject effect

Expectations Variable or Statistic	Model (1): Cournot Expectations				Model (2): Fictitious Play			
	Tsukuba B	Tokyo B	Tokyo B'	B and B' Pooled	Tsukuba B	Tokyo B	Tokyo B'	B and B' Pooled
(Expected Non-commitment Earnings)/ (Expected Commitment Earnings)	0.051* (0.022)	-0.044 (0.043)	0.086** (0.031)	0.028* (0.015)	-0.013 (0.054)	-0.132* (0.064)	0.131 (0.193)	-0.002 (0.025)
Intercept	0.618** (0.187)	1.279** (0.224)	1.126** (0.241)	1.010** (0.115)	0.915** (0.211)	1.312** (0.240)	0.935** (0.222)	1.073** (0.120)
Rho ^a	0.382* (0.165)	0.339 (0.221)	0.447* (0.191)	0.344** (0.100)	0.280* (0.137)	0.371* (0.176)	0.382* (0.193)	0.333** (0.094)
Number of Observations	300	300	300	900	300	300	300	900
Estimated Log-likelihood	-147.3	-123.7	-129.8	-405.6	-149.0	-122.7	-130.0	-406.7
Restricted (slopes=0) Log-likelihood	-161.6	-136.5	-145.8	-450.1	-161.6	-136.5	-145.8	-450.1

Standard errors in parentheses.

* denotes significantly different from zero at the 5-percent level; ** denotes significantly different from zero at the 1-percent level (all two-tailed tests except for (Expected Non-commitment Earnings)/(Expected Commitment Earnings)).

^aRho is a standard Hausman test statistic for the presence of random effects.

Table 8. Results of Probit Commitment Model

Average Investment Pattern

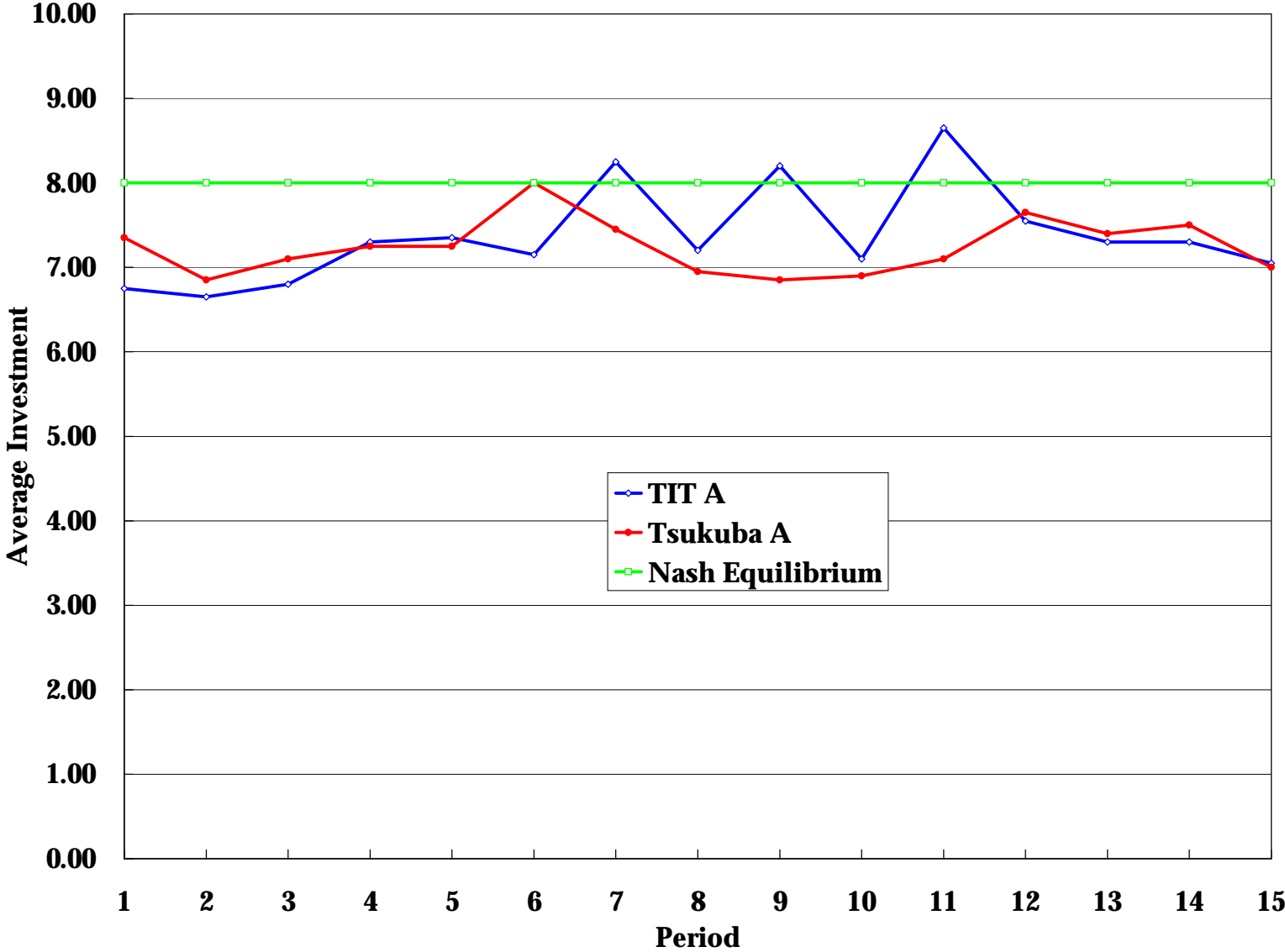


Figure 9. Average Investment Pattern for Treatment A.

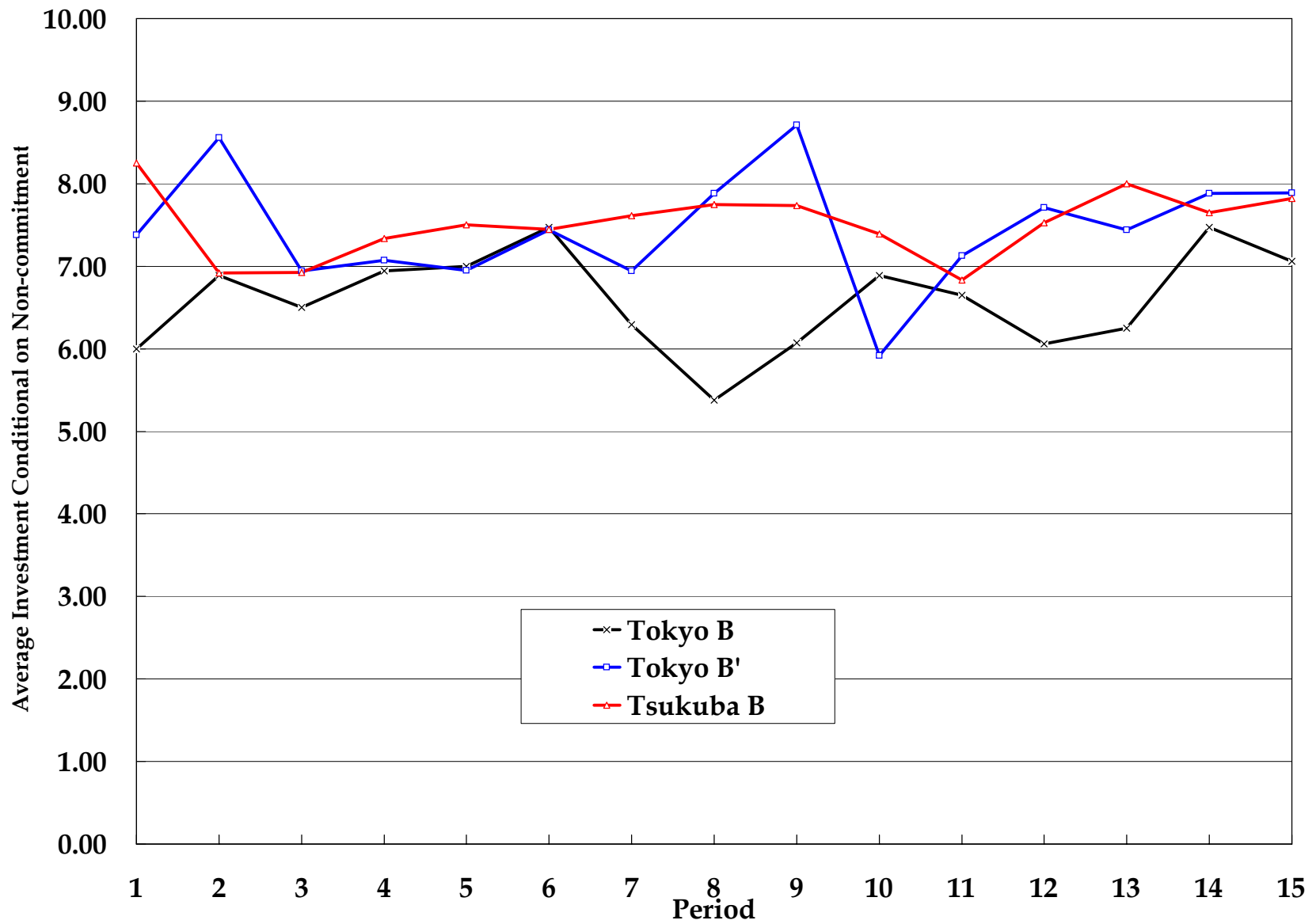


Figure 10. The Average Investment Patterns Conditional on Non-commitment in Tsukuba B, Tokyo B, and Tokyo B' Sessions.

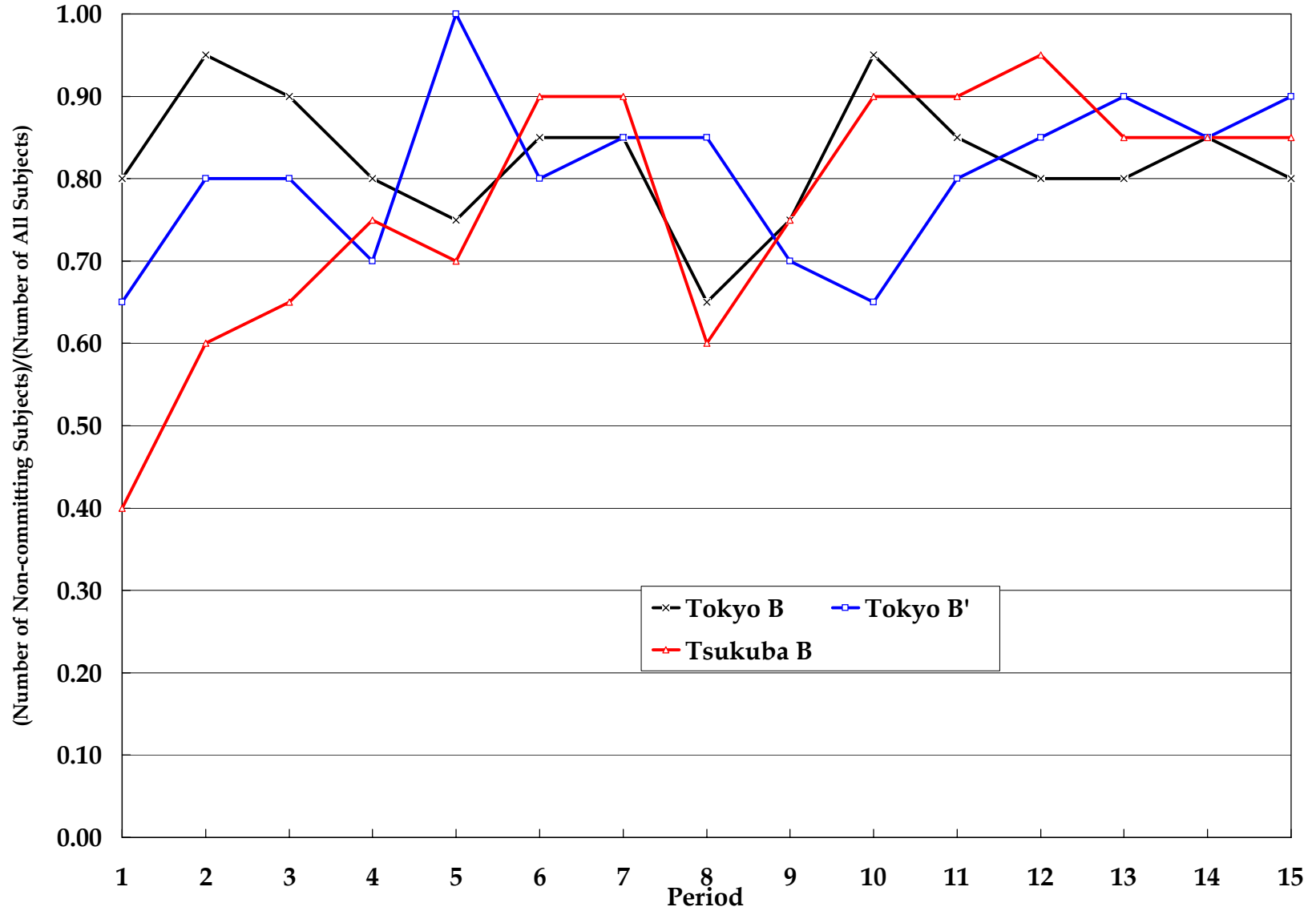


Figure 11. The Average Non-commitment Rate Patterns in Tsukuba B, Tokyo B, and Tokyo B' Sessions.