Experimental Economics, 5:133–153 (2002)

Voluntary Participation and Spite in Public Good Provision Experiments: An International Comparison

TIMOTHY N. CASON
Department of Economics, Krannert School of Management, Purdue University, West Lafayette, IN 47907-1310, USA
email: cason@mgmt.purdue.edu

TATSUYOSHI SAIJO
Inst. of Social and Economic Research, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan
email: saijo@iser.osaka-u.ac.jp

TAKEHIKO YAMATO
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
email: yamato@valdes.titech.ac.jp

Abstract

This paper studies voluntary public good provision in the laboratory, in a cross-cultural experiment conducted in the United States and Japan. Our environment differs from the standard voluntary contribution mechanism because subjects first decide whether or not to participate in providing this non-excludable public good. This participation decision is conveyed to the other subject prior to the subjects' contribution decisions. We find that only the American data are consistent with the evolutionary-stable-strategy Nash equilibrium predictions, and that behavior is significantly different across countries. Japanese subjects are more likely to act spitefully in the early periods of the experiment, even though our design changes subject pairings each period so that no two subjects ever interact twice. Surprisingly, this spiteful behavior eventually leads to more efficient public good contributions for Japanese subjects than for American subjects.

Keywords: fairness, other regarding preferences, subject pools, culture

JEL Classification: D70, C90, H41

1. Introduction

Culture and national character have played a central role in explaining differences in business management and performance across countries, both in the popular press and in management research. Theoretical research in economics, however, almost universally ignores cultural differences, although recent laboratory research has identified the potential role of cultural norms in influencing economic outcomes—particularly in the context of bargaining (e.g., Buchan et al., 1999; Kachelmeier and Shehata, 1992; Roth et al., 1991). The cultural influence may arise through differences in *fairness* norms. For example, these norms could lead to preferences for roughly equal earnings for equal effort, often referred to as difference or inequality aversion (e.g., Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). Or if someone feels that they are being treated unfairly, this feeling may trigger a *spiteful* response; i.e., “a malicious desire to harm . . . another person.”[1] If this desire is triggered by harmful actions of the other person towards oneself, this spiteful response should be viewed as (negative) reciprocal behavior.

In this paper we examine spiteful behavior and cultural differences by studying voluntary public good provision, in experiments conducted in the United States and Japan. Our environment differs in three ways from the standard linear voluntary contribution mechanism (e.g., Isaac and Walker, 1988). First, we employ two-person groups. Second, the payoffs from the public good are nonlinear. Third, and perhaps most importantly, subjects decide whether or not to participate in funding this non-excludable public good in an initial stage, and this participation decision is conveyed to the other subject prior to the subjects’ contribution (or “investment”) decisions. In other words, a non-participation decision is essentially an announced commitment to free-ride on the other’s contribution.[2] This two-stage structure allows reciprocity-motivated subjects to punish non-participation decisions by their opponents. We find that Japanese subjects are significantly more likely to punish their opponents for non-participation, even though our design changes subject pairings each period so that no two subjects ever interact twice. Surprisingly, this spiteful behavior eventually leads to more efficient outcomes for Japanese subjects than for American subjects.

Laboratory researchers often study reciprocal behavior using the ultimatum game. In the ultimatum game a proposer offers a split of some surplus to a respondent, and the respondent can either accept the proposed allocation, or reject it so that both players earn nothing. Although the ultimatum game might be the simplest environment to study spite, the public goods game that we employ is interesting because spiteful behavior can transform the incentives facing players. For example, we show that the spiteful behavior exhibited by Japanese subjects in the second stage contribution subgame following non-participation leads to average monetary payoffs that differ substantially from the Nash equilibrium. In fact, these realized payoffs make participation a dominant strategy in the first stage even for those subjects who care only about their own monetary payoffs.[3] By contrast, participation is not a dominant strategy in the first stage based on realized payoffs for American subjects, because they are less likely to act spitefully in the second stage.

While such cultural and nationality differences are inconsistent with the standard income-maximizing objective usually assumed in economics, such differences have been well documented in psychology, sociology and anthropology. A few researchers in economics and management have also studied cultural differences. For example, Buchan et al. (1999) present a detailed review of the literature on cultural differences in bargaining and in psychology, highlighting how this research suggests that culture influences fairness norms. They then provide laboratory evidence from ultimatum bargaining that supports the hypothesis that Japanese subjects prefer more equal earning distributions than American subjects prefer.[4] Their findings are consistent with the interpretation in our experiment that “unfair” non-participation decisions are more likely to invoke a spiteful response by Japanese subjects who chose to participate in funding the public good.[5]

Although Saijo and Nakamura (1995) and others document substantial spiteful behavior by Japanese subjects, not all evidence supports the hypothesis that Japanese subjects act more spitefully than their western counterparts. In their ultimatum game study, Roth et al. (1991) find that Japanese subjects' rejection rates (controlling for the offer amount) are no higher and are sometimes lower than American subjects' rejection rates. Brandts et al. (1999) do not find significant differences in public goods contributions across two European countries, the United States and Japan, and Okada and Riedl (1999) do not identify significant differences in behavior for a coalition formation game conducted in Japan and Austria.

One possible source of the mixed results in previous economics experiments could be other differences in university subject pools. In nearly all previous experimental economics research that studies cultural differences, subjects are recruited from one university in each country. We believe that it is important to determine whether between-country differences are greater than within-country differences in order to identify cultural or nationality effects confidently. We therefore collect data at two universities within each country. Our results indicate that nationality differences have a greater influence on outcomes than within-country university subject pool differences.

Of course, the mixed results regarding significant cultural differences in economics experiments could also be due to differences in experimental designs. For example, as noted above Brandts et al. (1999) do not identify significant cultural differences in a voluntary public goods contribution experiment, whereas we do identify differences in spiteful behavior between Japanese and American subjects. The two experiments differ in multiple dimensions, so one can only speculate regarding the source of the differences in observed outcomes. Brandts et al. (1999) employ a "contribution function" approach in which subjects make ten different contributions each period in a linear public goods environment, with each contribution corresponding to a different marginal per capita return (MPCR) to public goods contributions. Their public good is funded in a four-player single stage game, and they identify spiteful behavior as less than full contributions for the cases where the MPCR makes full contribution a dominant strategy, as in Saijo and Nakamura (1995). In all cases they consider, either full contribution or zero contribution is a dominant strategy.

By contrast, in the present experiment no contribution level is a dominant strategy because the public good environment is nonlinear, and subjects make a single contribution decision each period. But even more importantly, in our study subjects interact in two-player groups and we employ a two stage game with an initial participation decision, and we identify spiteful behavior through non-Nash contribution decisions in response to nonparticipation decisions of one's opponent. Thus, the two stage, two-player game in the present study makes it easier to identify spiteful behavior than in the one stage, four-player game of Brandts et al. (1999). Other differences between the two studies could also explain the differences in results.[6]

2. Experimental environment, design and hypotheses

2.1. Environment

In the first stage of our public goods provision game, subjects simultaneously choose whether or not they participate in the provision mechanism. In the second stage, knowing the other subject's participation decision, subjects who selected participation in the first stage choose contributions to the public good.

Two subjects, a and b, may each fund the public good and subject $i(=a, b)$ has w_i units of initial endowment of a private good. Each subject who participates in funding the public good must allocate w_i between her own consumption of the private good (x_i) and her public good investment (y_i). From the total public good 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 non-excludable public good available for consumption by each subject is the sum of the investments of two subjects and the initial level of the public good. Note that non-participants who choose not to fund the public good still consume it, as is the usual case for the voluntary contribution mechanism. Each subject's decision problem is

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

where $u_i(x_i, y)$ is subject i's payoff function. We use an identical Cobb-Douglas type payoff function to transform contributions toward the public good and consumption of the private good into each subject's payoffs: $u_i(x_i, y) = x_i^{\alpha} y^{1-\alpha}$, where $\alpha \in (0, 1)$. Using a monotonic transformation, we employ the following payoff function:

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

Our experiment sets $(w_a, w_b, w_y) = (24, 24, 3)$, $\alpha = 0.47$, and $\beta = 4.45$. With these parameters, if both subjects participate in funding the public good the Nash equilibrium investment pair of the game is $(\hat{y}_a, \hat{y}_b) = (7.69, 7.69)$, and the payoff level 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 $(y_a^*, y_b^*) = (12.02, 12.02)$. 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 contributions is less than the Pareto efficient level of the public good, which is the standard problem of this provision mechanism.

The situation just described represents the case in which subjects have already committed to participate during stage 1. However, Saijo and Yamato (1999) demonstrate that a wide class of mechanisms for funding public goods exists in which subjects have incentives not to participate, including this mechanism based on voluntary contributions. Therefore, this mechanism is not *voluntary* from the viewpoint of participation incentives.

Consider now the following two-stage game. In the first stage, subjects simultaneously decide whether or not to participate in funding the public good through this mechanism. In the second stage, subjects decide how many units of their initial endowment to invest after learning the other subject's participation decision. Notice that non-participation is different from zero investment with participation. Once a subject decides to participate in the mechanism, his opponent must choose her investment *without* knowing the other subject's investment. But if a subject chooses non-participation, then his opponent knows that he invests nothing.

In our experiment, subjects choose integer investments. If both subjects decide to participate in the mechanism, then the Nash equilibrium of that subgame is for each subject to

Table 1. First stage participation payoffs based on Nash equilibrium investments in the second stage.

	Player 2	
Player 1	Participate (p_2)	Not participate (1-p_2)
Participate (p_1)	7345, 7345	2658, 8278
Not participate (1-p_1)	8278, 2658	706, 706

Notes: Payoffs are based on Nash equilibrium second-stage investments of (8, 8) when both participate, (0, 11) and (11, 0) when only one participates, and (0, 0) when neither participates. The subgame perfect Nash equilibria of this initial stage participation game are participation 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).

contribute 8, and each earns 7345. No other Nash equilibria arise due to our use of a discrete strategy choice set. If one subject participates in the mechanism and the other does not, then the participant maximizes her payoff at $y_i = 11$ and earns 2658. The non-participant invests nothing and earns 8278. If both choose not to participate in the mechanism, both subjects receive 706. These subgame payoffs are displayed in the first stage normal form game payoff table shown in Table 1.

The game in Table 1 is a version of the well-known Hawk-Dove game, sometimes referred to as "chicken." Although the usual representation of the public good provision problem is a Prisoners' Dilemma game, the proper representation is a game of chicken once participation in the funding mechanism is a choice variable. Two pure strategy Nash equilibria exist: in one player 1 participates in funding the public good while player 2 opts to not participate, and in the other player 2 participates and player 1 does not. In the mixed strategy Nash equilibrium, each subject i chooses 0.68 as her participation probability p_i. Among these three equilibria, the mixed strategy equilibrium is the unique evolutionary-stable-strategy (ESS) equilibrium.[7]

Note that the subgame with one participant and one (announced) free-rider is similar in some respects to the sequential best-shot game (Harrison and Hirshleifer, 1989) although in our game the sequence of contribution decisions is endogenous. In the best-shot game, the total public good contribution is the maximum of the two players' contributions. In equilibrium the first player contributes zero—analogous to the non-participant's announced contribution of zero in the present game. The second player—like the participant in our game—then makes a contribution that will benefit both players. Harrison and Hirshleifer (1989) find support for the Nash equilibrium of this game, which as in the present game involves substantially higher payoffs for the free-rider than for the contributor. Prasnikar and Roth (1992) and Duffy and Feltovich (1999) show that this support for equilibrium play extends to complete information conditions. All of these studies use American subjects, and in the role of the second player these subjects exhibit very little spiteful behavior (punishment) when facing a zero contribution by the first player.[8] This finding is consistent with the non-spiteful behavior that we observe in our American data. Andreoni et al. (2002) present a two-player sequential summation public goods game, however, in which subjects frequently punish the other player. In this game the amount of the public good provided equals the sum of the players' individual contributions (as in our game), and the first player can commit to a low

Table 2. Summary of five laboratory sessions.

Session name	University	Country
Tsukuba	Univ. of Tsukuba	Japan
USC[a]	Univ. of So. Calif.	United States
Tokyo	Tokyo Metro. Univ.	Japan
Tokyo'	Tokyo Metro. Univ.	Japan
Purdue	Purdue University	United States

Notes: Each session employed 20 subjects for 4 practice periods and 15 actual periods. Subjects were randomly re-paired each period.
[a]One subject in the USC session misunderstood the instructions, and her choices are removed from the data prior to the analysis.
[b]The instructions wording in the Tokyo session differed slightly from the other 4 sessions as a robustness check. In the Tokyo session the phrase "the person you are paired with" replaced the phrase "your opponent" everywhere in the instructions, record sheets, questionnaires and payoff tables.

contribution. Unlike the best-shot game, such low contributions in the summation game are often punished by the second players, which leads the first players to deviate from the Nash equilibrium and contribute. Although Andreoni et al. (2002) use American subjects, this behavior is (qualitatively) similar to the spiteful behavior we observe in our Japanese data.

2.2. *Design and procedures*

We conducted three sessions in Japan and two sessions in the United States. As shown in Table 2, we conducted sessions at two different universities in each country.[9] Twenty subjects participated in each session for a total of 100 separate subjects. Each subject was randomly paired with each other subject one at a time—a so-called "strangers" design. The same game was repeated for 19 periods, 4 for practice and 15 for monetary reward, so as not to pair the same two subjects more than once. No subject had prior experience in a public good provision experiment.

The twenty subjects in each session were seated at desks in a relatively large room, and each had a randomly chosen identification number. These identification numbers were not displayed publicly. Each subject received an experimental procedure sheet, a record sheet, payoff tables, 15 investment sheets, and 4 practice investment sheets. Instructions were given by tape recorder to minimize the interaction between subjects and experimenters.

In each period we made ten pairs out of the twenty subjects. The pairings were anonymous and were determined in advance. Before choosing their investment, subjects decided whether or not they would participate in funding the public good. These decisions were collected by experimenters and then redistributed only to the paired subjects. No information (such as the total number of participants) or decisions were publicly announced. After this redistribution of the participation decisions, subjects who decided to participate in the mechanism chose their investment on an investment sheet by circling an integer between 0 and 24. In order to not reveal the number of participants and to obscure the participants' identity, non-participants also filled out this investment sheet, circling the phrase "Not Participate."

Experimenters collected these investment sheets and then redistributed them only to the paired subjects. During the redistribution, subjects were asked to fill out the reasons why they chose these numbers. After the redistribution, subjects calculated their payoffs from the payoff tables. Then the next period started.

It was common knowledge that every subject had the same payoff function. We distributed three kinds of payoff tables to minimize the likelihood of any possible misunderstanding, and all of the 100 subjects were able to readily calculate their payoffs following the instructions and practice periods.[10] Table 3 is the detailed payoff table provided to subjects: the rows are for the subject's own investment numbers and the columns are for the opponent's investment numbers. We also presented a rough payoff table summarizing average payoffs for sets of 9 or 12 payoff cells shown in Table 3, as well as an iso-payoff map. Most subjects indicated in their post-experiment questionnaire that they used the detailed payoff table (Table 3) only. We gave subjects three minutes to study these three payoff tables before the practice rounds and ten minutes to study them before the real rounds. The payoff function and tables used for practice and real rounds were different.

The sessions in Japan were conducted in Japanese, and the sessions in America were conducted in English. The instructions and forms were translated from Japanese to English by the two bilingual co-authors. The exchange rate used to translate payoffs from Japan to America was $1 = 100 Yen = 3030 experimental points.[11] Sessions required approximately 2 hours to complete. The mean payoff per subject was $26.75, the maximum payoff was $38.75, and the minimum payoff was $12.25.

As mentioned in the introduction, our use of two university subject pools in each country is an important feature of the design, given our objective to study cultural differences. Differences in results for any two universities could be due to subject pool effects unrelated to culture and nationality. To establish a significant cultural difference one must show that *between-country* differences are greater than *within-country* differences. We should also highlight the fact that there exist important similarities between sets of universities across countries. Purdue University and the University of Tsukuba both have a major emphasis on engineering and science and are both in (relatively) small "college towns" with predominantly university-resident students. By contrast, Tokyo Metropolitan University and the University of Southern California are both situated in major urban centers with many off-campus "commuter" students. Subject pool differences other than nationality and culture are therefore substantially lower within these sets of universities, and these other subject pool differences are considerably greater within countries.

2.3. Hypotheses

Our first task is to determine if behavior differs across university subject pools within the two countries. After establishing that within country differences are minor and usually insignificant, we then test the following null *cultural hypotheses*.

Hypothesis 1. The participation rate is equal across countries.

Hypothesis 2. The mean investment per subject is equal across countries, (a) conditional on one subject participating and (b) conditional on both subjects participating.

Table 3. Detailed payoff table provided to subjects.

Your payoff	Your investment number 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
Your opponent's investment number																									
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

The second set of *equilibrium hypotheses* is based on the theoretical discussion above. Because we reject Hypotheses 1 and 2 in favor of significant country effects, we test the following hypotheses separately for each country.

Hypothesis 3. The rate of participation is 68 percent, corresponding to the mixed strategy ESS participation rate.

Hypothesis 4. (a) Conditional on one subject participating, the mean investment for the participating subject is 11 units; and, (b) conditional on both subjects participating, the mean investment per subject is 8 units.

Finally, we test whether the total surplus realized in this public good provision mechanism is different across countries.

Hypothesis 5. The overall efficiency is equal across countries.

3. Results

3.1. Preliminaries—Within-country differences

Result 1. *The participation rates and the mean investments conditional on one subject participating are rarely different across universities within countries. The mean investments conditional on both subjects participating are sometimes statistically different across universities within countries, but these differences are typically less than one investment unit.*

Support. We compare behavior across sessions using random effects models of the participation decisions with a probit model, as well as separate investment choice models for the subgames with two participants and with one participant.[12] We estimate separate regressions for each pair of sessions, and include session dummy variables. When these dummy variables are not significantly different from zero we conclude that behavior does not differ across sessions. For the four within-country session comparisons (USC versus Purdue, Tsukuba versus Tokyo, Tsukuba versus Tokyo', and Tokyo versus Tokyo'), we find

- participation rates are never significantly different (t-statistics are always less than one);
- investment levels with one subject participating are only significantly different between the Tsukuba and Tokyo sessions ($t = 2.20$) and are not significantly different in the three other pairwise comparisons; and
- investment levels with both subjects participating are sometimes statistically different across sessions due to a low within-subject variance, but mean investments across all five sessions are quite similar (7.40 in Tsukuba, 7.00 in USC, 6.77 in Tokyo, 7.61 in Tokyo', and 8.07 in Purdue).

Cason et al. (1998) presents the details. We therefore conclude that differences within countries are minor, and turn our attention to the more significant differences across countries.

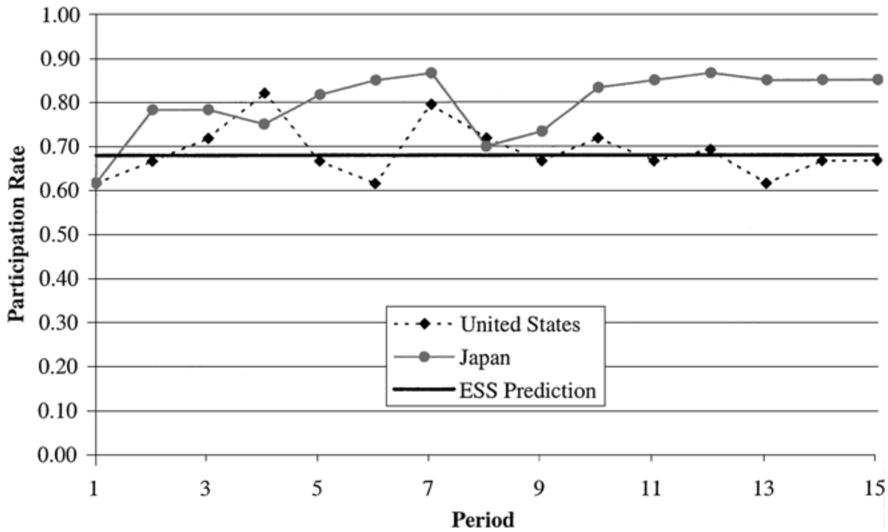

Figure 1. Participation rates.

3.2. Between country comparison

Hypothesis 1. The participation rate is equal across countries.

Result 2. *Hypothesis* 1 *is rejected in the final third of the sessions.*

Support. Figure 1 presents the participation rate by period, pooled across sessions within each country. In the first half of the sessions the participation rates are often similar across countries, but in the final third of the sessions the American participation rate is always below the Japanese participation rate.

A pooled random effects probit model strongly rejects the hypothesis of no country differences ($t = 3.05$). We also compared participation frequencies across countries for each individual period using Fisher's exact test. This nonparametric test rejects the null hypothesis of equal participation rates at the five-percent significance level in periods 6, 11, 12, 13, 14 and 15.[13]

Hypothesis 2. The mean investment per subject is equal across countries, (a) conditional on one subject participating and (b) conditional on both subjects participating.

Result 3. *Hypothesis* 2(a) *is rejected but Hypothesis* 2(b) *is not.*

Support. Figure 2 presents mean investments by the participating subject when only one subject participates, along with standard error bands. This figure indicates that investments with one participant are lower in the Japanese sessions, by about two units on average. The

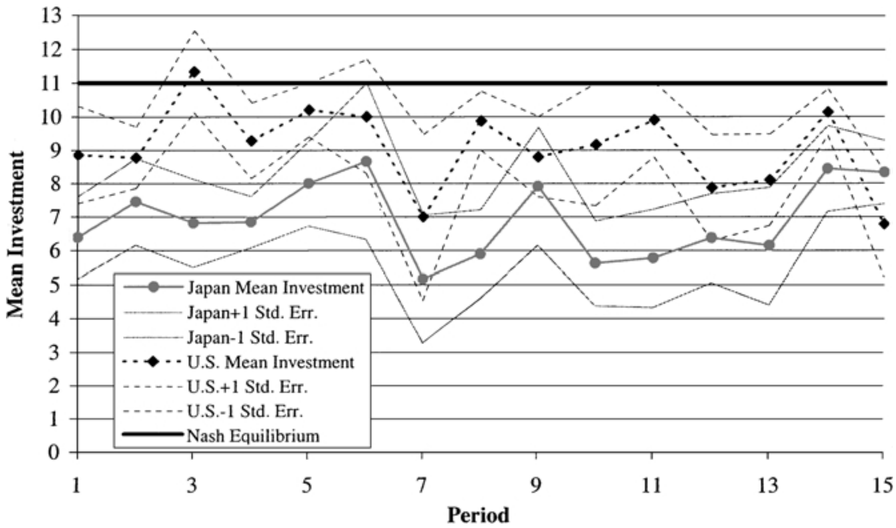

Figure 2. Mean investment when only one subject participates.

data strongly reject Hypothesis 2(a) when pooling across periods with a random effects model ($t = 4.56$). Moreover, a nonparametric Wilcoxon signed rank test conducted period-by-period rejects Hypothesis 2(a) at the five-percent level in periods 3, 8 and 11.

Figure 3 presents mean investments when both subjects participate. These data fail to reject Hypothesis 2(b), indicating no systematic differences in investments across countries

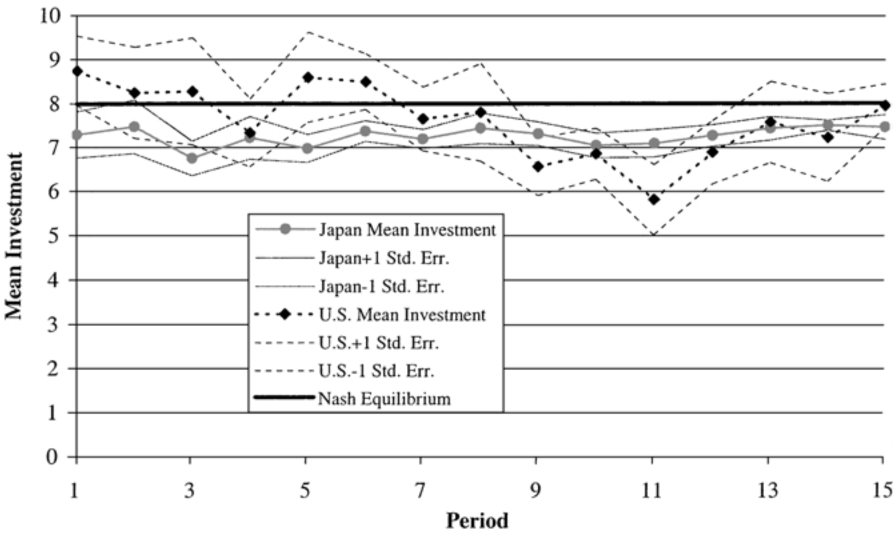

Figure 3. Mean investment when both subjects participate.

when both subjects participate (pooled random effects model $t = 1.89$). Taken together, the rejection of Hypothesis 2(a) and the failure to reject Hypothesis 2(b) suggest that the differences in behavior across countries arise mainly from a differing propensity to employ a negative reciprocal (i.e., spiteful) response to an opponent's non-participation.

3.3. Equilibrium hypotheses

Hypothesis 3. The rate of participation is 68 percent, corresponding to the mixed strategy ESS participation rate.

Result 4. *Hypothesis* 3 *is not rejected in the American data*, *but it is rejected in the Japanese data.*

Support. Figure 1 shows that the participation rate in the Japanese data exceeds the mixed strategy ESS prediction of 68 percent in every period after period 1. The participation rate in the American data fluctuates around this mixed strategy equilibrium, exceeding the equilibrium rate in about one-half of the periods. We test the null hypothesis that the participation rate is 68 percent using a binomial test, with individual subjects as the unit of observation. The Japanese sample has 60 subjects. Under the null hypothesis that the participation rate is 68 percent, the probability of observing 48 or more participation decisions out of 60 (80 percent) is less than five percent. The Japanese participation rate exceeds this critical 5-percent threshold in periods 5 through 7 and periods 10 through 15. The American sample has 39 subjects (recall that we omit the confused subject in the USC ssssion). Under the null hypothesis that the participation rate is 68 percent, the probability of observing 32 or more participation decisions out of 39 (82 percent) is less than five percent. The American participation rate exceeds this critical value only in period 4.

Tests of a mixed strategy equilibrium based on aggregate choice frequencies may mask differences across individuals' choice frequencies that may or may not be consistent with a mixed strategy equilibrium (Brown and Rosenthal, 1990). We therefore also examined the overall participation rates separately for each of the 99 subjects. The mean participation rate among the 60 Japanese subjects is 0.80 (12 of 15 decisions), and the median participation rate is 0.87 (13 of 15 decisions). The corresponding mean and median participation rates for the 39 American subjects are 0.69 and 0.73. Fifteen of the 60 Japanese subjects (25%) and 5 of the 39 American subjects (12.8%) were apparently using a pure strategy, as they participate in 15 out of 15 periods. Note that the ESS rate of 0.68 implies on average slightly more than 10 participation decisions. Only 14 of the 60 Japanese subjects (23.3%) participate 10 times or less, while the other 46 Japanese subjects (76.6%) participate 11 times or more. Using the 60 separate subject observations, the Japanese data reject the ESS prediction of 0.68 at better than the 0.0001 significance level using the non-parametric Wilcoxon signed-rank test. The frequency of participation rates below and above the ESS prediction is more equal among the 39 American subjects—17 (43.6%) below and 22 (56.4%) above. Consequently, a Wilcoxon test does not reject the ESS participation rate in the American data (Wilcoxon p-value = 0.61).

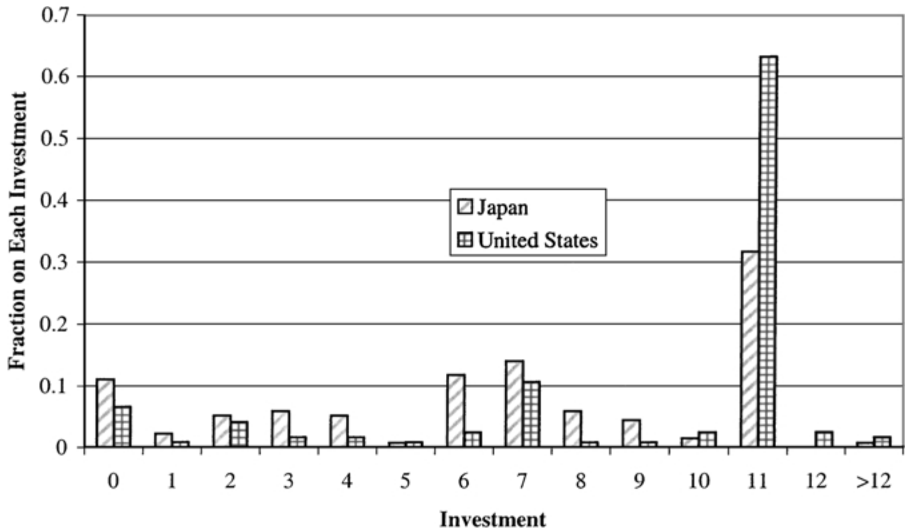

Figure 4. Distribution of investments when only one subject participates.

Hypothesis 4. (a) Conditional on one subject participating, the mean investment for the participating subject is 11 units; and, (b) conditional on both subjects participating, the mean investment per subject is 8 units.

Result 5. *Hypothesis* 4 *is not rejected in the American data but it is rejected in the Japanese data*, *although the evidence against Hypothesis* 4(b) *is relatively weak in the Japanese data.*

Support. Figure 2 shows that the average investment with one subject participating falls below the prediction of 11 for both countries, but investments are lower in the Japanese data (Result 3). Figure 4 displays the frequency distribution of investments with one subject participating, pooled over all periods. Although the modal choice is 11 in both countries, American subjects choose 11 about twice as frequently as Japanese subjects do. Over all 136 investments with one participant in the Japanese sessions, 43 (32%) were 11; over all 122 investments with one participant in the American sessions, 77 (63%) were 11. Period-by-period Wilcoxon signed rank tests reject the null hypothesis of 11 at the five-percent level for the Japanese data–in–periods 1 through 4, 8 and 10 through 12. This same test never rejects the null in the American data. The sample size is slightly smaller for these tests in the American data (an average of 8.1 observations per period for the American data versus an average of 9.1 for the Japanese data), but this lower power is unlikely to be the main reason that the American data fail to reject Hypothesis 4(a).

Figure 3 presents the average investment when both subjects participate. The American data reject the Nash prediction of 8 at the five-percent level only in periods 9, 10 and 11 (Wilcoxon test). The Japanese data reject the Nash equilibrium using this same test in all periods *except* periods 1 and 4. Nevertheless, compared to the case of one participant (figure 2), the Nash equilibrium has substantial drawing power overall. We therefore conclude that Hypothesis 4(b) is not rejected economically in both countries, since the mean investments differ from the Nash equilibrium by less than one investment unit in both countries (7.58 in the U.S. and 7.26 in Japan), and between one-half and two-thirds of the investments are within one unit of the Nash equilibrium in both countries.[14]

3.4. Efficiency

Hypothesis 5. The overall efficiency is equal across countries.

Result 6. *Hypothesis* 5 *is rejected in the final third of the sessions.*

Support. Efficiency is defined as the percentage of the maximum available earnings realized by subjects. Figure 5 presents the average efficiency for the Japanese and American sessions by period. If both subjects participate and choose the Pareto optimal investment of 12, they each earn 9090.[15] If both subjects participate and choose the Nash investment of 8, they each earn 7345, for an efficiency of 7345/9090 = 81 percent. This is displayed on the figure as a horizontal dashed line. The horizontal solid line on the figure displays the predicted efficiency of the ESS equilibrium, 5829/9090 = 64 percent.

The average efficiency differences in the two countries are not significant at the five-percent level in any individual period using a Wilcoxon test. Figure 5 suggests, however,

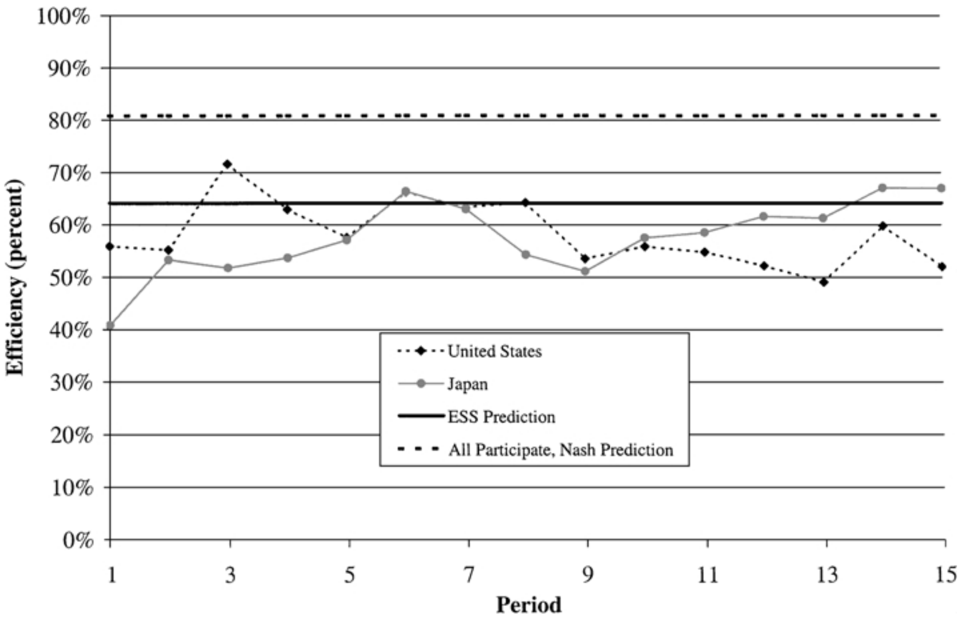

Figure 5. Average efficiency.

that efficiency in the Japanese data begins to exceed efficiency in the American data toward the end of the sessions. This is due to the greater participation among Japanese subjects (cf. figure 1). When pooling periods into the first, middle and final thirds of the session, a Wilcoxon test rejects the hypothesis that efficiency is equal across countries, but only in the final third of the session (p-value < 0.01).[16]

4. Summary and interpretation

The overall pattern of our results can be summarized as follows. The American data are roughly consistent with the mixed strategy ESS equilibrium of this two-stage game (Results 4 and 5). The Japanese data generally reject this equilibrium. Behavior of Japanese and American subjects is typically significantly different (Results 2 and 3). Relative to their American counterparts, Japanese subjects tend to participate more and invest less when only one subject of the pair chose to participate. This presents an important puzzle: What is different between Japanese and American subjects that could explain their differences in behavior? As discussed in the introduction, differences in the propensity to act spitefully can explain these differences across countries.

To see how these differences arise, consider the initial 5 periods of play. Participation rates are similar across countries for these initial periods (figure 1). However, the investment by the participating subject when only one subject participates is substantially higher in the American data than in the Japanese data for these initial periods (figure 2). That is, the Japanese subjects appear much more willing to "punish" their opponent for not participating. By investing, say, 7 instead of the best response of 11, the participating subject reduces her payoff from 2658 to 2210, a difference of 448. This spiteful behavior reduces the non-participating subject's payoff from 8278 (if his opponent invests the best response of 11) to 4018 (if his opponent instead invests 7), a difference of 4260. In this environment, a spiteful subject can sacrifice only a small amount to punish her opponent severely. This is similar to rejecting a relatively low offer in the ultimatum game.

A consequence of the low contributions by the Japanese subjects when they are the only participant is to considerably reduce the incentives for Japanese subjects to forego participation. Tables 4 and 5 summarize the normal form game for the first stage participation decision based on *realized* average monetary earnings during the first 5 periods in the two countries. Compare these payoff tables with the theoretical monetary payoffs based on the stage two Nash equilibrium shown in Table 1. The realized monetary payoff matrix based on the early Japanese data (Table 4) indicates that *participation is a dominant strategy*.

Table 4. First stage participation payoffs based on average payoffs in the second stage through period 5 (Japanese data).

	Player 2	
Player 1	Participate (p_2)	Not Participate (1-p_2)
Participate (p_1)	6570, 6570	2049, 4795
Not participate (1-p_1)	4795, 2049	706, 706

Table 5. First stage participation payoffs based on average payoffs in the second stage through period 5 (American data).

	Player 2	
Player 1	Participate (p_2)	Not participate (1-p_2)
Participate (p_1)	7167, 7167	2400, 7279
Not participate (1-p_1)	7279, 2400	706, 706

The early experience of Japanese subjects shows them that "non-participation doesn't pay," and we believe this is a primary explanation of the high participation rate (rejecting the ESS prediction of 68 percent) observed in the Japanese data. In other words, the behavior off the equilibrium path reinforces deviations from equilibrium participation decisions, similar to the reinforcement of generous (disequilibrium) offers in the ultimatum game and the punishment of low first-player contributions in the sequential summation public goods game of Andreoni et al. (2002), and contrary to the best shot game (e.g., Prasnikar and Roth, 1992). By contrast, for the American data (Table 5), all monetary payoffs are reduced compared to the theoretical predictions in Table 1 (except, of course when both fail to participate); however, the Hawk-Dove property of this payoff matrix is preserved.[17,18]

But why would subjects reduce their own earnings to punish non-participants, especially since they never encounter the same subject twice in this strangers design?[19] This is a problem familiar to experimental economists and it arises repeatedly in bargaining and other laboratory games. The evidence seems clear that subjects in many situations do not seek simply to maximize monetary earnings, and several recent models have introduced non-monetary factors into players' objective functions (Rabin, 1993; Dufwenberg and Kirchsteiger, 1998; Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Levine, 1998; Charness and Rabin, 2000). Our experiment was not designed to differentiate between these alternative models, so we shall not attempt to evaluate their performance using these new data. But we can offer two brief observations regarding the country difference observed in this study.

First, these models can make quite different predictions depending on different preference types and the fraction of subjects consistent with each type. These models may therefore be consistent with the differences between the Japanese and American results if the distribution of preference types differs across countries. Second, our data suggest that reciprocity in this game differs across countries. Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) model players' utilities as depending only on the final payoff allocations and not on players' intentions, whereas Rabin's (1993) fairness equilibrium, Dufwenberg and Kirchsteiger's (1998) sequential reciprocity equilibrium and Charness and Rabin's (2000) model of social preferences allow for reciprocal behavior based on how kind or unkind a player believes his opponent is treating him. Our experiment provides evidence that Japanese subjects act spitefully to punish non-participants, while American subjects seem to forgive (money-maximizing) non-participation choices more readily. It is plausible that the same action of non-participation could be *interpreted* differently in these two cultures, so that different equilibria are being played in each country.

To summarize, we find that Japanese subjects often punish opponents who fail to participate. This punishment occurs even though it is common knowledge that our design allows subjects to interact only once during a session. This (non-myopic best response) behavior by Japanese subjects increases their participation relative to their American counterparts who behave roughly consistent with the standard money-maximizing model. Perhaps most striking is that the increased participation observed for the Japanese subjects eventually increases efficiency compared to the efficiency of the money-maximizing Americans. Evaluated in isolation, spite may not seem to be a desirable cultural or personality trait; but in strategic environments such as this one, it improves efficiency.

5. Conclusion

In this paper we presented a laboratory experiment on two-person public goods provision, in which subjects first announce whether they will participate in a mechanism to fund a public good through voluntary contributions. The subgame perfect Nash equilibrium of this game applies, of course, to any economic decision-maker regardless of their cultural background. The evolutionary-stable-strategy equilibrium of this game involves participation about two-thirds of the time. This prediction, as well as many others we test, is supported only for the data gathered using American subjects. Japanese subjects participated more often than the ESS prediction. This "over-participation" generated efficiency that exceeds that observed for American subjects toward the end of the sessions.

Although the data from the Japanese sessions generally fail to support the theoretical equilibria based on only pecuniary payoffs, they are consistent in some respects with alternative utility payoffs that include nonpecuniary considerations. We observe Japanese subjects who frequently "punish" non-participants with low contributions, and this kind of (negative) reciprocal behavior is how we interpret the term "spite" in our context, as distinct from simple rivalistic (or "competitive") behavior in which a subject seeks merely to earn more than his opponent. The data are inconsistent with such simple rivalistic motivations, because such motivations are more likely to lead to *less* participation than the ESS prediction since non-participation 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 both subjects participate (figure 3) because rivalistic subjects have a strong incentive to reduce their public good investment.

Our results regarding the differences across countries are consistent with the hypothesis that Japanese subjects have a greater propensity to act spitefully towards opponents who fail to participate in funding the public good. This spiteful tendency of Japanese subjects has been identified in different previous public good environments (Saijo and Nakamura, 1995), but in the present setting it leads to more efficient outcomes than realized by American subjects. These individualized punishments are a special feature of two-person public goods environments, although punishments have also been observed recently in larger public goods groups when punishments can be directed to specific individuals—and this behavior also improves efficiency (Fehr and Gachter, 2000). In future research we plan to interact Japanese and American subjects in the *same* sessions in environments such as this and the ultimatum game. This will indicate whether the lack of a shared cultural background initially increases disequilibrium behavior (e.g., spiteful rejections).

Acknowledgments

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, the Tokyo Center for Economic Research, and the Krannert School of Management at Purdue University. We are grateful for the many helpful suggestions provided by participants at the 1997 Amsterdam Workshop on Experimental Economics, the Fall 1997 Economic Science Association conference, the Purdue University theory workshop, and Andreas Ortmann, Arthur Schram, Martin Sefton, Shyam Sunder, Jun Wako, the editor (Holt) and two anonymous referees. The usual caveat applies.

Notes

1. *Random House College Dictionary*, 1979.
2. Participation decisions are often made sequentially in the field, so it is common to learn others' participation decisions before making one's own decision. A specific, recent example of this commitment to free-ride might be the United States' announced intention not to reduce greenhouse gas emissions through a binding international agreement (the Kyoto Protocol). This example is not perfect, since the United States could choose to unilaterally reduce greenhouse gas emissions; this is highly unlikely in the short term, however, since the current administration has announced a goal to increase greenhouse gas emissions.
3. We intepret this spiteful "punishment" response as (negative) reciprocity, but difference aversion is also a plausible explanation, as is the case for the ultimatum game.
4. Buchan et al. actually use the *strategy method*, in which respondents' demands of their share of the surplus determine whether an offer is rejected. Their data indicate that Japanese respondents demand more than American respondents do, so the Japanese subjects are more likely to reject an offer of a given size. In another paper, Buchan et al. (1999) use an experimental trust game to study behavior across four countries (Japan, the United States, China and Korea). They find no differences attributable to subject nationality, but they find significant differences correlated with subject attitudes and beliefs elicited through a questionnaire, which they call culture. Croson and Buchan (1999) use the same data to identify greater reciprocal behavior for women than for men.
5. In a very different context with simple two-stage extensive form games, Beard et al. (2001) find that Japanese subjects are more willing to reject unequal payoff allocations than are American subjects. This leads to more "secure" play in the Jpanese treatments; in particular, the first decision maker is less likely to offer a Pareto superior but more asymmetric payoff distribution in the sessions conducted in Japan.
6. For example, our study provides a more detailed version of the payoff table to subjects than does the Brandts et al. (1999) experiment, and we also employ a strangers matching protocol in all sessions while Brandts et al. use a partners protocol in 12 of their 20 sessions. It is also possible that demographic differences in subject pools influenced our results, but unfortunately we did not collect demographic data to test this conjecture. [Brandts et al. collected demographic data but reject the hypothesis that demographics affected their results.]
7. See Maynard Smith (1982). Note that our use of a "strangers" design—randomly repairing subjects each period—also makes coordination on an asymmetric pure strategy equilibrium extremely difficult.
8. For example, the limited punishment observed in Prasnikar and Roth's (1992) study decreases over time and is lowest in the complete payoff information condition (see their figure 2), which is the information condition most similar to the present experiment.
9. We conducted two sessions in Tokyo to test whether a difference in instruction and record sheet wording—"your opponent" versus "the person you are paired with"—affects behavior. As we document below, results are substantially unaffected by this difference in phrasing. Saijo et al. (1999) presents additional analysis of the Japanese data (only) and provides further evidence that the wording does not systematically influence behavior.
10. Unfortunately, however, subject 11 in the USC session was confused regarding the subject identification numbers and investment choices. She was the only subject (out of 100) who appeared confused. She thought

the identification numbers were the investment choices, so she typically used her opponent's identification number (rather than her opponent's investment choice) when calculating her payoffs. Throughout the results presentation we remove this obviously confused subject from the data prior to analysis, but our qualitative conclusions are generally robust when statistics are recalculated using this confused subject.

11. This 100-to-1 exchange rate was the approximate prevailing exchange rate when the experiments were conducted, but the real value of these payoffs varied somewhat (in terms of local purchasing power) across sites. Nevertheless, the payoff variance across countries in real terms was probably no greater than the variance within countries. For example, cost-of-living differences between West Lafayette, Indiana and Los Angeles were at least as great as the differences between Los Angeles and Tokyo. Consequently, these payoff differences should not confound the main cultural comparison.
12. This random-effects approach we use here and in later analyses is similar to using subjects as the unit of observation—averaging subject's choices across all periods and comparing differences in means across sessions or treatments. It is an improvement over this basic procedure, however, since the random–effects error structure accounts explicitly for subject heterogeneity. Unfortunately, it cannot account for the dependency of choices across subjects within a session (e.g., the effect of one subject's actions on another's later behavior). We could completely eliminate this dependency by discarding all but the first period choices, but we do not pursue that strategy because we are interested in how subjects' behavior changes over time due to their experience.
13. Each subject contributes exactly one observation to these tests. But for the same reason discussed in the previous footnote, the significance levels of this Fisher's exact test are inexact because the common opponents shared by subjects introduces some statistical dependence in the data. We emphasized in the instructions that subjects play each opponent only once and we used a large number of subjects in each session (20). Nevertheless, the play of any two subjects is not completely independent because they both play the same opponents (in different periods) and they play each other in one period. A conservative analysis strategy to avoid this problem would be to consider each session as providing one statistically independent observation. Unfortunately, it would be prohibitively costly to employ the number of subjects necessary to follow this strategy and still obtain meaningful sample sizes.
14. Orbell and Dawes (1993) report an experiment in which subjects first decide whether or not to participate in playing a binary prisoner's dilemma game. They show that intending cooperators are more likely to participate, and as a result of this selection bias they frequently observe the cooperative outcome. Note that our subjects who agree to participate in funding the public good tend to play at or slightly below the noncooperative Nash equilibrium (figure 3), so a similar selection bias is not evident in our data.
15. Asymmetric collusion could generate even higher payoffs, but such coordination is extremely difficult in this strangers design that randomly reassigns pairs each period.
16. We would have preferred to use a random effects model for this test that pools observations across periods, as we do elsewhere in this paper. The efficiency observation is defined for *pairs* of subjects, however, and in our strangers design the pairs were randomly reassigned each period. Therefore, random subject or random pair effect specifications are not possible.
17. In an earlier version of this paper (Cason et al., 1998) we also presented a simple reinforcement learning model to demonstrate that subjects' participation decisions respond to their previous experience. In particular, we show that subjects are more likely to participate if their earnings when participating in previous periods increase relative to their earnings when not participating in previous periods.
18. Following the suggestion of an anonymous referee, we examined how often the low return to non-participation is affected by mutual non-participation (in which both subjects earn only 706), and whether this frequency varies across countries and could therefore contribute to the observed differences in participation rates. The rate that subjects *both* choose to not participate in the American data is about double the rate in the Japanese data, whether one looks at the first five periods (11.6 percent of pairs in the U.S. compared to 6.7 percent of pairs in Japan) or over all periods (9.5 percent compared to 4.9 percent). In the first five periods the average payoff for participants is lower than for non-participants in both countries, but the average *reduction* in earnings from non-participation in Japan (-1760) is greater than in the U.S. (-994). Since the American subjects experience the low payoff from mutual non-participation more frequently than do the Japanese subjects, the greater average reduction in payoffs from non-participation in Japan is mostly due to the punishments received from the other participating subject.

19. We even observe non-payoff-maximizing contributions in the final period 15, even though the final period was announced in the instructions (see figure 2).

References

Andreoni, J., Brown, P., and Vesterlund, L. (2002). "What Makes an Allocation Fair? Some Experimental Evidence." *Games and Economic Behavior*. 40, 1–24.

Beard, T.R., Beil, R., and Mataga, Y. (2001). "Reliant Behavior in the United States and Japan." *Economic Inquiry*. 39, 270–279.

Bolton, G. and Ockenfels, A. (2000). "ERC: A Theory of Equity, Reciprocity and Competition." *American Economic Review*. 90, 166–193.

Brandts, J., Saijo, T., and Schram, A. (1999). "A Four Country Comparison of Spite, Cooperation and Errors in Voluntary Contribution Mechanisms." Working paper, University of Amsterdam (CREED).

Brown, J. and Rosenthal, R. (1990). "Testing the Minimax Hypothesis: A Re-examination of O'Neill's Game Experiment." *Econometrica*. 58, 1065–1081.

Buchan, N., Croson, R., and Johnson, E. (1999). "Getting to Know You: An International Experiment of the Influence of Culture, Communication, and Social Distance on Trust and Reciprocation." Working paper 98-03-05, Wharton School.

Buchan, N., Johnson, E., and Croson, R. (1999). "Understanding What's Fair: Contrasting Perceptions of Fairness in Ultimatum Bargaining in Japan and the United States." Working paper 97-03-10, Wharton School.

Cason, T., Saijo, T., and Yamato, T. (1998). "Voluntary Participation and Spite in Public Good Provision Experiments: An International Comparison," Working paper, Purdue University Center for International Business Education and Research (CIBER).

Charness, G. and Rabin, M. (2000). "Social Preferences: Some Simple Tests and a New Model." Working paper, University of California at Berkeley.

Croson, R. and Buchan, N. (1999). "Gender and Culture: International Experimental Evidence from Trust Games." *American Economic Review Papers and Proceedings*. 89(2), 386–391.

Duffy, J. and Feltovich, N. (1999). "Does Observation of Others Affect Learning in Strategic Environments?" *International Journal of Game Theory*. 28, 131–152.

Dufwenberg, M. and Kirchsteiger, G. (1998). "A Theory of Sequential Reciprocity." Working paper, Uppsala University.

Fehr, E. and Gachter, S. (2000). "Cooperation and Punishment in Public Goods Experiments." *American Economic Review*. 90, 980–994.

Fehr, E. and Schmidt, K. (1999). "A Theory of Fairness, Competition and Cooperation." *Quarterly Journal of Economics*. 114, 817–868.

Harrison, G. and Hirshleifer, J. (1989). "An Experimental Evaluation of Weakest Link/Best Shot Models of Public Goods." *Journal of Political Economy*. 97, 201–225.

Isaac, R.M. and Walker, J. (1988). "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism." *Quarterly Journal of Economics*. 103, 179–200.

Kachelmeier, S. and Shehata, M. (1992). "Culture and Competition: A Laboratory Market Comparison Between China and the West." *Journal of Economic Behavior and Organization*. 19, 145–168.

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 University Press, Cambridge.

Okada, A. and Riedl, A. (1999). "When Culture Does Not Matter: Experimental Evidence from Coalition Formation Ultimatum Games in Austria and Japan," Working paper, University of Amsterdam (CREED).

Orbell, J. and Dawes, R. (1993). "Social Welfare, Cooperators' Advantage, and the Option of Not Playing the Game." *American Sociological Review*. 58, 787–800.

Prasnikar, V. and Roth, A. (1992). "Considerations of Fairness and Strategy: Experimental Data from Sequential Games." *Quarterly Journal of Economics*. 108, 865–888.

Rabin, M. (1993). "Incorporating Fairness into Game Theory and Economics." *American Economic Review*. 83, 1281–1302.

Roth, A., Prasnikar, V., Okuno-Fujiwara, M., and Zamir, S. (1991). "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study." *American Economic Review*. 81, 1068–1095.

Saijo, T. and Nakamura, H. (1995). "The "Spite" Dilemma in Voluntary Contribution Mechanism Experiments." *Journal of Conflict Resolution*. 39(3), 535–560.

Saijo, T. and Yamato, T. (1999). "A Voluntary Participation Game with a Non-Excludable Public Good." *Journal of Economic Theory*. 84, 227–242.

Saijo, T., Yamato, T., Yokotani, K., and Cason, T.N. (1999). "Voluntary Participation Game Experiments with a Non-Excludable Public Good: Is Spitefulness a Source of Cooperation?" Working paper, Osaka University.