Experimental Social Sciences

Toward Experimentally-based New Social Sciences for the 21st Century

ExpSS21

Working Paper

ExpSS-E-4

Designing a mechanism to cope with coordination failure for goods with network externalities: an experimental study

Tatsuhiro Shichijo
Osaka Prefecture University

Kenju Akai
Japan Society for the Promotion of Science

Takao Kusakawa
Hiroshima Shudo University

Tatsuyoshi Saijo
Osaka University and UCLA

May, 2008

Grant-in-Aid for Scientific Research on Priority Areas

Designing a mechanism to cope with coordination failure for goods with network externalities: an experimental study

Tatsuhiro Shichijo*

Kenju Akai†

Takao Kusakawa ‡

Tatsuyoshi Saijo§

March 24, 2008

Abstract

It is not easy to sell goods with network externalities when the number of consumers possessing these goods is small. This is because the consumers may not gain sufficient utility from the goods if the network size is small. Moreover, when consumers' preferences are private information, consumers have an incentive to delay their purchasing decisions in order to obtain the others' information. Although there exists a more efficient outcome, consumers cannot achieve it. We suggest a mechanism to resolve this problem. The mechanism has the characteristics of cheap talk

*Osaka Prefecture University. e-mail: shichijo@eco.osakafu-u.ac.jp

†Japan Society for the Promotion of Science. e-mail: k-akai@iser.osaka-u.ac.jp

‡Hiroshima Shudo University. e-mail: kusakawa@shudo-u.ac.jp

§Osaka University and UCLA. e-mail: saijo@iser.osaka-u.ac.jp

and costly signal. We also conducted experiments to the mechanism and obtained results that the mechanism resolves the coordination problem.

Keywords Coordination failure, Network externalities

Journal of Economic Literature classification numbers: C91, H23

1 Introduction

The characteristic of a good with network externalities is that the utility to a purchaser from the good is increasing in the number of consumers possessing the good. Once the number of purchasers of a good exceeds critical mass, the good is sold easily[1]. However, when only a few people buy it, the utility is very small, which makes it difficult to sell the good. Therefore, there is a possibility that the diffusion of the good stagnates even if diffusion is desired by society.

For example, newly developed land such as a new industrial park is a good with network externalities. If, on the one hand, many individuals and/or firms establish their business within the park, then the value of the land increases because of the agglomeration economy. On the other hand, if only a few individuals or firms buy sites, then the value of the land is lower. For example, the northeast industrial park in San Angelo, Texas has many sites that sit idle. In Japan too, there is some newly developed land that remains unsold. These failures cause large losses for the local government that developed the land. For example, the Bureau of Public Enterprise of Osaka Prefecture in Japan, which develops new land, was abolished after accumulating a 22 billion yen deficit. One of the reasons was the coordination failure from the network externalities.

As another example, the fuel cell car, which uses hydrogen, is a good with network externalities. When sales of fuel cell cars increase, the number of hydrogen filling stations increases and the convenience of fuel cell cars increases. The diffusion of the fuel cell car is socially desired. However, even if the price of the fuel cell car falls, it is not easy to sell the car because the number of hydrogen filling stations is small initially.

There are many theoretical studies about network externalities that describe coordination failure. For example, Farrell and Saloner (1985) analyzed the tech-

[1] We refer to a person who buys the good, as a purchaser.

nology adoption problem using a model of incomplete information about firms' preferences for new technology[2]. They showed that, in some cases, a socially beneficial technology will not be adopted as the outcome of a perfect Bayesian Nash equilibrium ("excess inertia"). This situation can arise because each firm fears to adopt the technology on its own and decides to wait and observe other firms' behavior, with the result that all firms remain in the status quo. In addition, Farrell and Saloner showed that nonbinding communication about preferences could remove some, but not all, types of the inefficiency.

However, there are only a few theoretical studies about how to solve these coordination problems of network externalities. Rauch (1993) studied the problem of industrial parks and obtained the result that an introductory price helps to avoid coordination failure. Park (2004) developed an inducement mechanism for the good with network externalities. Shichijo and Nakayama (2006) developed other dynamic inducement mechanism model, which uses a uniform price and refundable reservation tickets.

There are many experimental studies about coordination problems in complete information games. However there are few experimental studies about coordination problems in incomplete information games. Van Huyck, Battalio, and Beil (1990, 1991) demonstrated that the coordination failures arise in the so-called Minimum and Median games with multiple Pareto-ranked equilibria. For complete information games, two types of resolutions were proposed: cheap talk, and costly signal. For example, Duffy and Feltovich (2002) showed that cheap talk solves some kinds of coordination problems. In addition, Duffy and Feltovich (2002) and Van Huyck, Battalio, and Beil (1993) showed that costly signals are effective too. However, for a game with incomplete information such as Farrell and Saloner (1985), these solutions cannot eliminate the possibility of

[2]There are many other studies about network externalities, e.g., Katz and Shapiro (1986) and Rohlfs (1974).

miscoordination.

Compared with the large number of experiments about coordination failure, there are only a few experiments about network externalities. Chakravarty (2003a,b) conducted experiments that involve both sellers and consumers. Etziony and Weiss (2002) studied consumer behavior. Mak and Zwick (2006) studied network externalities with continuous demand. However, there are no experiments that considered a new mechanism to overcome the coordination failure from network externalities.

We design a simple mechanism to sell the good with network externalities and conduct experiments to investigate its performance. We should note that the mechanism should satisfy a participation constraint, because we cannot force consumers to participate in the mechanism. As we will see in subsection 2.2, efficient mechanisms, such as the Groves mechanism, cannot be used by the participation constraint. Thus, we propose a new mechanism, which uses particular reservation tickets and has two characteristics resolving coordination failure: cheap talk and costly signal. The reservation tickets are not free. Agents have to pay for them. Thus, it has the characteristic of a costly signal. However, they are refundable if only a few agents buy the tickets. That is, agents do not have to bear the risk of small network size when they buy them. Thus, it has the characteristic of cheap talk too. In equilibrium, agents reveal their types by buying the reservation tickets and know the other's type by seeing the sales of the reservation tickets. Thus, they can coordinate and improve their welfare.

From the experiment, we have the statistically significant result that the mechanism improves social welfare. In the experiments of the mechanism with reservation tickets, subjects chose the alternatives, which are predicted by theory, except for a few observations. Moreover, we do not find any systematic change over time. That is, most of subjects did not need to learn from experience and could choose the theoretically predicted alternative from the first round.

The paper is organized as follows: In Section 2, we consider the theoretical model. In Section 3, we show the procedures of the experiment. In Section 4, we report the experimental results. Section 5 concludes.

2 A Theoretical Model

2.1 Original Game

Let $I = \{1, 2\}$ be the set of agents and let $\Theta_1 = \Theta_2 = \{L, H\}$ be the set of agents' types[3]. We denote the type profile by $\Theta \equiv \Theta_1 \times \Theta_2$. We denote by $q \in (0, 1)$ the probability that the other member in the same group[4] is type H, and hence $1 - q$ is the probability of type L. Although we can extend the model to a finite type and finite players game (Shichijo (2007)), for simplicity we consider the two type and two player case.

First, we consider the following incomplete information game. There are two stages. The decision flow of the agents is illustrated in Fig. 1. First of all, nature determines an agent's type. Each agent knows his/her type, but does not know his/her partner's type. Each agent has two choices, "B" and "N". "B" represents buying the good and "N" represents not buying the good[5]. After that, they go to stage 2. At the beginning of stage 2, each subject receives information about the choice of his/her partner at stage 1. Each agent has two choices "B" and "N" again. However, this time, the agent who has chosen "B" at stage 1 cannot change his/her decision of "B" at stage 2. That is, we assume

[3] "L" and "H" mean Low and High type respectively. However, we use the neutral terms "A" and "B" in experiments instead of "L" and "H".

[4] In order to avoid possible framing effects, we used "the other member in the same group" rather than the words "partner" or "opponent" in our experiment. We will use "partner" instead of "the other member in the same group" in the following sentences since the latter is quite cumbersome.

[5] Instead of "B" and "N", we use the neutral terms "X" and "Y" in the experiments.

that “Buy the good” is an irreversible action as in Gale (1995, 2001) and that the agent who chooses “B” at stage 1 keeps holding the good because the good is durable. On the other hand, the subject who has chosen “N” at stage 1 may choose either “B” or “N” at stage 2. For simplicity, we assume that each agent can buy only one good.

[[Insert Figure 1 around here]]

The payoff to agents is determined by whether they buy the good and the number of purchasers at stage 2. If the agents do not buy the good in both stages, we assume their payoff is zero for normalization. On the other hand, if agents buy the good at either stage, then the utility from the good is $v(k|\theta)$ where θ is the agent's type and k is the number of agents who have bought the good by the end of stage 2. Because of the network externalities, $v(k|\theta)$ is increasing in k.

For simplicity, we assume that the cost to produce one good is fixed at c and that the price of the good is equal to c. Moreover, we assume $v(1|L) - c < v(2|L) - c < 0$. That is, the type L agents do not have the incentive to buy the good if the price is equal to the production cost. We also assume that $v(1|H) - c < 0 < v(2|H) - c$. That is, the type H agents have the incentive to buy the good if the price is equal to the production cost and their partners buy the good. Finally, we assume that $v(2|H) + v(2|L) - 2c > 0$. That is, the social welfare with both agents' buying some goods is greater than that when they do not buy any. We refer to the game defined above as the *original game*.

On the one hand, if a type L agent buys the good, then his/her payoff is negative. Thus, the strategy that “Buy the good when his/her type is L” is a strictly dominated strategy. On the other hand, if his/her type is H and the partner buys the good at stage 1, then buying the good at stage 2 gives him/her a positive payoff, $v(2|H) - c > 0$. Therefore, the strategy “Do not buy the good at stage 2 when his/her type is H and the partner buys the good at stage 1" is a strictly dominated strategy. We can delete these strategies to obtain the perfect Bayesian equilibrium. Then, the expected payoff of type H from buying the good at stage 1 is $(1-q)v(1|H) + qv(2|H) - c$. We can easily verify the following facts.

Fact 1: If $(1-q)v(1|H) + q\,v(2|H) - c > 0$, then every perfect Bayesian equilibria in the original game involve some agents buying the good.

Fact 2: If $(1-q)v(1|H) + q\,v(2|H) - c < 0$, then every perfect Bayesian equilibria in the original game involve no agent buying the good.

In the case of Fact 2, the expected payoff to agents in equilibria is zero, even though both of them are type H. If they knew the partners' type and could coordinate, then they could improve the social welfare level.

2.2 Ex post incentive compatibility and participation constraints

In this section, we consider mechanisms that overcome the coordination failure explained in Section 2.1.

An outcome consists of two components: a decision rule about allocation of the good, and the payments of agents for the good. Let us denote by $d_i : \Theta \to \{0, 1\}$ a decision rule about the allocation to agent i. $d_i(\theta) = 1$ if agent i is allocated the good whenever the type profile is θ. $d_i(\theta) = 0$ if agent i does not have the good whenever the type profile is θ. Let a decision rule be $d(\theta) = (d_1(\theta), d_2(\theta))$ and the number of purchasers of the good be $\sharp d(\theta_1, \theta_2) = d_1(\theta_1, \theta_2) + d_2(\theta_1, \theta_2)$. We say a decision rule d satisfies *efficiency* if $\sum_i v(\sharp d(\theta)|\theta_i) \geq \sum_i v(\sharp d'(\theta)|\theta_i)$ for any decision rule d'. Note that in order for a mechanism to be efficient, its decision rule has to be efficient.

Let us denote by $\xi_i(\theta)$ agent i's payment when the type is θ. We say the scheme satisfies *feasibility* if total payments are larger than or equal to the total production cost, i.e., $\xi_1(\theta) + \xi_2(\theta) \geq \sharp d(\theta)c$. Because it is unnatural for agents who are not given the good to pay some money, we assume that $\xi_i(\theta) > 0$ only if $d_i(\theta) = 1$. We refer to $f = (d, \xi)$ as a social choice function. The utility of agent i from the social choice function is $\pi_i(f(\theta) \mid \theta_i) = d_i(\theta)v(\sharp d(\theta) \mid \theta_i) - \xi_i(\theta)$.

We consider a robust mechanism, which is not sensitive to the belief agents have. Thus, we require the mechanism to satisfy the following ex post incentive compatibility condition instead of Bayesian incentive compatibility[6].

Definition 2.1. *We say a social choice function f satisfies ex post incentive compatibility, if $\pi(f(\theta_i, \theta_{-i}) \mid \theta_i) \geq \pi(f(\hat{\theta}_i, \theta_{-i}) \mid \theta_i)$ for all $\hat{\theta}_i \in \Theta_i$.*

On the other hand, the participation constraint is important to the mechanism to ensure that selling occurs, because government or companies cannot usually force consumers to participate in the mechanism to ensure selling.

The participation constraint imposes strict restriction on mechanisms. Saijo and Yamato (1999) considered the case that the agents who do not participate in the mechanism can obtain benefit from the outcome due to non-excludability. Saijo and Yamato (1999) theoretically showed that there is no mechanism in which every agent participate in equilibrium. Cason, Saijo, Yamato, and Yokotani (2004) confirmed it by experiments. Moreover, even if the concerned good is excludable, the participation constraint imposes restriction on mechanisms. There exist both negative results for implementation with participation constraints (Myerson and Satterthwaite (1983), Schweizer (2006)) and positive results (Matsushima (2007)). We consider a social choice function that satisfies the following ex post participation constraint[7].

[6]Ex post incentive compatibility is required by robust concepts of implementation such as implementation in dominant strategies, ex post implementation and robust implementation.

[7]We have the same result even if we use an interim participation constraint, which requires the expected payoff to all types to be larger than or equal to zero.

Definition 2.2. *We say a social choice function f satisfies the ex post participation constraint, if $\pi_i(f(\theta) \mid \theta_i) \geq 0$ for all $\theta \in \Theta$ and all i.*

As stated above, there was some positive results for participation constraints. However, these studies use Bayesian incentive compatibility, which is sensitive to the precise beliefs of agents. Because we use the stricter condition, ex-post incentive compatibility, we have the following negative results.

Fact 3: There is no social choice function which satisfies ex post incentive compatibility, feasibility, the participation constraint, and efficiency of the decision rule.

The proof is in the Appendix.

Because we cannot implement an efficient decision rule with constraints, we consider the second-best social choice function $f = (d, \xi)$ such that $d(H, H) = (1, 1)$, $d(H, L) = d(L, H) = d(L, L) = (0, 0)$, $\xi_i(\theta) = d_i(\theta)c$ for $i = 1, 2$ in the next subsection.

The direct mechanism of the social choice function f is a good candidate for the inducement mechanism. However, we have to note that consumers are anonymous agents in many cases[8]. In that case, they can escape from the mechanism at any time. Because it usually takes time to gather information from agents and to produce the good, agents who send a message can escape from the mechanism before the principal supplies the good and makes them pay for it. Consider the case where the principal prepares reservation tickets and gives them free to those who want to buy the good in the future. In this case, some agents perhaps get reservation tickets without intending to buy the good, even if the principal says "You have to buy the good if you get a ticket and

[8]In some cases, we can know agents' identity. However, we consider a mechanism which can be used without agents' identity, because it is robust and relatively easy to introduce.

many other people also get tickets." The principal cannot force them to buy the good if they are anonymous agents. Thus, we have to impose a positive price for the reservation tickets. In the next section, we consider a particular type of reservation ticket with a positive price.

2.3 Refundable reservation tickets game

In order to overcome the problem in the original game, we suggest a mechanism with *refundable reservation tickets* (RRT), which has both costly signal and cheap talk characteristics. In this mechanism, a seller sells the RRT before the sale of the good and sells the good to the agents who bought the RRT. We add stage 0 to the original game as Fig 2 shows [9]. At stage 0, each subject is given two choices, "b" and "n"[10]. "b" represents buying the RRT and "n" represents not buying the RRT. The price of RRT is $\eta > 0$. If an agent buys the RRT and his/her partner does not, the seller refunds η. If both agents buy RRT, then they go into stage 1 and both agents choose "B" or "N" at stage 1. After that, they are notified of their partner's action at stage 1 and at stage 2 they choose "B" or "N" again. As in the original game, we assume that once they choose "B" at stage 1, they have to choose "B" at stage 2 again. Moreover we add the following refund policy[11]. The seller refunds η to an agent if he/she chooses "B" at stage 1. On the other hand, when he/she chooses "N" at stage 1, the seller does not refund η even if he/she chooses "B" at stage 2. We refer to this game as the RRT game.

[[Insert Figure 2 around here]]

[9] Instead of stages 0, 1 and 2, we use the more natural terms stages 1, 2, and 3 in the experiments.

[10] Instead of "b" and "n", we use the neutral terms, "Go to the next stage" and "Do not go to the next stage", in the experiment.

[11] This refund policy is not necessary in theory, however, we believe that it would help agents buy the good.

Because agents have to buy the RRT at a positive price, each agent will guess that type L agents will not buy the RRT. That is, buying RRT has the characteristic of a costly signal. However, buying RRT does not give a negative payoff when the partner does not buy the RRT. That is, the agents need not risk miscoordination regarding the purchase of RRT and the RRT game has the characteristic of cheap talk.

2.4 The solution of the RRT game

In this subsection, we discuss the theoretical solution of the RRT game informally. We discuss the formal model in the Appendix. We consider a solution concept that does not depend on beliefs about agents' types for robustness of the solution. Dominant strategy implementation is a desirable implementation concept. However, an equilibrium with dominant strategies does not exist in many cases and the RRT game does not have such an equilibrium. Thus, in this paper, we consider an implementation concept with iterative deletion of dominance. There are various notions of dominance such as ex ante, interim, and ex post. Of these, we use the most conservative concept, ex post dominance, which does not assume common knowledge about agents' beliefs[12]. If a strategy is ex post dominated, then it is ex ante and interim dominated for any full support beliefs.

We refer to *dominance solvable equilibrium* as the set of strategy profiles that remains after the iterative deletion of ex post weakly dominated strategies. We should note that the dominance solvable equilibrium may depend on the order of eliminations. Thus, as discussed in Chung and Ely (2000), we say that a mechanism *implements* a social choice function in *iteratively weakly undominated strategies* only if the dominance solvable equilibrium does not depend on the

[12] We cannot adapt the results with iterated deletion of *interim* weakly dominated strategies, e.g., Moulin (1979), Tomas (1994), Cabrales and Ponti (2000) and Abreu and Matsushima (1994), to our model.

order of eliminations. Using this implementation concept, a second-best social choice function is implementable with an RRT mechanism as follows.

Theorem 2.1. *Let a social choice function $f = (d, \xi)$ where $d(H, H) = (1, 1)$, $d(H, L) = d(L, H) = d(L, L) = (0, 0)$ and $\xi_i(\theta) = d_i(\theta)c$. Then the RRT mechanism implements f in iteratively weakly undominated strategies.*

Proof. If an agent is type L and buys the RRT, the payoff is zero at best and negative at worst. That is, the strategies "Buy the RRT at stage 0 when his/her type is L" are weakly dominated strategies. If an agent's type is H and his/her partner buys the good at stage 1, then buying the good at stage 2 gives a positive payoff. That is, not buying in this situation is a dominated strategy. Deleting these two behavioral strategies, the following strategy becomes a strictly dominant strategy: "If type is L, then do not buy the RRT at stage 0. If type is H, then buy the RRT at stage 0 and buy the good at stage 1." The strategy profile that both agents use the strategy becomes the dominance solvable equilibrium. Moreover, there is no other dominance solvable equilibrium. Therefore, the unique dominance solvable equilibrium is the profile that both agents use the strategy.

□

That is, we can estimate that coordination failure does not occur in equilibrium. When we consider the welfare of the RRT game, we should be careful about the payment for the RRT. In the RRT game, if an agent buys the RRT and does not buy the good at stage 1, then η is lost. Because η is transferred to the seller, we can include this transfer into the social welfare. From this, the welfare of the RRT can be determined by the last state of the game and it is the same as the original game. For example, if both agents are type H and buy the good by the end of stage 2, then the social welfare is $2\{v(2|H) - c\}$

regardless of at which stages they buy. Therefore, the expected social welfare in the equilibrium in the RRT game is $2q^2\{v(2|H) - c\}$.

3 Experimental procedures and theoretical predictions

3.1 Outline of the experiments

In each session, we employed 24 subjects who were volunteer undergraduate and graduate students at Osaka University[13]. After the subjects assembled, we randomly assigned seats by lottery. We instructed the subjects that the payoffs in the games were their rewards.

Each session had nine rounds. We assigned each subject to one of the types at the beginning of each round. We notified each subject which type he/she was assigned to and did not notify which types the other subjects were. After finishing the type decision, we made 12 groups of two subject members each. Two subjects played a game, which we called a round. A round comprised some stages. The original game had two stages and the RRT game had three stages.

At the beginning of each round, the subjects were reassigned to either type. They were told that in each round they would be randomly matched with another subject and that at no time in the experiment would they be rematched with the same subject.

3.2 Payoff of the original game

The payoff of the original game was determined as follows. At the beginning of each session, each of the subjects received 3,300 yen[14]. The payoff changed,

[13]We recruited more than 24 subjects and selected 24 subjects from them by lot.
[14]US$1 was about 110 yen. That is, 3,300 yen was about US$30.

depending on the choices in each round. Because of the irreversibility of the choice "B", there were three possible choices.

- Choose "B" at stage 1 and choose "B" at stage 2.
  We denote the choice by "BB".
- Choose "N" at stage 1 and choose "B" at stage 2.
  We denote the choice by "NB".
- Choose "N" at stage 1, choose "N" at stage 2.
  We denote the choice by "NN".

Table 1 and Table 2 are the payoff tables of each round where each row represents the subject's own choices and each column represents the partner's choices. We assume that the price of the good is equal to the marginal cost c and that the payoffs are the net payoff to subjects[15]. Then we have $v(2|H) - c = 150$ yen, $v(1|L) - c = -300$ yen, $v(2|L) - c = 40$ yen, $v(1|L) - c = -310$ yen.

[[Insert Table 1 and Table 2 around here]]

3.3 Payoff of the RRT game

In the RRT game, each subject received 3,300 yen at the beginning of each session. The price of the reservation tickets, η, was 50 yen. We have five possible choices in the RRT game.

- Choose "b" at stage 0, "B" at stage 1 and "B" at stage 2.
  We denote the choice by "bBB".
- Choose "b" at stage 0, "N" at stage 1 and "B" at stage 2.
  We denote the choice by "bNB".

[15] We can get similar results in this paper under different assumptions, if the assumptions do not change the order of the social welfare.

- Choose “b” at stage 0, “N” at stage 1 and “N” at stage 2.
  We denote the choice by “bNN”.

- Choose “n” at stage 0 and make no choice at stages 1 and 2.
  We denote the choice by “nNN”.

Table 3 is the payoff table for type H subjects and Table 4 for type L subjects[16]. In these payoff tables, each row represents the subject’s own choices and each column represents the choices of the partner.

[[Insert Table 3 and 4 around here]]

3.4 Type of environments

We had two environments, good environment and bad environment. On the one hand, in the good environment, 16 of the 24 subjects were assigned to type H and eight subjects to type L in each round. In the next round, the subjects were reassigned to either type. However, the total number of type H subjects was fixed to 16 and type L subjects to 8. In the bad environment, on the other hand, eight of the 24 subjects were assigned to type H and 16 subjects to type L in each round. The rules about the types were explained to the subjects. Therefore, these rules were common knowledge.

Because we have two environments and two game rules, we have four treatments. We denote the four treatments as in Table 5, where “O” represents the original game, “R” the RRT game, “good” the good environment and “bad” the bad environment.

[[Insert Table 5 around here]]

We conducted two sessions for each of the treatments O-good and R-good, and one session for the each of the treatments O-bad and R-bad. That is, the

[16] Payoffs in the tables include money from the refund of the reservation tickets.

total number of sessions was six.

Because there were 24 subjects in each session, we had 12 pairs of subjects in each round. There are three possible combinations of types: both of them are H, only one of them is H, or both of them are L. We denote these by (H,H), (H,L), (L,L) respectively.

We used combinations of types in Table 6 in one session for the treatment O-good and two sessions for the treatment R-good[17]. We used the combinations of types in Table 7 in one session for the treatment O-good[18]. On the other hand, we use the combinations of types in Table 8 in all two sessions under the bad environment.

[[Insert Table 6, Table 7 and Table 8 around here]]

3.5 Theoretical predictions

In this subsection, we assume that subjects are selfish profit maximizers. We obtain the theoretical predictions of the original game at first and then consider those of the RRT game. Let q be the probability that the partners' type is H on the condition that his/her type is H. In our treatments, $q = 15/23$ in the good environment, $q = 7/23$ in the bad environment. Suppose that type L subjects do not buy the good because their payoff from buying the good is inevitably negative. Then, the expected payoff to type H subjects from buying the good at stage 1 is at most $150q - 300(1 - q)$, which is negative in both environments.

Then from Fact 2, the theoretically predicted behaviors in the original game, if subjects are risk neutral or risk averse, are as follows.

- Type L subjects do not buy the good at any stage.

[17] We did not show the table to the subjects

[18] Although we planned to use the combinations of types in Table 6 in all of four sessions under the good environment, we used the combinations of types in Table 7 in one session for the treatment O-good because of programmer's mistakes. That is, the combination of types in the third round in Table 6 is moved to the last round in Table 7. However, we have not found any evidence of effects that these mistakes had on the experimental results.

- Type H subjects do not buy the good at stage 1. If their partners do not buy the good at stage 1, then they do not buy the good at stage 2. However, if their partners buy the good at stage 1, then they buy the good at stage 2.

On the other hand, regardless of the risk preference of the subjects, the behaviors in the RRT game predicted by the dominance solvable equilibrium are as follows.

- Type L subjects do not buy the RRT at stage 0 and do not go to stage 1.
- Type H subjects buy the RRT at stage 0 and buy the good at stage 1 if their partners buy the RRT at stage 0.

The theoretically expected welfares are zero in the original game and $2\{v(2|H)-c\} \times$ (the percentage of type H pairs) in the RRT game. The dotted lines in Figure 3 report the theoretically expected welfare. The circles represent welfare under the bad environment and the squares represent welfare under the good environment. As can be seen, the theoretically expected welfare of the RRT game is larger than that of the original game in each environment.

[[Insert Figure 3 around here]]

4 Experimental results

4.1 Comparison of welfare levels between the original game and the RRT game

In this section, we report our experimental results and compare them with the theoretical predictions.

We firstly consider whether RRT have the effect of increasing the welfare level. The solid lines in Figure 3 represent the social welfare associated with the experimental results. From the figure, we found that the average welfare of the RRT game is larger than that of the original game in each environment. That is, RRT have the effect of improving the welfare level not only in the theoretical prediction but also in the experimental results.

In order to compare welfare levels more strictly, let us consider the possible welfare level first. When both subjects in a group buy the good, there are three possible welfare levels, 300, 110 and -80 yen.[19] When only one of the subjects in a group buys the good, there are two possible welfare levels, -310 and -300 yen [20]. When neither of the subjects buys the good, there is only one possible welfare level, zero. Thus, we have six possible welfare levels, but there is no observation of -80 yen.

We used the exact combinations of types for each environment in the experiments. Therefore, in the same environment, we can compare welfare levels between the original game and the RRT game. Table 9 reports the welfare data of the RRT game and the original game. In the table, "cumulative number" describes the number of observations whose welfare level is less than, or equal to, the corresponding value. For example, the cumulative number for the original game at welfare 0 in the bad environment is 107, which means that there are 107 observations such that the welfare level in a group is less than or equal to zero. Figure 4 reports the cumulative distribution of the welfare. As we can see, any cumulative numbers in the original game are larger than corresponding numbers in the RRT game. From the table and figure, we make the following observation.

[[Insert Table 9 around here]]

[[Insert Figure 4 around here]]

[19] If both subjects in a group are type H, then it is 150 + 150 = 300. If one of them is type H, it is 150 - 40 = 110. If neither of them is type H, then it is -40 - 40 = -80.

[20] If the subject is type L, then the welfare is -310 + 0 = -310. If he/she is type H, then it is -300 + 0 = -300.

Observation 1: In each environment, the social welfare of the RRT game is significantly larger than that of the original game ($p < 0.01$ for the Mann Whitney's U test).

4.2 Comparison of the theoretical predictions and the experimental results

We next consider whether subjects buy the good in the RRT game. We use the notation about behaviors defined in section 3.3 to classify the behaviors of subjects in the RRT game. Furthermore, we add "bXX" to the categories of behaviors. "bXX" means that the subject buys the RRT at stage 0 and cannot go to the next stage because the partner does not buy the RRT. As we mentioned in subsection 3.5, "bBB" and "bXX" of type H subjects are consistent with the theoretical equilibrium behaviors in the RRT game. Furthermore, "nNN" of type L subjects is consistent with the theoretical equilibrium behavior. We refer to such behaviors, which are consistent with the theoretical prediction as *equilibrium behaviors.*

[[Insert Table 10 and Table 11 around here]]

Table 10 and Table 11 report the frequency of behaviors in the RRT game. In the tables, "*" indicates the equilibrium behaviors. The percentage of equilibrium behaviors is greater than 80% for all types in all environments. The numbers in the columns labeled "random" represent the expected frequencies assuming that subjects randomly choose an alternative at each stage with equal probability.

Observation 2: In the RRT game, the frequencies of the equilibrium behaviors are larger than those of the random behavior for each type in each environment. The differences are statistically significant ($p < 0.01$ for the goodness of fit).

We consider the details of the experimental results of the original game. Because the theoretically predicted behaviors of the original game depend on the partner's choice (see section 3.5), we use new notations. We use three letters to describe behaviors. The first letter describes the choice of a concerned subject at stage 1. The second letter describes the choice of his/her partner at stage 1. In order to distinguish the partner's choice from the concerned subject's choice, we enclose the partner's choice in brackets. The third letter describes the choice of the concerned subject at stage 2. For example, N[B]B means that the concerned subject does not buy the good at stage 1, his/her partner buys the good at stage 1 and the concerned subject buys the good at stage 2. Moreover, we use X to describe any choice. B[X]B means that the concerned subject buys the good at stage 1 and has to keep it regardless of his/her partner's behaviors.

[[Insert Table 12 and Table 13 around here]]

Table 12 and Table 13 report the frequency of each behavior and the percentage of equilibrium behaviors.

Observation 3: In the original game, the frequencies of the equilibrium behaviors are larger than those of the random behavior for each type in each environment. The differences are statistically significant. ($p < 0.01$ for the goodness of fit).

4.3 Learning process

As shown above, the experimental data fits the theoretical predictions. However, if it takes a long time for subjects to learn the equilibrium strategy, it is hard to adopt the mechanism in the real world. Figure 5 reports the percentage of equilibrium behavior in each round. From this, we can examine the learning process and make the following observation:

[[Insert Figure 5]]

Observation 4: The percentages of equilibrium behaviors do not appear to change systematically over time. That is, in our experiments, the subjects did not learn from experience.

4.4 The details of the experimental results in the treatment O-good

On the one hand, in the treatments R-good, R-bad and O-bad, more than 80% of the observed behaviors corresponded to the equilibrium behaviors. On the other hand, the percentage of the observed behaviors in the treatment O-good, which corresponded to the equilibrium behavior, was relatively smaller than that in the other three treatments. As we observed in observation 3, the experimental result in the treatment O-good was significantly different from the random behaviors. However, it is worthwhile studying the detailed data from this treatment.

In the treatment O-good, 39.2% (113 of 288) of type H subjects bought the good at stage 1. These B[X]B behaviors are inconsistent with the theoretical prediction. When a type H subject played B[X]B, he/she earned 42.5 yen on average in this round. On the other hand, when a type H subject played the equilibrium behaviors, i.e., N[B]N and N[N]N, he/she earned 56.4 yen on average in this round. The difference between the payoff from B[X]B and that from the equilibrium behaviors was small. Therefore, some subjects might consider the payoff from B[X]B and that from the equilibrium behaviors as being indifferent.

As we see in subsection 3.5, type H's expected payoff from B[X]B will be negative if type L subjects rationally choose "Do not buy" at any stage. That is, the irrational behaviors of type L subjects are needed for B[X]B's average payoff to be positive. In fact, 11.1% of type L subjects bought the good in the treatment O-good. Although the percentage of these irrational behaviors was not so large, such irrational behaviors of type L subjects might lead type H subjects to B[X]B.

We do not know why type L subjects bought the good despite their negative payoff. We consider the following two reasons. Firstly, some subjects might seek social optimality. Two subjects bought the good at stage 1 in all rounds regardless of their types. In the questionnaires that we conducted after the experiments, they provided the following reasons: "If everybody buys the good regardless of their types, then the total payoff will be maximized. So I bought the good at stage 1 regardless of my type." Secondly, some subjects seem to avoid negative payoffs for their partners. After their partners bought the good at stage 1, 16.1% of type L subjects bought the good at stage 2. Because type L subjects got a negative payoff from the choice, these behaviors may be interpreted as altruistic behavior for their partners.

5 Concluding Remarks

In experimental studies, costly signal and cheap talk are considered to be useful in overcoming coordination failure (Duffy and Feltovich (2002), Van Huyck et al. (1990)). We introduce the RRT mechanism, which has both of these characteristics.

Not only does the RRT mechanism have the characteristics of cheap talk and costly signal, but also there is a possibility that it is useful in coping with coordination failure for goods with network externalities. Therefore, whether the RRT mechanism works is a question of both theoretical and practical importance.

As we see in section 4, the results in the RRT game are clearly consistent with theoretical predictions. Because most of the subjects use the equilibrium behaviors from the first round, the learning process is not necessary for the experimental result. Moreover, RRT enhance the social welfare of the game in the experiment. That is, the RRT game is useful for selling the good with network externalities in our experiments.

If the principal tells a lie about his/her partner's choice, an unjustified profit results. Therefore, it is hopeful that the principal of the RRT mechanism is the government or a responsible company, which will be harmed from a bad reputation[21]. On the other hand, the RRT mechanism does not use explicit contract of agents, but uses specific reservation tickets. Therefore, the RRT mechanism does not need any additional enforcement rules about agents. That is, the mechanism can be used for anonymous agents. Because of this property, it is likely that the mechanism is relatively easy to introduce.

However, the number of members in a group is only two in our experiments. This can be a problem, because the number of members represents the number of potential consumers in theory. This experiment is the first step of the new mechanism, and experiments with more members would be appropriate for future work.

References

Abreu D., Matsushima H., 1994. Exact implementation. Journal of Economic Theory 64 (1), 1–19.

Cabrales A., Ponti G., 2000. Implementation, elimination of weakly dominated

[21] It is unnecessary for the seller of the good with network externalities to be the government. Even if the seller of RRT is the government and the seller of the good is a private company, the RRT mechanism will work.

strategies and evolutionary dynamics. Review of Economic Dynamics 3 (2), 247–282.

Cason T.N., Saijo T., Yamato T., Yokotani K., 2004. Non-excludable public good experiments. Games and Economic Behavior 49 (1), 81–102.

Chakravarty S., 2003a. Experimental evidence on product adoption in the presence of network externalities. Review of Industrial Organization 23 (3), 233–254.

Chakravarty S., 2003b. An experimental market with network effects and one sponsored seller. in Michael Baye, ed., Advances in Applied Microeconomics, Volume 12: Organizing the new industrial economy .

Chung K.S., Ely J.C., 2000. Efficient and dominance solvable auctions with interdependent valuations (1313).

Duffy J., Feltovich N., 2002. Do actions speak louder than words? an experimental comparison of observation and cheap talk. Games and Economic Behavior 39 (1), 1–27.

Etziony A., Weiss A., 2002. Coordination and Critical Mass in a Network Market–An Experimental Evaluation. Working paper, Department of Economics, Bar Ilan University, October, Xerox .

Farrell J., Saloner G., 1985. Standardization, Compatibility, and Innovation. The RAND Journal of Economics 16 (1), 70–83.

Gale D., 1995. Dynamic coordination games. Economic Theory 5, 1–18.

Gale D., 2001. Monotone games with positive spillovers. Games and Economic Behavior 37 (2), 295–320.

Katz M.L., Shapiro C., 1986. Technology adoption in the presence of network externalities. Journal of Political Economy 94 (4), 822–41.

Mak V., Zwick R., 2006. Betting on the trend: An experimental study on network externalities with continuous demand. working paper .

Matsushima H., 2007. Mechanism design with side payments: Individual rationality and iterative dominance. Journal of Economic Theory 127 (1), 1–30.

Moulin H., 1979. Dominance solvable voting schemes. Econometrica 47 (6), 1137–51.

Myerson R.B., Satterthwaite M.A., 1983. Efficient mechanisms for bilateral trading. Journal of Economic Theory 29 (2), 265–281.

Park I.U., 2004. A simple inducement scheme to overcome adoption externalities. Contributions to Theoretical Economics 4 (1).

Rauch J., 1993. Does history matter only when it matters little? the case of city- industry location. Quarterly Journal of Economics 108 (3), 843–867.

Rohlfs J., 1974. A theory of interdependent demand for a communications service. Bell Journal of Economics 5 (1), 16–37.

Saijo T., Yamato T., 1999. A voluntary participation game with a non-excludable public good. Journal of Economic Theory 84 (2), 227–242.

Schweizer U., 2006. Universal possibility and impossibility results. Games and Economic Behavior 57 (1), 73–85.

Shichijo T., 2007. A scheme to sell the good with network externalities. Mimeo.

Shichijo T., Nakayama Y., 2006. A two-step subsidy scheme to overcome network externalities in a dynamic game. mimeo .

Tomas S., 1994. Implementation in undominated nash equilibria without integer games. Games and Economic Behavior 6 (3), 502–511.

Van Huyck J.B., Battalio R.C., Beil R.O., 1990. Tacit coordination games, strategic uncertainty, and coordination failure. American Economic Review 80 (1), 234–48.

Van Huyck J.B., Battalio R.C., Beil R.O., 1991. Strategic uncertainty, equilibrium selection, and coordination failure in average opinion games. The Quarterly Journal of Economics 106 (3), 885–910.

Van Huyck J.B., Battalio R.C., Beil R.O., 1993. Asset markets as an equilibrium selection mechanism: Coordination failure, game form auctions, and tacit communication. Games and Economic Behavior 5 (3), 485–504.

6 Appendix

6.1 Proof of Fact 3

Fact 3 There exists no mechanism that satisfies ex post incentive compatibility, feasibility, the ex post participation constraints and efficiency of allocation.

Proof. If truth telling is ex post incentive compatible, $\pi_1(f(H,H)|H) \geq \pi_1(f(L,H)|H)$ and $\pi_1(f(H,H)|L) \leq \pi_1(f(L,H)|L)$. If the allocation is efficient, $d_i(H,H) = d_i(H,L) = d_i(L,H) = 1$ for all i. Therefore, we have $\xi_1(H,H) = \xi_1(L,H)$. On the other hand, from feasibility $\sum_i \xi_i(\theta) - \sharp d(\theta)c \geq 0$ for all $\theta \in \Theta$. That is, $\xi_1(H,H) + \xi_2(H,H) \geq 2c$. Therefore, $\xi_1(H,H) \geq c$ or $\xi_2(H,H) \geq c$. We consider the case $\xi_1(H,H) \geq c$. From the ex post participation constraint $v(2|L) - \xi_1(L,H) \geq 0$. From $\xi_1(L,H) = \xi_1(H,H)$ and $\xi_1(H,H) \geq c$, we have $v(2|L) - c \geq 0$. This contradicts our assumption in section 2.1. A contradiction obtains.

□

6.2 Formal definitions

In this section, we consider the formal definition of implementation discussed in Section 2.4. The implementation concept is a finite version of Chung and Ely (2000).

Let us denote the set of outcomes by $X \equiv \{0,1\} \times \{0,1\} \times \mathbb{R} \times \mathbb{R}$. Each element of X determines who gets the goods and how much they pay for the good. Let M be a finite action space. A mechanism consists of action space M and the function $g : M \to X$. The function $s_i : \Theta_i \to M$ is a strategy in the mechanism. Let S_i be the set of strategies for agent i and $S \equiv S_1 \times S_2$.

Definition 6.1. *We say that s_i is* ex post weakly dominated *within a given subset S^k of strategy profiles, if there exists $s'_i \in S_i^k$ such that $\pi_i(g(s(\theta)) \mid \theta_i) \leq \pi_i(g(s'_i(\theta_i), s_{-i}(\theta_{-i})) \mid \theta_i)$ for all $s_{-i}(\theta_{-i})$ with strict inequality for some $s_{-i}(\theta_{-i})$.*

Definition 6.2. *We say that a subset S^k of strategy profiles is internally undominated if there is no strategy $s_i \in S^k$ that is ex post weakly dominated within S^k.*

Definition 6.3. *We say that a strategy profile s^* is a dominance solvable equilibrium if there exists some sequence of strategy profile sets $\{S^1, \ldots, S^K\}$ such that $S^0 = S$, $S^{k+1} \subset S^k$, $\forall s_i \in S_i^k \setminus S_i^{k+1}$ is ex post weakly dominated within S^k, S^K is internally weakly undominated.*

Definition 6.4. *We say that a mechanism implements social choice function f in iteratively weakly undominated strategies, if, for every dominance solvable equilibrium s^*, $g(s^*(\theta)) = f(\theta)$.*

Note that the definition of implementation does not depend on the order of deletion of weakly dominated strategies.

7 Tables and Figures

Figure 1: The decision flow of the original game

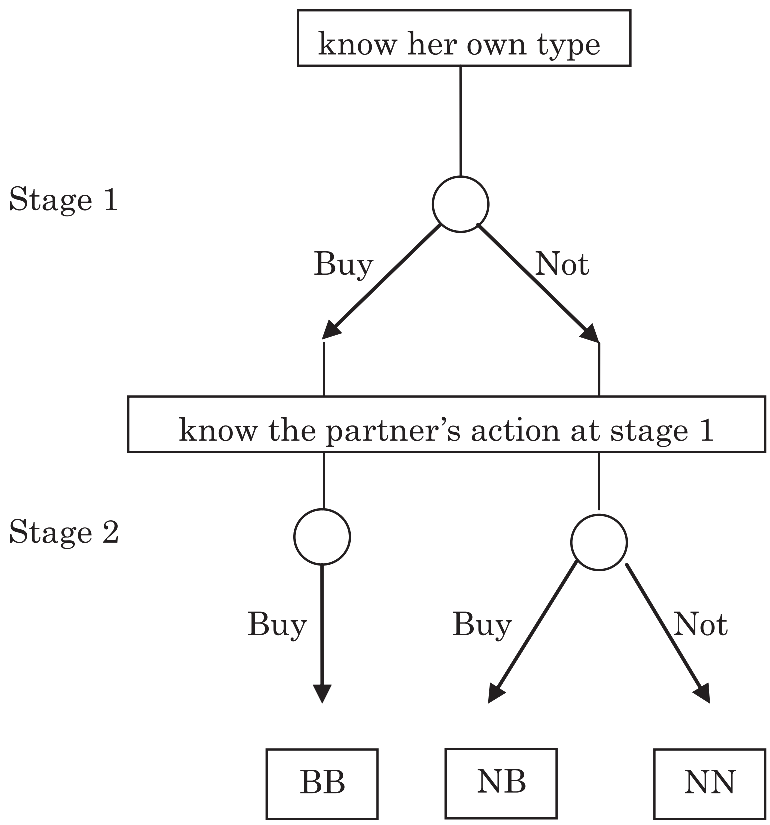

Figure 2: The decision flow of the RRT game

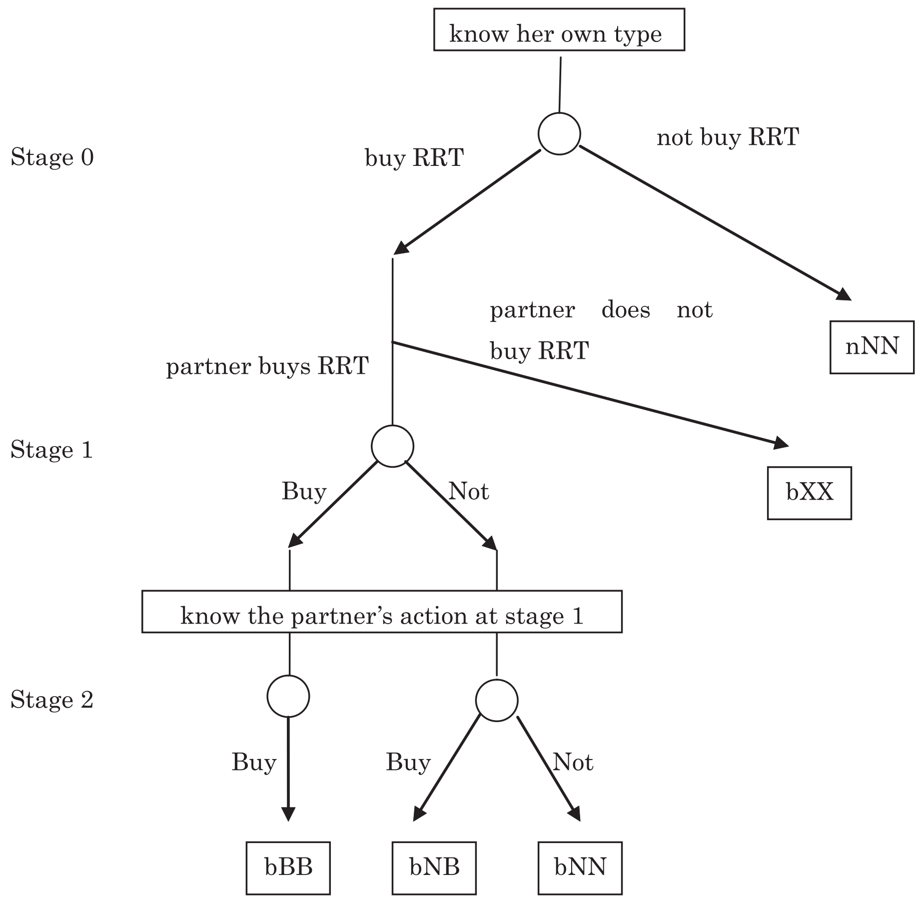

Figure 3: The average social welfare in the different treatments

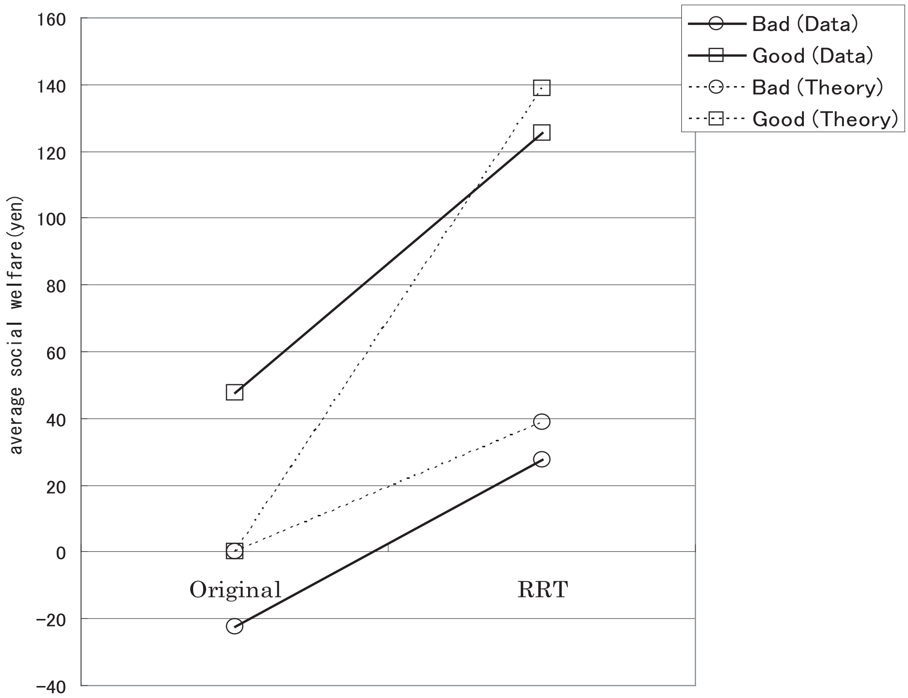

Figure 4: The cumulative data distributions of social welfare

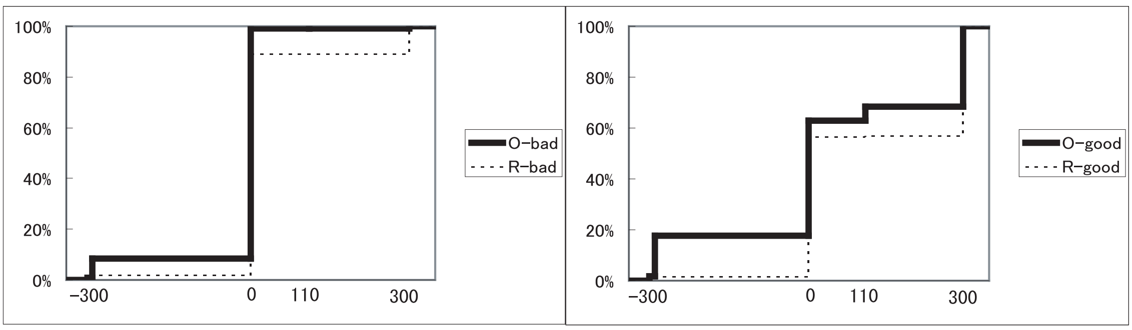

Figure 5: The percentage of behaviors which consistent with the theoretical prediction

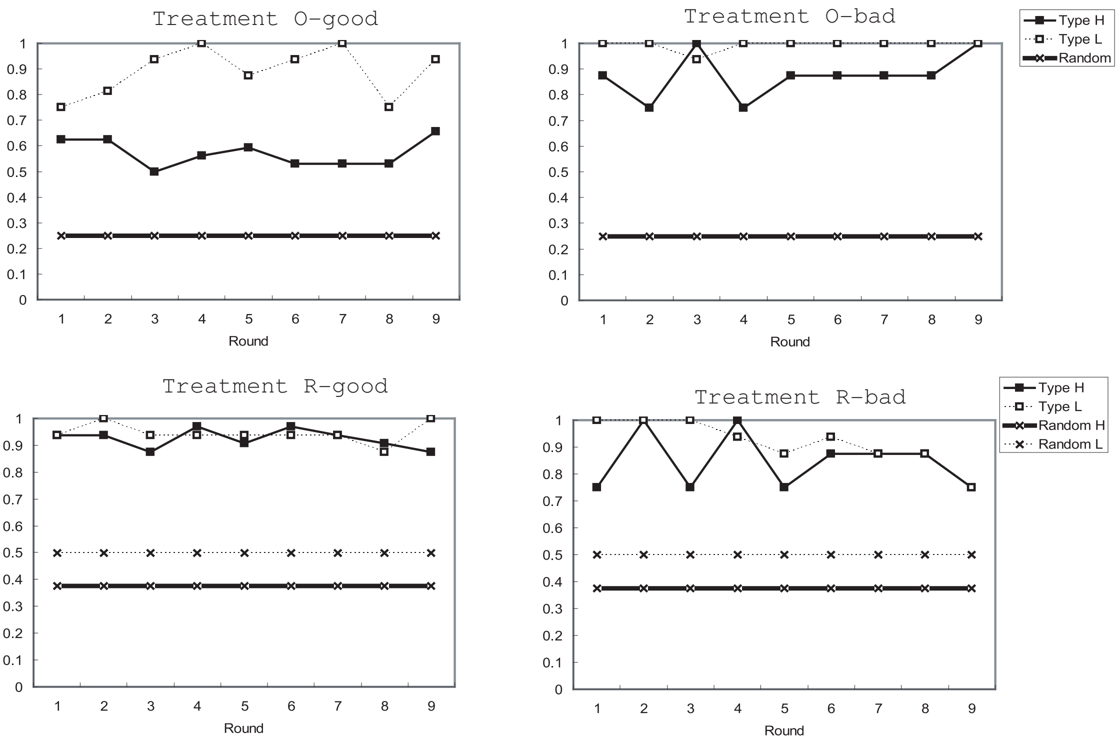

Table 1: The payoff matrix for type H in the original game

	BB	NB	NN
BB	150 yen	150 yen	-300 yen
NB	150 yen	150 yen	-300 yen
NN	0 yen	0 yen	0 yen

Table 2: The payoff matrix for type L in the original game

	BB	NB	NN
BB	-40 yen	-40 yen	-310 yen
NB	-40 yen	-40 yen	-310 yen
NN	0 yen	0 yen	0 yen

Table 3: The payoff matrix for type H in the RRT game

	bBB	bNB	bNN	nNN
bBB	150 yen	150 yen	-300 yen	0 yen
bNB	100 yen	100 yen	-350 yen	0 yen
bNN	-50 yen	-50 yen	-50 yen	0 yen
nNN	0 yen	0 yen	0 yen	0 yen

Table 4: The payoff matrix for type L in the RRT game

	bBB	bNB	bNN	nNN
bBB	-40 yen	-40 yen	-310 yen	0 yen
bNB	-90 yen	-90 yen	-360 yen	0 yen
bNN	-50 yen	-50 yen	-50 yen	0 yen
nNN	0 yen	0 yen	0 yen	0 yen

Table 5: The notations of the treatments

	Bad Env.	Good Env.
Original game	O-bad	O-good
RRT game	R-bad	R-good

Table 6: The frequency of combinations of types in two sessions of the treatment R-good and in one session of the treatment O-good.

Round	(H,H)	(H,L)	(L,L)	Total of H
1	4	8	0	16
2	5	6	1	16
3	7	2	3	16
4	5	6	1	16
5	6	4	2	16
6	6	4	2	16
7	5	6	1	16
8	6	4	2	16
9	6	4	2	16

Table 7: The frequency of combinations of types in a session of the treatment O-good

Round	(H,H)	(H,L)	(L,L)	Total H
1	4	8	0	16
2	5	6	1	16
3	6	4	2	16
4	7	2	3	16
5	5	6	1	16
6	6	4	2	16
7	6	4	2	16
8	5	6	1	16
9	6	4	2	16

Table 8: The frequency of combinations of types under the bad environment

Round	(H,H)	(H,L)	(L,L)	Total H
1	0	8	4	8
2	1	6	5	8
3	3	2	7	8
4	1	6	5	8
5	2	4	6	8
6	2	4	6	8
7	1	6	5	8
8	2	4	6	8
9	2	4	6	8

Table 9: The frequency of each social welfare level

	Bad environment					Good environment				
Welfare	-310	-300	0	110	300	-310	-300	0	110	300
Frequency										
Original	1	8	98	0	1	4	34	98	12	68
RRT	0	2	94	0	12	0	3	119	1	93
Cumulative number										
Original	1	9	107	107	108	4	38	136	148	216
RRT	0	2	96	96	108	0	3	122	123	216
Mann–Whitney p-value	< 0.01					< 0.01				

"cumulative number" describes the number of observations such that the social welfare is larger than or equal to the concerned welfare level.

Table 10: The frequency of each behaviors in the treatment R-good

Subject's Type	Type H				Type L			
Data Type	Experimental		Random		Experimental		Random	
bBB	175*	(60.8%)	36*	(12.5%)	1	(0.7%)	18	(12.5%)
bNB	15	(5.2%)	18	(6.3%)	0	(0.0%)	9	(6.3%)
bNN	0	(0.0%)	18	(6.3%)	3	(2.1%)	9	(6.3%)
bXX	91*	(31.6%)	72*	(25.0%)	4	(2.8%)	36	(25.0%)
nNN	7	(2.4%)	144	(50.0%)	136*	(94.4%)	72*	(50.0%)
Total	288	(100.0%)	288	(100.0%)	144	(100.0%)	144	(100.0%)
Equilibrium behaviors	266	(92.4%)	108	(37.5%)	136	(94.4%)	72	(50.0%)
p-value	< 0.01				< 0.01			

* indicates the equilibrium behaviors. The last row is the p-value of goodness of fit.

Table 11: The frequency of each behaviors in the treatment R-bad

Subject's Type	Type H				Type L			
Data Type	Experimental		Random		Experimental		Random	
bBB	22*	(30.6%)	9*	(12.5%)	0	(0.0%)	18	(12.5%)
bNB	4	(5.6%)	4.5	(6.3%)	0	(0.0%)	9	(6.3%)
bNN	0	(0.0%)	4.5	(6.3%)	2	(1.4%)	9	(6.3%)
bXX	39*	(54.2%)	18*	(25.0%)	10	(6.9%)	36	(25.0%)
nNN	7	(9.7%)	36	(50.0%)	132*	(91.7%)	72*	(50.0%)
Total	72	(100.0%)	72	(100.0%)	144	(100.0%)	144	(100.0%)
Equilibrium behaviors	61	(84.7%)	27	(37.5%)	132	(91.7%)	72	(50.0%)
p-value	< 0.01				< 0.01			

* indicates the equilibrium behaviors. The last row is the p-value of goodness of fit.

Table 12: The frequency of each behaviors in the treatment O-good

Subject's Type	Type H				Type L			
Data Type	Experimental		Random		Experimental		Random	
B[X]B	113	(39.2%)	144	(50.0%)	9	(6.3%)	72	(50.0%)
N[B]B	62*	(21.5%)	36*	(12.5%)	5	(3.5%)	18	(12.5%)
N[B]N	3	(1.0%)	36	(12.5%)	26*	(18.1%)	18*	(12.5%)
N[N]B	7	(2.4%)	36	(12.5%)	2	(1.4%)	18	(12.5%)
N[N]N	103*	(35.8%)	36*	(12.5%)	102*	(70.8%)	18*	(12.5%)
Total	288	(100.0%)	288	(100.0%)	144	(100.0%)	144	(100.0%)
Equilibrium behaviors	165	(57.3%)	72	(25.0%)	128	(88.9%)	36	(25.0%)
p-value	< 0.01				< 0.01			

* indicates the equilibrium behaviors. The last row is the p-value of goodness of fit.

Table 13: The frequency of each behaviors in the treatment O-bad

Subject's Type	Type H				Type L			
Data Type	Experimental		Random		Experimental		Random	
B[X]B	5	(6.9%)	36	(50.0%)	1	(0.7%)	72	(50.0%)
N[B]B	1*	(1.4%)	9*	(12.5%)	0	(0.0%)	18	(12.5%)
N[B]N	0	(0.0%)	9	(12.5%)	5*	(3.5%)	18*	(12.5%)
N[N]B	4	(5.6%)	9	(12.5%)	0	(0.0%)	18	(12.5%)
N[N]N	62*	(86.1%)	9*	(12.5%)	138*	(95.8%)	18*	(12.5%)
Total	72	(100.0%)	72	(100.0%)	144	(100.0%)	144	(100.0%)
Equilibrium behaviors	63	(87.5%)	18	(25.0%)	143	(99.3%)	36	(25.0%)
p-value	< 0.01				< 0.01			

* indicates the equilibrium behaviors. The last row is the p-value of goodness of fit.