THE MINIMUM APPROVAL MECHANISM IMPLEMENTS
THE EFFICIENT PUBLIC GOOD ALLOCATION
THEORETICALLY AND EXPERIMENTALLY

Takehito Masuda
Yoshitaka Okano
Tatsuyoshi Saijo

The Minimum Approval Mechanism Implements the Efficient Public Good Allocation Theoretically and Experimentally

September 2013

Takehito Masuda [a,b,*], Yoshitaka Okano [c], and Tatsuyoshi Saijo [c, d]

[a] *Graduate School of Economics, Osaka University 1-7, Machikaneyama, Toyonaka, Osaka 560-0043, Japan*

[b] *Research Fellow of the Japan Society for the Promotion of Science, Kojimachi Business Center Building, 5-3-1 Kojimachi, Chiyoda, Tokyo 102-0083, Japan*

[c] *Research Center for Social Design Engineering, Kochi University of Technology, Tosayamada, Kami, Kochi 782-8502, Japan*

[d] *Center for Environmental Innovation Design for Sustainability, Osaka University, 2-1 Yamada-oka, Suita, Osaka 565-0871, Japan*

Abstract

We propose the minimum approval mechanism (MAM) for a standard linear public good environment with two players. Players simultaneously and privately choose their contributions to the public good in the first stage. In the second stage, they simultaneously decide whether to approve the other's choice. Both contribute what they choose in the first stage if both players approve; otherwise, both contribute the minimum of the two choices in the first stage. The MAM multiply implements the Pareto-efficient allocation in backward elimination of weakly dominated strategies (BEWDS), limit logit agent quantal response equilibrium, subgame perfect minimax regret equilibrium, level-k thinking, and diagonalization heuristics. Moreover, the MAM is unique under plausible conditions. Overall, contributions in the MAM experiment averaged 94.9%. Quantifying subjects' responses to the questionnaire revealed the heterogeneity of reasoning processes to be consistent with the model.

---

* Corresponding author. Postal address: Research Center for Social Design Engineering, Kochi University of Technology, Tosayamada, Kami, Kochi 782-8502, Japan. Tel.: +81-887-57-2795; Fax: +81-887-57-2366. E-mail addresses: tmasuda@iser.osaka-u.ac.jp (T. Masuda), okano.yoshitaka@kochi-tech.ac.jp (Y. Okano), and tatsuyoshisaijo@gmail.com (T. Saijo).
Abbreviations: MAM = minimum approval mechanism, BEWDS = backward elimination of weakly dominated strategies, MCM = mate choice mechanism.

1. Introduction

The free-rider problem lies at the heart of public good provision. To solve this problem, theorists have traditionally concentrated on designing mechanisms yielding efficient outcomes, given that players are self-interested. In this mechanism design or implementation approach, individual theorists have discretion over which equilibrium concept or behavioral rule to use, such as dominant strategy equilibrium, Nash equilibrium, or subgame perfect Nash equilibrium (SPNE). However, recent experimental results have raised questions concerning this fundamental assumption.[1] It is also widely observed that experimental subjects are heterogeneous in their motive to choose their contributions to the public good.

We offer an alternative approach, multiple implementations, in order to bridge gaps between the theory and experiments. That is, public good mechanisms should be robust to as many as possible equilibrium concepts or behavioral rules. More precisely, we design the minimum approval mechanism (MAM) that multiply implements the symmetric Pareto-efficient allocation for linear public good environments with two players.

The MAM consists of two stages. Each player chooses his/her contribution in the first stage, following which each player observes the other's choice and approves or disapproves the same. If both players approve, their final contributions are what they chose in the first stage. If either disapproves, both finally contribute the minimum of their first-stage choices.

Under the MAM, both players fully contribute in all five equilibrium concepts or heuristics: (i) backward elimination of weakly dominated strategies (BEWDS; Saijo et al., 2013)[2], (ii) limit logit agent quantal response equilibrium (limit LAQRE; Mckelvey and Palfrey, 1995; 1998), (iii) subgame perfect minimax regret equilibrium (SPMRE; Renou and Schlag, 2011), (iv) level-k thinking (Costa-Gomes and Crawford, 2006), and (v) diagonalization heuristics.

Moreover, we show that the MAM is a unique BEWDS-implementing approval mechanism, satisfying the following natural properties. *Forthrightness* requires that the outcome must be what players choose whenever both approve the other's first-stage choices. *Voluntariness* mandates that no player is forced to finally contribute more than what he/she chooses in the first stage.[3] *Monotonicity* states that final contributions increase as their first-stage choices increase.

In experiments under the perfect stranger matching protocol, we observe, as predicted, a rapid convergence to efficient allocation in the MAM. Subjects successfully

---

[1] See Chen (2008) for an excellent survey.

[2] As Saijo et al. (2013) noted, Kalai (1981) used BEWDS.

[3] In other words, we only provide refunds to players, rather than using coercive power.

sustained high average contribution rates, beginning with 76.9% in the first period and ending with 89.7% in the final period. The overall average contribution rate is 94.9%.[4] Our experimental results of the MAM are in sharp contrast with those of existing mechanisms that need dozens of iterations to converge to the efficient Nash equilibrium in experiments. We also conduct experiments for the mate choice mechanism (MCM) of Saijo et al. (2013), which theoretically does not work in our public good environment and observe that the average contribution fluctuates within a middle range.

In order to capture subjects' reasoning processes, we apply a coding scheme similar to the one of Cooper and Kagel (2005). Two coders independently quantified subjects' responses to open questions during and after the session. The coding confirms that subjects are heterogeneous in their motive. In particular, 20.0% of subjects in the MAM are deemed to have used diagonalization heuristics. Moreover, 16.7% and 11.7% of subjects in the MAM and MCM, respectively are deemed to minimize regret. Finally, 21.7% and 45.0% of subjects in the MCM are deemed to have followed level-1 and level-2 thinking, respectively.

Our approach differs from much of the experimental research on the public good mechanism, which thus far has mainly focused on evaluating mechanism designs based on dominant strategy equilibrium, Nash equilibrium, or SPNE. For dominant strategy mechanisms, Attiyeh et al. (2000) found that subjects in the pivotal mechanism frequently misrepresent their private information and become stuck at a weakly dominated Nash equilibrium. To solve this problem, Saijo et al. (2007) introduced secure implementation (or double implementation in dominant strategy equilibrium and Nash equilibrium). Further, Cason et al. (2006) experimentally showed that the frequency of dominant strategy play was significantly increased in a secure Groves–Clarke mechanism relative to a nonsecure pivotal mechanism, although the former's frequency was only slightly over 80%.

The Groves–Ledyard (Groves and Ledyard, 1977) and Walker (1981) mechanisms have difficulty in attaining efficient Nash equilibria in one-shot play, because of subjects' bounded rationalities.[5] Andreoni and Varian (1999) reported similar results for Varian's (1994) extensive form mechanism with the unique SPNE. Chen and Plott (1996), Chen and Tang (1998), Chen and Gazzale (2004), Healy (2006), and Healy and Mathevet (2012) characterized the stable mechanisms to achieve efficient Nash equilibria as rest points of various learning dynamics. Nonetheless, these studies still required dozens of repetitions, which may not be feasible in practical situations. A notable exception is Falkinger et al. (2000), who observed that under Falkinger's (1996) mechanism, contributions approached

---

[4] We obtained a similar result in the simplified MAM (SMAM) treatment, in which the player with the higher first-stage choice alone can proceed to the second stage.

[5] See Chen and Plott (1996) and Chen and Tang (1998).

the efficient Nash equilibrium in only a few periods.[6]

The remainder of this paper is organized as follows. Section 2 overviews the MCM proposed by Saijo et al. (2013), introduces the MAM with examples, and presents the implementation results. Section 3 describes the proof of the uniqueness of the MAM. Section 4 argues in favor of the MAM's robustness. Section 5 describes the experimental design. Section 6 discusses the experimental results. Section 7 provides an analysis of subjects' responses to the open-ended questionnaire. Section 8 discusses further implications and concludes.

2. The model and the MAM

Consider a voluntary contribution mechanism (VCM) in the provision of a public good with two players. Each player i has an initial endowment $w > 0$, and he/she must decide the contribution $s_i \in S = [0, w]$. The sum of the contribution is multiplied by $\alpha \in (0.5, 1)$, and the benefit passes to every player, which expresses the non-rivalness of the public good. That is, player i's payoff is $u_i(s_i, s_j) = (w - s_i) + \alpha(s_i + s_j), j \neq i$. Then, $s_1 = s_2 = 0$ is the dominant strategy equilibrium, and hence, no public good is provided.

When there are only two levels of the contribution (i.e., 0 or w) when players face the prisoner's dilemma, Saijo et al. (2013) introduced the notion of the MCM following the dilemma decision. By knowing the other player's choice between cooperation (contributing w and abbreviated by C) and defection (contributing 0 and abbreviated by D) in the dilemma or the first stage, each player must approve (y) or disapprove (n) it. If both players approve the choices of their counterparts in the first stage, the outcome or payoff vector is what they choose. If either player disapproves, however, the outcome is that when both choose D.

<center><Figure 1 about here.></center>

Let us explain how BEWDS works. Figure 1 shows an example where $w = 10$ and $\alpha = 0.7$. Note that this is an unusual extensive form game tree. Consider the case where player 1 chooses D and player 2 chooses C (termed subgame (D,C) in Figure 1). If both choose y in the second stage, players 1 and 2 obtain 17 and 7, respectively. Otherwise, both obtain 10. Then, player 1 compares (17,10) with (10,10) and chooses y, since y weakly dominates n. A vector x weakly dominates z if every component in x is at least greater than or equal to that in z, and there is at least one component with strict inequality. Similarly, player 2 chooses n since (10,10) weakly dominates (7,10). Therefore, the outcome in subgame (D,C) is (10,10). By repeating the same procedure for every subgame, players can

---

[6] Contributions averaged 90.5% of subjects' endowments in experiments with linear payoffs. However, this mechanism proportionally taxes or subsidizes to the extent of the deviation of a player's contribution from the mean contribution, and hence, some players are forced to contribute more than they choose.

construct the normal form game. Then, by comparing (14,10) with (10,10), the player chooses *C* since the former weakly dominates the latter. Thus, BEWDS achieves (*C,C*) in the prisoner's dilemma game with the MCM.

Saijo et al. (2013) showed that subjects achieve nearly full cooperation in the MCM experiment, where each subject was randomly paired with another, and the same two subjects were not paired more than once. With the help of a related experiment, the authors also showed that BEWDS is most consistent with the data among the behavioral principles based upon Nash equilibrium, SPNE, neutrally stable strategies, and BEWDS.[7] Following Saijo et al. (2013), we use BEWDS as a basic behavioral principle and restrict ourselves to the class of mechanisms where each player chooses $s_i \in S$ in the first stage and *y* or *n* in the second stage. We call this mechanism an *approval mechanism*. In other words, the MCM is an approval mechanism.

However, the simple extension of the MCM with two strategies 0 and *w* to a case with many strategies cannot implement the symmetric Pareto-efficient outcome. For simplicity, hereafter, we call the extended MCM with many strategies as the MCM. In the MCM, if both players approve the other player's choice in the first stage, then both contribute what they choose, whereas if either one disapproves, then both contribute 0. To see why the MCM breaks down, consider the case with *w* = 10 and $\alpha = 0.7$ once more. For simplicity, each player can choose a contribution from {0,2,4,6,8,10}. Table 1 shows player 1's (the row player) payoff table. The symmetric Pareto-efficient outcome is (14,14) when both players contribute 10. Consider subgame (10,8). Then, both choose *y*, since for player 1 (12.6,10) weakly dominates (10,10), and for player 2 (14.6,10) weakly dominates (10,10). Thinking in reverse, player 2 will find choosing 8 a profitable deviation, which shows that the MCM cannot achieve efficiency.

<Table 1 about here.>

In order to implement the symmetric Pareto-efficient outcome, we introduce the MAM. In this mechanism, if both players approve the contribution of the other in the first stage, then both contribute what they choose, whereas if either one disapproves, then both contribute the minimum of the two contributions in the first stage.

Using Table 1, let us consider two examples of subgames under the MAM. In subgame (6,2) on the left-hand side of Table 2, if either disapproves the other's choice, the mechanism returns 4 (= 6 − 2) to player 1, and both players contribute 2 and obtain 10.8. Therefore, player 1 chooses *n* since (10.8,10.8) weakly dominates (9.6,10.8), and player 2 chooses *y* since (13.6,10.8) weakly dominates (10.8,10.8). Therefore, both obtain 10.8.

---

[7] Neutrally stable strategies also attain (*C,C*), though the outcomes of the Nash equilibrium and the SPNE show every possibility, namely (*C,C*), (*C,D*), (*D,C*), and (*D,D*). As shown in Saijo et al. (2013), the strategies of BEWDS are a subset of neutrally stable strategies even though both achieve (*C,C*).

<Table 2 about here.>

Next, consider subgame (8,10) on the right-hand side of Table 2. Player 1 chooses *y* and player 2 chooses *n*; hence, the payoff is (13.2,13.2). In general, player *i* chooses *y* (*n*) if his/her first-stage choice is smaller (larger) than that of player $j \neq i$: $s_i <(>)s_j$. Table 3 shows player 1's payoff matrix in the reduced normal form game.

Clearly, choice 10 weakly dominates the rest, and hence, the MAM implements the symmetric Pareto-efficient outcome in BEWDS. Table 3 presents a special structure: the payoff in subgame (*m,m*) is the same as that in (*m,n*) and (*n,m*) for all *n* > *m*, whereas the payoff in (*n,n*) is greater than that in (*m,m*) for all *n* > *m*. If the payoff matrix satisfies these two conditions, we say that it has an *echelon structure*.

<Table 3 about here.>

Proposition 1. *The MAM implements the symmetric Pareto-efficient outcome in BEWDS.*

Proof. First, we show that player 1's payoff matrix in the reduced normal form game under the MAM has an echelon structure. Fix $a \in S$, and choose any $b \in S$ with *b* > *a*. Consider subgame (*a,b*). Since $\{w - a + \alpha(a + a)\} - \{w - b + \alpha(a + b)\} = (b - a)(1 - \alpha) > 0$, player 2 chooses *n* in the second stage. Then, player 1's payoff becomes $w - a + \alpha(a + a)$. In other words, for all *b* > *a*, the payoff in subgame (*a,b*) is the same as that in (*a,a*) under the MAM. Similarly, we find that for all *b* > *a*, the payoff in subgame (*b,a*) is the same as that in (*a,a*) under the MAM. Choose any *a*, $b \in S$ with *b* > *a*. Then, the payoff in subgame (*b,b*) is greater than that in (*a,a*) since $w - b + 2\alpha b - (w - a + 2\alpha a) = (2\alpha - 1)(b - a) > 0$.

We further show that choice *w* weakly dominates $s_1 \in [0, w)$. Consider the case with $s_2 \in [0, s_1]$. From the first part of the echelon structure, we have $u_1(s_1, s_2) = u_1(s_2, s_2)$ $= u_1(w, s_2)$. [8] Consider the case with $s_2 \in (s_1, w]$. Using the two properties of the echelon structure, we have $u_1(s_1, s_2) = u_1(s_1, s_1) < u_1(s_2, s_2) = u_1(w, s_2)$. In other words, choice *w* weakly dominates $s_1 \in [0, w)$. ∎

3. The MAM is unique

There could be various ways to implement the symmetric Pareto outcome using approval mechanisms under BEWDS. For example, consider the average approval mechanism (AAM): if either player disapproves, then both contribute the average of the two choices in

---

[8] Let $t = (t_1, t_2)$ be the decision in the second stage. Then, we must write $u_i((s_1, t_1), (s_2, t_2))$. Thus, $u_i(s_1, s_2)$ should be $u_i(s_1, s_2) = u_i((s_1, y), (s_2, n))$. In other words, we must specify the decision in the second stage in order to obtain the payoff of player *i*. By slight abuse of notation, we can also use it both ways whenever the decision in the second stage is clear.

the first stage. The symmetric Pareto outcome is attained by the AAM under BEWDS.[9]

Under the AAM, however, the player who chooses a smaller number must contribute more. In order to exclude this coercion, we state that an approval mechanism is *voluntary* if each player's contribution does not exceed his/her first-stage choice. Clearly, the MAM satisfies voluntariness and the AAM does not.[10]

Moreover, we show that the MAM is a unique mechanism for satisfying voluntariness and several other conditions. For this purpose, let us introduce the following definitions. We refer to the players' choices in each stage $((s_1,t_1),(s_2,t_2))$ as a *path*. An approval mechanism g is a function that associates a contribution $g((s_1,t_1),(s_2,t_2)) = (g_1((s_1,t_1),(s_2,t_2)),g_2((s_1,t_1),(s_2,t_2)))$ for every path $((s_1,t_1),(s_2,t_2))$. First, we require that approval mechanisms be *forthright* in the following manner: if both make the same first-stage choice or if both choose y in the second stage, then they contribute what they choose in the first stage; moreover, given the first-stage choices, the outcomes when either chooses n are the same.[11] Formally, we have:

Definition 1. An approval mechanism g is *forthright* if:
(i) $g((s_1,\cdot),(s_2,\cdot)) = (s_1,s_2)$ if $s_1 = s_2$, where "$\cdot$" indicates either y or n;
(ii) $g((s_1,y),(s_2,y)) = (s_1,s_2)$ for every (s_1,s_2); and
(iii) $g((s_1,n),(s_2,y)) = g((s_1,y),(s_2,n)) = g((s_1,n),(s_2,n))$ for every (s_1,s_2).

In the following, vector inequality is taken component-wise as usual. The next property is voluntariness, as argued above.

Definition 2. An approval mechanism g is *voluntary* if $g((s_1,t_1),(s_2,t_2)) \leq (s_1,s_2)$ for every path $((s_1,t_1),(s_2,t_2))$.

The third property, *monotonicity*, is a plausible condition that reflects players' willingness to pay more. It states that if players weakly increase their first-stage choices, then their contributions weakly increase regardless of their decisions in the second stage.

Definition 3. An approval mechanism g is *monotonic* if $(s_1',s_2') \geq (s_1,s_2)$ implies

---

[9] It is easy to check that the AAM implements a Pareto-efficient allocation in BEWDS. More generally, for any $\lambda \in [0,1]$, $\lambda-$ weighted AAM, where if either says no then both players finally contribute $\lambda s_1 + (1-\lambda)s_2$, also implements a Pareto-efficient allocation in BEWDS.

[10] The budget-balanced mechanism proposed by Falkinger et al. (2000), under which subjects contribute over 90% of endowments, is similar to the AAM, because it taxes (subsidizes) players with a contribution lower (higher) than the group mean.

[11] This definition is slightly different from forthrightness introduced by Saijo et al. (1996).

$$g((s_1', t_1), (s_2', t_2)) \geq g((s_1, t_1), (s_2, t_2)) \text{ for every } (t_1, t_2).$$

The above properties are sufficient to prove the uniqueness of the MAM.[12]

Proposition 2. *Suppose that a forthright approval mechanism implements the Pareto-efficient allocation in BEWDS. If it is voluntary and monotonic, then it must be the MAM.*

Proof. See Appendix. ∎

4. Multiple implementations of the MAM

In this section, we explore other equilibrium properties and the outcome of heuristics under the MAM. First, we present the main result of implementation in several equilibrium concepts and heuristics.[13] Then we derive the result step by step.

Proposition 3. *The MAM multiply implements the symmetric Pareto-efficient outcome in limit LAQRE, SPMRE, level-*k *thinking, and diagonalization heuristics.*

Proof. It immediately follows from Lemmata 2-5. ∎

The concept of SPNE is often used to analyze multistage games. We denote player i's strategy by $p_i = (s_i, t_i(\cdot))$, where $s_i \in S$, and $t_i(\cdot)$ is a function that associates y or n to each first-stage choice (s_1, s_2). Let $P_1 = P_2 = P$ be the pure strategy set. With a slight abuse of notation, we write i's payoff function induced by the MAM as u_i. A strategy profile $p = (p_1, p_2)$ is a Nash equilibrium if for each i, j with $j \neq i$, $u_i(p) \geq u_i(p_i', p_j)$ for all $p_i' \in P$. A Nash equilibrium p is *subgame perfect* if it constitutes a Nash equilibrium in every subgame.

Lemma 1. *In the MAM, the set of SPNE contributions is the set of symmetric contributions.*

Proof. See Appendix. ∎

---

[12] Although it is artificial, the following forthright mechanism is voluntary, but not monotonic. Consider a forthright rule in which a player who announces a bigger contribution than the other in the first stage must contribute 0, and the other player must contribute what he/she chooses in the first stage if either one chooses n in the second stage. Apparently, this mechanism is voluntary. However, it is not monotonic, since $g_1((6,n),(2,y)) = 0 < g_1((2,n),(2,y)) = 2$. It is easy to check that the mechanism implements a Pareto-efficient allocation in BEWDS.

[13] We thank an anonymous referee for pointing out this possibility.

Mckelvey and Palfrey (1995, 1998) introduced logit quantal response equilibrium (QRE) concept for the normal and extensive form games. They modeled the players who make noisy best response to other players. The probability of choosing an action is proportional to the exponential function of the observed payoff of the action, divided by a parameter $\mu > 0$. Even in such a case, the MAM can implement the efficient outcome in *limiting* equilibrium, which survives as players become fully sensitive to their payoffs: $\mu \to 0$. Assume μ is common to both players and all their information sets. To ensure the existence of the limiting equilibrium by Theorem 1 of Mckelvey and Palfrey (1998), we focus on the case where players contribute in integer.

Let $\sigma_i = (\sigma_i^1, \sigma_i^2)$ be player i's behavioral strategy with the superscript denoting stage, and let $\sigma = (\sigma_1, \sigma_2)$ be a strategy profile. Let $\pi_i(s_i)$ be i's expected payoff from when he/she chooses $s_i \in S_d = \{1, 2, \ldots, w\}$, while all the other actions of the players are prescribed by σ. Similarly, let $\pi_i(t_i) = \pi_i(t_i | s_i, s_j)$, $t_i = y, n$ be i's expected payoff when he/she chooses t_i at subgame (s_i, s_j), while $j \neq i$ is following σ in the subgame, where the vertical " $|$ " indicates "conditional on." A strategy profile σ is an *LAQRE* if it solves the following equations: for each $i = 1, 2$, $j \neq i$,

$$\sigma_i^1(s_i) = \frac{\exp(\pi_i(s_i)/\mu)}{\sum_{x \in S_d} \exp(\pi_i(x)/\mu)} \quad \text{for each } s_i \in S_d, \text{ and } \sigma_i^2(y | s_1, s_2) = \frac{\exp(\pi_i(y)/\mu)}{\exp(\pi_i(y)/\mu) + \exp(\pi_i(n)/\mu)}$$

for each $s_i, s_j \in S_d$. A strategy profile σ is a limit LAQRE, if it is an LAQRE as $\mu \to 0$.

In the limit, since the player with a higher first-stage choice surely chooses *n*, the other player randomizes *y* and *n* by indifference. Still, the reduced normal form game approaches the echelon form.[14] Then, analogous to Proposition 5 of Anderson et al. (2001) for minimum effort games, the full contribution is the unique equilibrium. Therefore, we have Lemma 2.

Lemma 2. *The MAM implements the symmetric Pareto-efficient outcome in limit LAQRE.*

Proof. See Appendix. ∎

The third alternative equilibrium concept is *SPMRE*. Some subjects imagine foregone payoffs from unchosen strategies. This is close to the simplest version of ε-minimax regret equilibrium introduced by Renou and Schlag (2011), who incorporated into game-theoretic models experimental evidence that people avoid ambiguity and the decision-theoretic formulation of such behavior. The basic idea is that a player imagines

---

[14] For a more general discussion of the QRE, see Mckelvey and Palfrey (1995, 1998).

foregone payoffs from unchosen strategies. For every pure strategy profile $p = (p_1, p_2)$, player i's *regret* at p is defined by $R_i(p) = \max_{p_i'} u_i(p_i', p_j) - u_i(p_i, p_j)$, $j \neq i$. This is the difference between his/her maximal payoff when player i best responds to j's strategy p_j and his/her payoff when he/she plays p_i. A strategy profile p is a *minimax regret equilibrium* if for each i, j with $j \neq i$ and for all p_i', $\max_{p_j'} R_i(p_i, p_j') \leq \max_{p_j'} R_i(p_i', p_j')$. In other words, each player chooses the strategy that minimizes the largest possible regret in the equilibrium.[15] We say that a minimax regret equilibrium p is *subgame perfect* if the action profile induced by p constitutes a minimax regret equilibrium in every subgame.

Lemma 3. *The MAM implements the symmetric Pareto-efficient outcome in SPMRE.*

Proof. See Appendix. ∎

We now propose two alternative heuristics that predict that players make the full contribution. The first one is level-*k* thinking. We apply such heuristics to the reduced normal form game induced by elimination of weakly dominated strategies. Consider Table 3 as an illustration. Assume that player 1 expects player 2 to choose each number with equal probability (level-0) and to eliminate weakly dominated strategies in the second stage. In the case of the MAM, player 1's best response to level-0 is to choose 10. Thus, efficiency can be achieved by any pair of players of level-1 or above.[16]

Lemma 4. *The MAM implements the symmetric Pareto-efficient outcome in level-1 or above.*

Proof. Straightforward. ∎

The fifth possibility is a heuristic called *diagonalization,* of which we find notable descriptions given by the subjects in the questionnaires during and after the experiment. Some subjects wrote that the player who chooses the higher first-stage choice will disapprove the other's choice, and both would obtain the payoffs on the diagonal line in the payoff table. Then, subjects can easily find that full contributions attain the maximum payoff among the payoffs on the diagonal line. Thus, we have the following simple proposition.

---

[15] Here, we assume that each player believes that his/her opponent may choose any strategy.
[16] Lemma 4 also holds if we consider the level-0 player who chooses w.

Lemma 5. *If both players follow diagonalization in the MAM, then the outcome is symmetric Pareto-efficient.*

Proof. Straightforward. ∎

Now, Proposition 3 is verified for the MAM. Proposition 4 for the MCM provides the counterparts to Propositions 1 and 3 for the MAM. Moreover, the results are quite contrasting. Consider Table 1 again. Let us begin with BEWDS. Since both obtain 10 if either chooses n, player 1 will know that player 2 disapproves if and only if $(s_1, s_2) = (0, x)$ with $x \geq 2$, $(2, x)$ with $x \geq 6$, or $(4, 10)$. Then, the choices of 0, 8, and 10 are eliminated, but any choice of 2, 4, and 6 survives under BEWDS. Every pair of contributions such that both players get at least 10 can be supported as an SPNE.[17] Consider level-k thinking. Given that player 2 randomly chooses any number, player 1 can see that the best response is to choose 6. This also minimizes regret. Another step of iterated reasoning leads player 1 to contribute 4, which is the best response to player 2's 6. The third iteration result uniquely selects the choice of 2.[18] (2,2) is also a limit LAQRE contributions. Let $\beta = (1 - \alpha) / \alpha$.

Proposition 4. *Under the MCM, the following holds.*
i) A player following BEWDS chooses any contribution in $(0, \beta w]$.
ii) There exists an SPNE where players contribute (g_1, g_2) if and only if $\beta g_2 \leq g_1 \leq \beta^{-1} g_2$.
iii) A limit LAQRE supports contributions $(1,1)$ if players choose their contributions in integers.
iv) The unique SPMRE contribution is βw.
v) Level-k player's contribution is $\beta^k w$, which decreases as k increases.

Proof. See Appendix. ∎

5. Experimental design

We conducted four treatments in order to test the performances of the approval mechanisms. The VCM was the control treatment, while the other three treatments were the MCM, MAM, and simplified MAM (SMAM). Under the SMAM, if both players make the same first-stage choice, the game ends. Otherwise, only the player with the higher first-stage choice proceeds to the second stage. If he/she approves, they contribute what they choose in the first stage. If he/she disapproves, both players contribute the minimum of their choices in the first stage. The SMAM also implements the efficient allocation in

---

[17] That is, (x,x) with $x \in \{0, 2, 4, 6, 8, 10\}$, (4,2), (6,4), (8,4), (8,6), (10,6), (10,8), and their permutations.
[18] Proposition 4-iv) also holds if we consider the level-0 player who chooses w.

BEWDS.[19] Each of the above-mentioned four treatments had three experimental sessions. These 12 sessions were conducted at Osaka University in March, April, September, and November 2011. Subjects were recruited from Osaka University through campus-wide advertisements and were inexperienced in this particular type of experiment. No individual participated in more than one session. The experiment was computerized using the experimental software z-Tree (Fischbacher, 2007). Twenty subjects participated in each session. Each was seated at a computer terminal assigned by lottery. All terminals were separated by partitions. No communication was allowed between subjects.

Each subject had a set of printed instructions, a record sheet, and a payoff table (called the points table).[20] In the distributed payoff table, each subject was considered to be player 1. Player i's payoff is given by $u_i = 300\{(24 - g_i) + 0.7(g_1 + g_2)\}$ for each pair of contribution (g_1, g_2). Since we consider two players with identical utility functions and endowments, the payoff table is identical for all subjects. Subjects were informed that they had identical payoff tables.

Instructions were read aloud by an experimenter for approximately 10 minutes. After that, subjects were given another 10 minutes to ask questions. Then, subjects proceeded to the payment periods. There was no practice period. Each session consisted of 19 periods under the perfect stranger matching protocol. Subjects were informed that each of them would meet any other subject just once.

The approval mechanism treatments continued as follows. At the beginning of each period, all subjects endowed with 24 tokens were anonymously paired. In the first stage, subjects were asked to enter their contributions as nonnegative integers into a box in the display and to write down their choices along with their reasoning in the corresponding row in the record sheet. Once all subjects had finished their tasks, they clicked the *OK* button. In the second stage, the first-stage choices of both players and the payoff matrix in the second stage were displayed. After all players wrote down the first-stage choices of their counterparts, they chose to approve or disapprove by clicking the relevant radio button and recording the decision along with the reason. If a subject in the SMAM did not proceed to the second stage, he/she simply wrote down the first-stage choice of his/her counterpart before the other subjects finished the second stage and circled "none" while facing the waiting screen. Once all the subjects in the second stage had finished the above tasks and clicked the *OK* button, they proceeded to the results screen. The results screen included the first-stage choices of both players, their "approve" or

---

[19] Recall that under the MAM, the decision by the player with the higher first-stage choice alone determines the outcome in every second stage. The SPNE, SPMRE, limit LAQRE, and level-k thinking outcomes of the SMAM are the same as those of the MAM.

[20] See Online Appendix.

"disapprove" decisions, and their payoffs (points earned) in the period. No information on the choices of the other nine groups was provided to the subjects. Finally, the subjects wrote down the decisions of their counterparts and the points they had earned and clicked the *Next* button to begin the next period. The treatment without the second stage became the VCM treatment.

After playing for 19 periods, the subjects completed a questionnaire and were immediately paid privately in cash. Each subject was paid an amount proportional to the sum of the points that he/she had earned for the 19 periods. Individual payments ranged from $49.91 to $87.97.

6. Experimental results

Figure 2 shows the time path of average contributions over the 19 periods arranged by mechanism. Individual contributions were evaluated after the second stage in the MAM, MCM, and SMAM.

<Figure 2 about here.>

We first assess the results of the VCM. The time path of average contribution in the VCM shows a similar pattern to that of previous VCM experiments (e.g., Ledyard, 1995). In the first period, subjects contributed 15.3% of endowments (3.68 tokens), with contribution rates gradually decreasing to 5.7% (1.37 tokens) in the last period. The hypothesis test based on Spearman's rank correlation coefficient shows that the downward trend in average contributions was statistically significant $(\rho = -0.916, p < 0.001)$. Data pooled across all 19 periods and 3 sessions show that the subjects contributed 10.2% of endowments (2.44 tokens). Out of 570 outcomes (10 pairs × 19 periods × 3 sessions), there is no case where both players contributed all endowments.

Introducing the MAM facilitates almost full contributions among subjects, even in the earlier period. Contributions in the first period average 76.9% of endowments (18.45 tokens). They then rise repeatedly for the first five periods, achieving 97.8% (23.47 tokens) in period 5. Then, cooperation is sustained with an average contribution rate of over 95% in every period except the final one. Moreover, the convergence to the efficient allocation under the MAM is statistically supported by Spearman's rank correlation test, which helps examine the upward trend in average contributions $(\rho = 0.557, p = 0.009)$. When the data are pooled across all 19 periods and 3 sessions, subjects in the MAM contribute on average 94.9% of endowments (22.78 tokens). Out of 570 outcomes, there are 475 outcomes where both players contributed all endowments. Table 4 compares the contributions for the four treatments using a two-tailed Mann–Whitney test. This test shows that subjects in the MAM contributed significantly more than those in the VCM, with the test statistic $z = 9.445$

13

($p < 0.001$).[21]

<Table 4 about here.>

The contributions in the SMAM show similar patterns to those in the MAM. It took five periods to achieve a contribution rate of over 90%, and almost all subjects maintained full contributions until the end of the session. Spearman's rank correlation test shows that the upward trend in average contributions is statistically significant ($\rho = 0.625, p = 0.004$). Overall, the average contribution rate of the SMAM is 89.9%. Of the 570 outcomes, there were 424 outcomes where both players contributed all endowments. The contribution rate of the SMAM was significantly lower than that of the MAM ($z = 3.799, p < 0.001$).[22] There is no significant difference, however, in the first-stage choices between the SMAM and MAM ($z = 0.456, p = 0.649$). The contribution rate of the SMAM is significantly higher than that of the VCM ($z = 9.275, p < 0.001$).

In contrast to the MAM and the SMAM, only 7 of the 540 pairs of the MCM achieved full contributions and the approval of both players. The average contribution under the MCM fluctuated within a relatively narrow range, but decreased repeatedly, and it remained higher than that of the VCM in every period. Spearman's rank correlation test shows that the downward trend in average contributions is statistically significant ($\rho = -0.525, p < 0.001$). The overall average contribution rate of the MCM is 28.4% (6.83 tokens), which is significantly higher than that of the VCM ($z = 7.284, p < 0.001$). However, the contribution rate was significantly lower than that of the MAM and the SMAM ($z = 9.450$ and $p < 0.001$ for the MCM vs. the MAM and $z = 9.133$ and $p < 0.001$ for the MCM vs. the SMAM).

In reality, the average contributions by subjects in one session are mutually dependent, although in a two-tailed Mann–Whitney test we assume that they are independent. In order to account for this dependence, we also run a regression to predict contribution after the approval stage, using three approval mechanism dummies and clustering standard errors by session. The estimation results are similar to those of a two-tailed Mann–Whitney test, indicating that all three approval mechanisms significantly increase contribution relative to the VCM. In particular, the MAM increases contribution by 20.34 tokens with a standard error of 0.633, followed by the SMAM which increases

---

[21] The null hypothesis is that the contribution rates are the same between the MAM and the VCM. We use the method presented by Andreoni and Miller (1993). We first calculate the average contribution rate of each subject across periods and then calculate the test statistic using these averages, in order to eliminate cross-period correlation.

[22] The statistical difference exists mainly because the lower choices of two of the subjects in the SMAM were disapproved. One of them mistakenly thought that she had to allocate 24 tokens across 19 periods. The other subject intentionally chose low contributions, expecting that his/her matched subjects might approve by giving up a small number of their points.

14

contribution by 19.1 tokens and a standard error of 0.886, and the MCM which increases contribution by 4.38 tokens and a standard error of 0.659.

7. Heterogeneous reasoning processes coexist

We report a third-party assessment and classification of descriptions in the questionnaire during and after the experiment, called *coding*, in order to investigate which type of reasoning process (including diagonalization, regret minimization, and level-*k* thinking) was used by the subjects.[23] The coding proceeded as follows. First, we read the responses in the MAM and the MCM separately. Then, we created 30 sentence-based categories from typical responses (see Online Appendix). Two research assistants (hereafter *coders*) were independently instructed, and each coder created the coding for all 120 subjects in the MAM and the MCM. For each subject, the coders were required to read subjects' descriptions and determine the categories to which his/her descriptions belong and to check all relevant categories (0 = description does not fall into the category, 1 = description falls into the category). We refer to a coder's binary decision as the *rating*. Coders never met face-to-face, and no efforts were made to reconcile the differences in individual coding.

For each category of our interest, we regard a subject rated positively by at least one coder as a follower of the equilibrium strategy or heuristics in the category. In particular, Category 10 is designed to capture subjects in the MAM who directly mentioned the diagonal line of the payoff table. Such subjects are deemed to follow diagonalization heuristics. Category 12 is designed to capture regret minimization in the first stage in the MAM and MCM. Categories 16 and 17 are set to capture level-1 and level-2 in the MCM, respectively.[24,25]

Let us summarize the coding result. In the MAM, 20.0% (= 12/60) of the subjects are deemed to follow diagonalization heuristics, and 16.7% (= 10/60) subjects fall in the regret minimization category. In the MCM, on the other hand, 21.7% (= 13/60) and 45.0% (= 27/60) of the subjects are deemed to follow level-1 and level-2 thinking, respectively. In addition, 11.7% (= 7/60) of the subjects are deemed to follow regret minimization, which appears commonly in both treatments. Although no category of our interest records an absolute majority, these results suggest that subjects are heterogeneous in reasoning

---

[23] This method is similar to that presented by Cooper and Kagel (2005). Note that our experiments did not allow communication among subjects, while Cooper and Kagel (2005) investigated specific words and attitudes in informal communication that facilitated cooperation among group members.

[24] We did not make level-*k* categories for the MAM because we could not find such responses in the questionnaires. The same applies to regret minimization in the second stage for both the MAM and the MCM.

[25] At the start of the coding process, Category 12 and 13 were classified as regret minimization. However, we found that the sentence of Category 13 was too imprecise to capture regret minimization. Therefore, we do not consider Category 13 to describe regret minimization in the analysis.

processes, and the appearance frequency of each reasoning process may depend on the mechanisms.[26]

8. Discussion and conclusion

Prior theoretical work on public good mechanisms has given theorists broad latitude the choice of equilibrium concept or behavioral rule, such as dominant strategy equilibrium, Nash equilibrium, or SPNE. However, recent experimental studies reveal that there is a wide gap between the theory and subjects' behavior.

This paper, in contrast, theoretically showed that the MAM multiply implements the Pareto-efficient provision of the public good in five equilibrium concepts and heuristics, namely BEWDS, limit LAQRE, SPMRE, level-k thinking, and diagonalization heuristics. Consistent with our predictions, we observed rapid convergence towards efficient allocation in the MAM sessions. Further, our inclusive coding of subjects' responses helped pinpoint their reasoning processes. The coding results suggest that the subjects are heterogeneous in reasoning processes. Although our public good environment is simple, our result is in stark contrast to those of recent experimental studies, which have often needed dozens of repetitions to attain efficiency in the laboratory.

To our knowledge, this is the first theoretical and experimental study to successfully design a public good mechanism robust to several behavioral rules. This paper should at least stimulate constructive discussion on mechanism design, to emphasize the alignment of players with various behavioral rules towards a social optimum.

One direction for further research is to evaluate experimental approval mechanisms in more general public good environments. First, if we consider asymmetric cases, Proposition 1 still holds if two players have moderately different preferences towards the public good.[27] Further, if the endowment is heterogeneous, the problem reduces to a symmetric case, through the definition of a new strategy space, where each player/subject chooses a contribution rate relative to his/her endowment in the first stage. If the public good is harmful for some player, the MAM with transfer implements the symmetric Pareto-efficient allocation in BEWDS (see Masuda, 2013). The second extension is toward the n-person case. In public good environments with binary choices, Huang et al.

---

[26] We did not obtain sufficiently high agreement in ratings between any two coders for the following two reasons. First, even though our coders had excellent and comparable game theoretic skills, the sensitivity of coders to track subjects' reasoning processes, or to pick up any subtle nuance in Japanese, is heterogeneous. Second, we made coders work independently throughout the coding process to prevent coordination among them, though experiments in psychology permit communication between two coders before rating each subject.

[27] If we write player i's benefit per public good as α_i, the conditions are $\alpha_1 \neq \alpha_2$, $\alpha_1, \alpha_2 \in (0,1)$, and $\alpha_1 + \alpha_2 > 1$.

(2013) succeeded in constructing a variant of an *n*-person approval mechanism. Masuda (2013) also generalized the mechanism of Huang et al. (2013) to homogeneous *n*-person public good environments with continuous contributions and linear payoffs in the spirit of the SMAM, where players other than those with the minimum choice can proceed to the second stage. Third, in a more realistic setting, we may consider the ratification process for international agreements involving countries; nature determines the order of the player who may decide to approve or otherwise. Our results hold in such cases also.

Appendix

This appendix gives the proofs for Propositions 2 and 3.

Proof of Proposition 2. Let g be a forthright approval mechanism that implements the Pareto-efficient allocation in BEWDS and that satisfies voluntariness and monotonicity. Let (s_1, s_2) be any first-stage choice. If $s_1 = s_2$, $g((s_1, \cdot), (s_2, \cdot)) = (s_1, s_2)$ according to condition (i) of forthrightness, which shows the outcome of the MAM. Consider the case where $s_1 \neq s_2$. Without loss of generality, assume that $s_1 > s_2$, and consider two cases: $s_1 = w$ and $s_1 < w$. First, consider subgame (w, s_2) with $w > s_2$. In order to implement the efficient allocation in BEWDS, w must weakly dominate s_2 for player 2 in the reduced normal form game. In other words,

(A.1) $\quad u_2^g((w, \cdot), (w, \cdot)) \geq u_2^g((w, t_1), (s_2, t_2))$

where u_i^g is i's payoff function induced by g, and t_i is player i's decision under BEWDS.

Suppose that $t_1 = t_2 = y$. Then, by conditions (i) and (ii) of forthrightness, we have

(A.2) $\quad u_2^g((w, y), (s_2, y)) = (w - s_2) + \alpha(w + s_2) > (w - w) + \alpha(w + w) = u_2^g((w, \cdot), (w, \cdot))$,

which contradicts to (A.1). Thus, $t_1 = n$ or $t_2 = n$. Together, condition (iii) of forthrightness and (A.2) show that $u_2^g((w, y), (s_2, y)) > u_2^g((w, y), (s_2, n)) = u_2^g((w, n), (s_2, y)) = u_2^g((w, n), (s_2, n))$. Hence, y weakly dominates n for player 2 in the second stage, namely $t_2 = y$, which implies $t_1 = n$.

Consider now the reduced normal form game under g, and compare the outcomes in subgames (s_2, s_2) and (w, s_2). Since player 1 chooses w by assumption, we have $u_1^g((w, n), (s_2, y)) \geq u_1^g((s_2, \cdot), (s_2, \cdot))$. Let $g((w, n), (s_2, y)) = (g_1, g_2)$. Then, $u_1^g((w, n), (s_2, y)) = w - g_1 + \alpha(g_1 + g_2) \geq w - s_2 + \alpha(s_2 + s_2) = u_1^g((s_2, \cdot), (s_2, \cdot))$. Since g is voluntary, $(g_1, g_2) \leq (w, s_2)$. Because of the monotonicity of g, $(g_1, g_2) \geq (s_2, s_2)$. The intersection of these three inequalities shows $(g_1, g_2) = (s_2, s_2)$. Thus, according to

17

condition (iii) of forthrightness, we have

(A.3) $g((w,n),(s_2,y)) = g((w,y),(s_2,n)) = g((w,n),(s_2,n)) = (s_2,s_2).$

In other words, both players contribute s_2, which is the minimum of (w,s_2), when at least one player chooses n.

Now take any $s_1 \in (s_2,w)$. Together with $(w,s_2) \geq (s_1,s_2) \geq (s_2,s_2)$ and (A.3), since g is monotonic, we have $(s_2,s_2) = g((w,n),(s_2,y)) \geq g((s_1,n),(s_2,y)) \geq g((s_2,n),(s_2,y)) = (s_2,s_2).$ Hence, we have $g((s_1,n),(s_2,y)) = (s_2,s_2)$. In other words, g is the MAM. ∎

Proof of Lemma 1. Suppose, on the contrary, that for some SPNE $p = ((s_1,t_1(\cdot)),(s_2,t_2(\cdot)))$, players contribute differently, i.e., $g_1 \neq g_2$. Then, by construction of the MAM, $s_1 \neq s_2$ and $t_i(s_1,s_2) = y$ for $i = 1,2$. Define $p' = (p_1', p_2)$ with $p_1' = (s_1', t_1(\cdot))$ as the strategy profile where player 1 deviates from p only in the first stage. Without loss of generality, assume $s_1 > s_2$. Assume $s_1' = s_2$, then $u_1(p') = w - s_2 + \alpha(s_2 + s_2) > w - s_1 + \alpha(s_1 + s_2) = u_1(p)$ by $s_1 > s_2$, which is a contradiction. Thus, $g_1 = g_2$. Next, we show that any symmetric pair of contributions can be supported as an SPNE. Let $s \in S$. Consider the strategy profile $q = ((s,t(\cdot)),(s,t(\cdot)))$, where $t(\cdot)$ is a function such that $t(s_1,s_2) = n$ for all $s_1,s_2 \in S$. Clearly, p constitutes a Nash equilibrium (n,n) for every second-stage subgame. Now, consider the reduced normal form game. Suppose that player 1 unilaterally deviates to $s_1' > s$ in the first stage. Let $q_1' = (s_1',t(\cdot))$ and let $q' = (q_1',q_2)$. Since both players choose n at (s_1',s) and by $s_1' > s$, we have $u_1(q') = w - s + \alpha(s+s) = u_1(q)$. Suppose $s_1' < s$. Then, we have $u_1(q') = w - s_1' + \alpha(s_1' + s_1') < w - s + \alpha(s+s) = u_1(q)$. The above argument also applies to player 2. Hence, q is an SPNE. ∎

Proof of Lemma 2. First, consider the second stage. Take any (s_1,s_2). Note that $\sigma_j^2(n) = 1 - \sigma_j^2(y)$, $u_i((s_i,y),(s_j,y)) = (w - s_i) + \alpha(s_i + s_j)$ and if $(t_i,t_j) \neq (y,y)$, $u_i((s_i,t_i),(s_j,t_j)) = (w - \underline{s}) + \alpha(\underline{s} + \underline{s})$ $= \pi_i(n)$ where $\underline{s} = \min\{s_i,s_j\}$. Then,

(A.4) $\pi_i(y) = \sum_{t_j \in \{y,n\}} \sigma_j^2(t_j) u_i((s_i,y),(s_j,t_j)) = \sigma_j^2(y)\{(w - s_i) + \alpha(s_i + s_j)\} + \{1 - \sigma_j^2(y)\}\pi_i(n).$

By substituting (A.4) into the second equation satisfied by the LAQRE, we have

(A.5) $\sigma_i^2(y) = \left[1 + \exp\{(\pi_i(n) - \pi_i(y))/\mu\}\right]^{-1} = \begin{cases} \left[1 + \exp\{\alpha(s_i - s_j)\sigma_j^2(y)/\mu\}\right]^{-1} & \text{if } s_i < s_j, \\ 0.5 & \text{if } s_1 = s_2, \\ \left[1 + \exp\{(1-\alpha)(s_i - s_j)\sigma_j^2(y)/\mu\}\right]^{-1} & \text{if } s_i > s_j. \end{cases}$

For each $s_i \in S_d$, $\pi_i(s_i) = \sum_{s \in S_d} \sigma_j^1(x) \pi_i(s_i, x)$, where $\pi_i(s_i, x) = \sigma_i^2(y) \sigma_j^2(y) u_i((s_i, y), (x, y)) +$

$(1 - \sigma_i^2(y) \sigma_j^2(y)) u_i((s_i, n), (x, n))$. Since

$$\pi_i(s_i, s_j) = \sigma_i^2(y) \sigma_j^2(y) \{(w - s_i) + \alpha(s_i + s_j)\} + (1 - \sigma_i^2(y) \sigma_j^2(y)) \{(w - \underline{s}) + \alpha(\underline{s} + \underline{s})\}$$

$$= \begin{cases} \alpha \sigma_i^2(y) \sigma_j^2(y)(s_j - s_i) + (2\alpha - 1)\underline{s} + w & \text{if } s_i < s_j, \\ (2\alpha - 1)\underline{s} + w & \text{if } s_1 = s_j, \\ (1 - \alpha)\sigma_i^2(y) \sigma_j^2(y)(s_j - s_i) + (2\alpha - 1)\underline{s} + w & \text{if } s_i > s_j, \end{cases}$$

we have

$$\pi_i(s_i) = \sum_{x \in S_d} \sigma_j^1(x) \pi_i(s_i, x) = \sum_{x \le s_i} \sigma_j^1(x) \pi_i(s_i, x) + \sum_{x > s_i} \sigma_j^1(x) \pi_i(s_i, x)$$

(A.6)
$$= \sum_{x \in S_d} \sigma_j^1(x) \{(2\alpha - 1) \min\{s_i, x\} + w\} + (1 - \alpha) \sum_{x \le s_i} \sigma_j^1(x) \sigma_i^2(y) \sigma_j^2(y)(x - s_i)$$

$$+ \alpha \sum_{x > s_i} \sigma_j^1(x) \sigma_i^2(y) \sigma_j^2(y)(x - s_i).$$

Now, let us consider limiting equilibrium. If $s_1 = s_2$, by (A.5), $(\sigma_1^2(y), \sigma_2^2(y))$

$= (0.5, 0.5)$ for any μ. Consider next that $s_1 < s_2$. Consider player 2's equilibrium strategy.

Note by $s_1 < s_2$, and $\alpha(s_1 - s_2) \sigma_2^2(y)/\mu < 0$, we have $0.5 < \sigma_1^2(y) < 1$ for any μ. Define a

function of x, $f(x; s_1, s_2) = [1 + \exp\{(1 - \alpha)(s_2 - s_1)x\}]^{-1}$. Then, by (A.5), $0.5 < \sigma_1^2(y) < 1$ for any

μ, and since f is decreasing, $f(0.5/\mu; s_1, s_2) > \sigma_2^2(y) = f(\sigma_1^2(y)/\mu; s_1, s_2) > 0$ for any μ.

Moreover, $\lim_{\mu \to 0} f(0.5/\mu; s_1, s_2) = 0$. Then, $\lim_{\mu \to 0} \sigma_2^2(y) = 0$. Consider player 1's equilibrium

strategy. Note $f(0.5/\mu; s_1, s_2)/\mu > \sigma_2^2(y)/\mu = f(\sigma_1^2(y)/\mu; s_1, s_2)/\mu > 0$. By letting $y = 0.5/\mu$,

$\lim_{\mu \to 0} f(0.5/\mu; s_1, s_2)/\mu = \lim_{y \to \infty} 2y / \{1 + \exp((1-\alpha)(s_2 - s_1)y)\} = \lim_{y \to \infty} 2 / \{1/y + \exp((1-\alpha)(s_2 - s_1)y)/y\} = 0$.

Then, $\lim_{\mu \to 0} \sigma_2^2(y)/\mu = 0$. By substituting this into (A.5), we have

$\lim_{\mu \to 0} \sigma_1^2(y) = \left[1 + \exp\{\alpha(s_1 - s_2) \lim_{\mu \to 0}(\sigma_2^2(y)/\mu)\} \right]^{-1} = [1 + \exp(0)]^{-1} = 0.5$. Therefore, $(\sigma_1^2(y), \sigma_2^2(y))$

$= (0.5, 0)$ in limiting equilibrium. Similarly, if $s_1 > s_2$, $(\sigma_1^2(y), \sigma_2^2(y)) = (0, 0.5)$ in limiting

equilibrium.

Next, we will show that for each i, $\sigma_i^1(w) \to 1$ as $\mu \to 0$. For each i, define a

function of x, $h_i(x; s_i, \mu) = \sigma_j^1(x) \sigma_i^2(y) \sigma_j^2(y)(x - s_i)$. By the above argument, for any

$s_i, x \in S_d$ with $x \ne s_i$, we have

$\lim_{\mu \to 0} h_i(x; s_i, \mu) = \lim_{\mu \to 0} \sigma_j^1(x) \cdot \lim_{\mu \to 0} \sigma_m^2(y) \cdot \lim_{\mu \to 0} \sigma_k^2(y) \cdot \lim_{\mu \to 0}(x - s_i) = \lim_{\mu \to 0} \sigma_j^1(x) \cdot 0.5 \cdot 0 \cdot \lim_{\mu \to 0}(x - s_i) = 0$

where $(m, k) = (i, j)$ if $x > s_i$ and $(m, k) = (j, i)$ otherwise. Finally, $h_i(s_i; s_i, \mu) = 0$. Hence,

$h_i(x; s_i, \mu)$ converges to the constant function 0 on S as $\mu \to 0$. Then, the second and third

terms in (A.6) also converge to 0. Then, for each i and any $s_i \in S_d$, we have

$\pi_i(s_i) \to \sum_{x \in S_d} \sigma_j^1(x) \{(2\alpha - 1) \min\{s_i, x\} + w\}$ as $\mu \to 0$.

Now we assume on the contrary that there exists i and $s < w$ with $s = \min\{x \in S_d \mid \sigma_i^1(x) > 0\}$ in the limiting equilibrium. Let $s' \in \{s+1, \ldots, w\}$. Then, we have

(A.6)
$$\pi_i(s') - \pi_i(s) \to \sum_{x \in S_d} \sigma_j^1(x)(2\alpha - 1)(\min\{s', x\} - \min\{s, x\})$$
$$= \sum_{s \le x \le s'-1} \sigma_j^1(x)(2\alpha - 1)(x - s) + \sum_{s' \le x \le w} \sigma_j^1(x)(2\alpha - 1)(s' - s) \text{ as } \mu \to 0.$$

Together with $\alpha \in (0.5, 1)$, $s < s'$, and $\sigma_i^1(x) \ge 0$ for each i and any $x \in S_d$, (A.6) is nonnegative. By the definition of limit LAQRE and the assumption of $\sigma_i^1(s) > 0$, (A.6) is nonpositive. Then, (A.7) must be 0. Then, $\sigma_i^1(s') = \sigma_i^1(s) > 0$ in the limiting equilibrium. Together with $s = \min\{x \in S_d \mid \sigma_i^1(x) > 0\}$ and the definition of limit LAQRE, it implies that σ_i^1 is the uniform distribution over $\{s, \ldots, w\}$, that is, $\sigma_i^1(x) = 1/(w-s+1)$ if $x \in \{s, \ldots, w\}$ and $\sigma_i^1(x) = 0$ otherwise. By substituting $\sigma_j^1(x) = 1/(w-s+1)$ into (A.6), $\pi_i(s') - \pi_i(s)$ converges to some strictly positive value, a contradiction to $\sigma_i^1(s') = \sigma_i^1(s) > 0$. Hence, for each i, $\sigma_i^1(w) \to 1$ as $\mu \to 0$. Therefore, the unique limit LAQRE outcome is efficient. ∎

Proof of Lemma 3. Let $p = ((w, t_1(\cdot)), (w, t_2(\cdot)))$ with for each i, $t_i(s_1, s_2) = y$ if $s_i < s_j$, $t_i(s_1, s_2) = y$ or n if $s_i = s_j$, $t_i(s_1, s_2) = n$ otherwise. If $s_1 = s_2$, $g((s_1, \cdot), (s_2, \cdot)) = (s_1, s_2)$, and hence both players are indifferent between y and n. Next, consider the case $s_1 \ne s_2$. Since t_i prescribes weakly dominant choice in every second stage, for each i, $j \ne i$ and each $t_j' \in \{y, n\}$, $u_i(t_i(s_1, s_2), t_j' \mid s_1, s_2) = \max_{t_i'} u_i(t_i', t_j' \mid s_1, s_2)$. Then, for each i, $t_i'' \ne t_i$ implies $R_i(t_i(s_1, s_2), \cdot \mid s_1, s_2) = 0 < R_i(t_i'', y \mid s_1, s_2)$. Hence, p prescribes the SPMRE choice in every second stage. Then, the reduced normal form game has an echelon structure. Since $u_i(w, w \mid t_1, t_2) = 2\alpha w > w + (2\alpha - 1)s_i = u_i(s_i, w \mid t_1, t_2)$ for each i, $j \ne i$ and any $s_i \ne w$, we have $\max_{s_j \in S} R_i(w, s_j \mid t_1, t_2) = 0 < (2\alpha - 1)(w - s_i) = R_i(s_i, w \mid t_1, t_2) = \max_{s_j \in S} R_i(s_i, s_j \mid t_1, t_2)$. ∎

Proof of Proposition 4. Let g be the mate choice mechanism (MCM).
(i) BEWDS. Consider the second stage. Take any $a, b \in S$. Note that $(w - a) + \alpha(a + b) \ge w$ if and only if $a \le \beta^{-1}b$. Then, $t_1(a, b) = y$ if $a \le \beta^{-1}b$ and $t_1(a, b) = n$ otherwise.[28] Then, in the reduced normal form game, $u_1(a, b) = (w - a) + \alpha(a + b)$ and $u_2(a, b) = (w - b) + \alpha(a + b)$ if $\beta b \le a \le \beta^{-1}b$ and $u_1(a, b) = u_2(a, b) = w$ otherwise. A similar argument holds for player 2.

---

[28] We assume that a player chooses y if y and n are indifferent.

First, we show that for each i, choice βw weakly dominates any choice $s_i \in \{0\} \cup (\beta w, w]$. Take any i and $s_i \in (\beta w, w]$. If $s_j \in [\beta s_i, w]$, then $\beta s_j < s_i \leq \beta^{-1} s_j$. Together with $\beta s_j \leq \beta w \leq \beta^{-1} s_j$, it implies that $u_i(\beta w, s_j) = (w - \beta w) + \alpha(\beta w + s_j) > (w - s_i) + \alpha(s_i + s_j) = u_i(s_i, s_j)$. If $s_j \in (\beta^2 w, \beta s_i)$, then we have $\beta s_j < \beta w < \beta^{-1} s_j$ and $s_i > \beta^{-1} s_j$. Then, $u_i(\beta w, s_j) = (w - \beta w) + \alpha(\beta w + s_j) > w = u_i(s_i, s_j)$. If $s_j \in [0, \beta^2 w]$, then we have $\beta w \geq \beta^{-1} s_j$ and $s_i > \beta^{-1} s_j$. Then, $u_i(\beta w, s_j) = w = u_i(s_i, s_j)$. Next, assume $s_i = 0$. Then, $\beta s_j > s_i$, for any $s_j > 0$. Then, we have $u_i(\beta w, s_j) \geq w = u_i(0, s_j)$ for any $s_j \geq 0$, with strict inequality for $s_j > \beta^2 w$. Second, for any $s_i \in (0, \beta w]$, s_i is the unique best response to $\beta^{-1} s_i$ in the reduced normal form game and hence s_i survives under BEWDS. Therefore, we get the desired result.

(ii) SPNE. Consider the second stage. Take any SPNE $p = ((s_1, t_1(\cdot)), (s_2, t_2(\cdot)))$. If $s_1 > \beta^{-1} s_2$ or $s_2 > \beta^{-1} s_1$, some player chooses n in subgame (s_1, s_2), otherwise the player does not provide the best response. Hence, the players finally contribute nothing in p. Next, take any (s_1, s_2) such that $\beta s_2 \leq s_1 \leq \beta^{-1} s_2$. We show that there exists an SPNE where both players approve and hence contribute (s_1, s_2). Consider the strategy profile $q = ((s_1, t(\cdot)), (s_2, t(\cdot)))$ with $t(s_1, s_2) = y$ and n otherwise. Clearly, q constitutes a Nash equilibrium for every second-stage subgame. Now, consider the reduced normal form game. Assume player 1 deviates to $s_1' \neq s_1$. By construction of q, such a deviation results in (n, n). Together with $s_1 \leq \beta^{-1} s_2$, this implies that $u_1(q) = (w - s_1) + \alpha(s_1 + s_2) \geq w = u_1(s_1', q)$. Hence, choosing s_1 is the best response to s_2. This is also true for player 2. Hence, q is an SPNE with contributions (s_1, s_2).

(iii) LAQRE. Let $S_d = \{0, 1, \dots, w\}$. Take any $s_1, s_2 \in S_d$. If $(t_i, t_j) \neq (y, y)$, $u_i((s_i, t_i), (s_j, t_j)) = w = \pi_i(n)$. Then, $\pi_i(y) = \alpha \sigma_j^2(y)(s_j - \beta s_i) + \pi_i(n)$. Together, this and the definition of LAQRE imply that

(A.8) $\sigma_i^2(y) = \{1 + \exp((\pi_i(n) - \pi_i(y))/\mu)\}^{-1} = \{1 + \exp(-\alpha \sigma_j^2(y)(s_j - \beta s_i)/\mu)\}^{-1}$.

On the other hand, by the first equation satisfied by LAQRE, for each $s_i \in S$,

we have $\pi_i(s_i) = \sum_{x \in S_d} \sigma_j^1(x) \pi_i(s_i, x)$, where $\pi_i(s_i, x) = \sigma_i^2(y) \sigma_j^2(y) u_i((s_i, y), (x, y)) + (1 - \sigma_i^2(y) \sigma_j^2(y)) u_i((s_i, n), (x, n))$. Since $\pi_i(s_i, x) = \sigma_i^2(y) \sigma_j^2(y)\{(w - s_i) + \alpha(s_i + s_j)\} + (1 - \sigma_i^2(y) \sigma_j^2(y)) w = \alpha \sigma_i^2(y) \sigma_j^2(y)(s_j - \beta s_i) + w$, we have

(A.9) $\pi_i(s_i) = \sum_{x \in S_d} \sigma_j^1(x) \pi_i(s_i, x) = \alpha \sum_{x \in S_d} \sigma_j^1(x) \sigma_i^2(y) \sigma_j^2(y)(x - \beta s_i) + w \sum_{x \in S_d} \sigma_j^1(x)$

Now let us consider limiting equilibrium. Consider first $(s_1, s_2) = (0, 0)$. Then,

$(\sigma_1^2(y),\sigma_2^2(y))=(0,0)$ constitutes the unique limiting equilibrium in the second stage. Consider next $(s_1,s_2)\neq(0,0)$. If $s_i\in(\beta s_j,\beta^{-1}s_j)$ for each $i=1,2$, $j\neq i$, $(\sigma_1^2(y),\sigma_2^2(y))=(1,1)$ constitutes the unique limiting equilibrium. Similarly, if $\beta^{-1}s_i<w$ and $s_j\in(\beta^{-1}s_i,w]$, then $(\sigma_i^2(y),\sigma_j^2(y))=(0.5,0)$ constitutes the limiting equilibrium. If $s_i=\beta^{-1}s_j$, then $(\sigma_i^2(y),\sigma_j^2(y))=(0.5,1)$ constitutes the limiting equilibrium.

In order to get the first-stage behavioral strategy in limiting equilibrium, for each i, let us see $\sigma_i^2(y)\sigma_j^2(y)(x-\beta s_i)$ as a function of $x\in S_d$. Then, by the limiting equilibrium choices in the second stage, we have $\sigma_i^2(y)\sigma_j^2(y)(x-\beta s_i)\to x-\beta s_i$ as $\mu\to0$ if $x\in[\beta s_i,\beta^{-1}s_i)$, $\sigma_i^2(y)\sigma_j^2(y)(x-\beta s_i)\to 0.5(\beta^{-1}-\beta)s_i$ as $\mu\to0$ if $x=\beta^{-1}s_i$, and $\sigma_i^2(y)\sigma_j^2(y)(x-\beta s_i)\to0$ as $\mu\to0$ otherwise. By substituting these results into (A.9), if $\beta^{-1}s_i$ is an integer, $\pi_i(s_i)\to\alpha\sum_{x\in S_d\cap[\beta s_i,\beta^{-1}s_i)}\sigma_j^1(x)(x-\beta s_i)+0.5\alpha\sigma_j^1(\beta^{-1}s_i)(\beta^{-1}-\beta)s_i+w$ as $\mu\to0$

and otherwise $\pi_i(s_i)\to\alpha\sum_{x\in S_d\cap[\beta s_i,\beta^{-1}s_i)}\sigma_j^1(x)(x-\beta s_i)+w$ as $\mu\to0$. Next, we show that there exists a limiting equilibrium where both players choose 1 with probability 1. Suppose that for each i, $\sigma_i^1(1)\to1$ as $\mu\to0$. By $\beta<1$, we have for each i,

$$\pi_i(1)-\pi_i(0)\to\alpha\sum_{x\in S_d\cap[\beta,\beta^{-1})}\sigma_j^1(x)(x-\beta)+w-w=\alpha(1-\beta)>0.$$

Let $s_i>1$. We have for each i,

$$\begin{aligned}\pi_i(1)-\pi_i(s_i)&\to\alpha\sum_{x\in S_d\cap[\beta,\beta^{-1})}\sigma_j^1(x)(x-\beta)+w-\alpha\sum_{x\in S_d\cap[\beta s_i,\beta^{-1}s_i)}\sigma_j^1(x)(x-\beta s_i)-w\\ &\geq\alpha\cdot1\cdot(1-\beta)-\alpha\cdot1\cdot(1-\beta s_i)\\ &=\alpha\beta(s_i-1)>0\end{aligned}$$

where the equality holds if $\beta s_i\leq1$. Therefore, we get the desired result. ∎

(iv) SPMRE. Let $p=(\beta w,t_1(\cdot)),(\beta w,t_2(\cdot)))$ be a strategy profile with for each i, $j\neq i$, $t_i(s_1,s_2)=y$ if $s_i\leq\beta^{-1}s_j$ and $t_i(s_1,s_2)=n$ otherwise. If $s_i=\beta^{-1}s_j$, player i is indifferent between y and n. Next, consider the case $s_i\neq\beta^{-1}s_j$. Since t_i prescribes weakly dominant choice in every second stage, for each i, $j\neq i$ and each $t_j'\in\{y,n\}$, $t_i'\neq t_i(s_1,s_2)$ implies $u_i(t_i(s_1,s_2),t_j'\mid s_1,s_2)=\max_{t_i''}u_i(t_i'',t_j'\mid s_1,s_2)>u_i(t_i',y\mid s_1,s_2)$. Hence, p prescribes the unique SPMRE choice in every second stage.

Consider the reduced normal form game, which is identical to the one induced by BEWDS. Take any $s_1,s_2\in S$. As shown in part (i), player i best responds to player j if $s_i=\beta s_j$. Then, for each i, $\max_{s_i}u_i(s_i,s_j\mid t_1,t_2)=w-\beta s_j+\alpha(\beta s_j+s_j)=w+(1-\beta)s_j$. Consider player 1's regret. If $s_1>\beta^{-1}s_2$ or $s_2>\beta^{-1}s_1$, either player chooses n and hence $u_1(s_1,s_2)=w$. Then, $R_1(s_1,s_2)=(w+(1-\beta)s_2)-w=(1-\beta)s_2$. If $\beta s_2\leq s_1\leq\beta^{-1}s_2$, both players choose y in the second stage. Then, we have $u_1(s_1,s_2)=w-s_1+\alpha(s_1+s_2)$ and $R_1(s_1,s_2)=(w+(1-\beta)s_2)-\{w-s_1+\alpha(s_1+s_2)\}=(\beta s_1-\beta^2 s_2)/(1+\beta)$, which is maximized at

$s_2 = \beta s_1$. Hence, player 1's maximum regret at s_1 is given by

$$\max_{s_2} R_1(s_1, s_2) = \begin{cases} R_1(s_1, w) = (1-\beta)w \text{ if } s_1 < \beta w, \\ R_1(s_1, \beta s_1) = \beta(1-\beta)s_1 \text{ otherwise.} \end{cases}$$

By $\alpha \in (0.5, 1)$, we have $\beta < 1$. Together with $s_1 \leq w$, it implies that $s_1 = \beta w$ minimizes the maximum regret and gives the solution. The same argument holds for player 2. Therefore, $(\beta w, \beta w)$ constitutes the unique minimax regret equilibrium of the reduced normal form game.

(v) Level-k. First, we show that βw is the best response to level-0 player. Given that the other player uniformly randomizes in the reduced normal form game, by a similar argument in (iii), the expected payoff $\pi_i'(s_i)$ when player i chooses s_i is

$$\pi_i'(s_i) = \int_{\beta s_i}^{\min\{w, \beta^{-1}s_i\}} (x - \beta s_i) / w \, dx + 1 = (\min\{w, \beta^{-1}s_i\} - \beta s_i)^2 / 2w + 1.$$

Then, $s_i = \beta w$ maximizes $\pi_i'(s_i)$. Let $B(k)$ be a sequence of the level-k player's contribution where $k \geq 1$. Then, $B(1) = \beta w$ and $B(k) = BR(B(k-1))$ until $B(k) = B(k-1)$, where $BR(.)$ is the best response to $B(k-1)$ in the reduced normal form game induced by elimination of weakly dominated strategies. Since $BR(s) = \beta s$ for any $s \in S$, we get the desired result. ∎

References
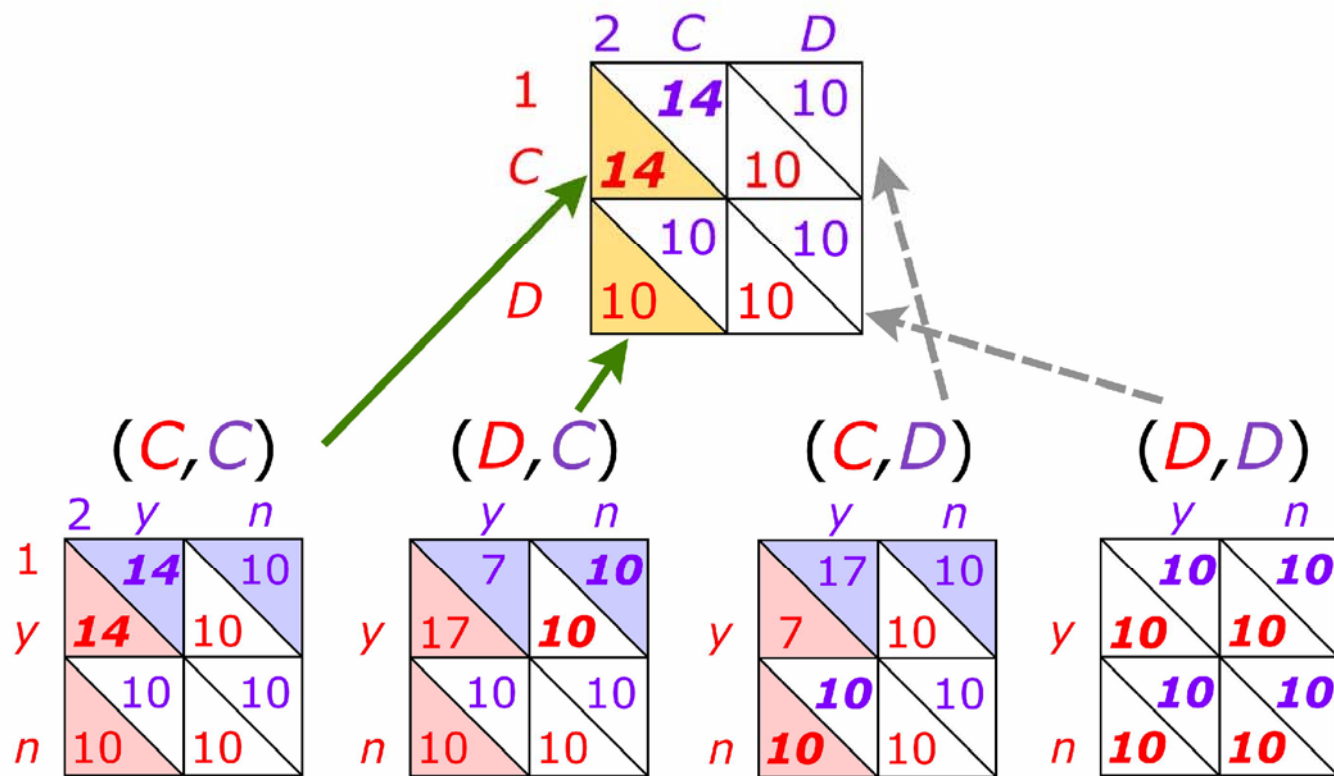
Anderson, S., Goeree, J.K., Holt, C.A., 2001. Minimum-effort coordination games: stochastic potential and logit equilibrium. Games Econ. Behav. 34 (2), 177-199.

Andreoni, J., Miller, J.H., 1993. Rational cooperation in the finitely repeated prisoner's dilemma: experimental evidence. Econ. J. 103 (418), 570-585.
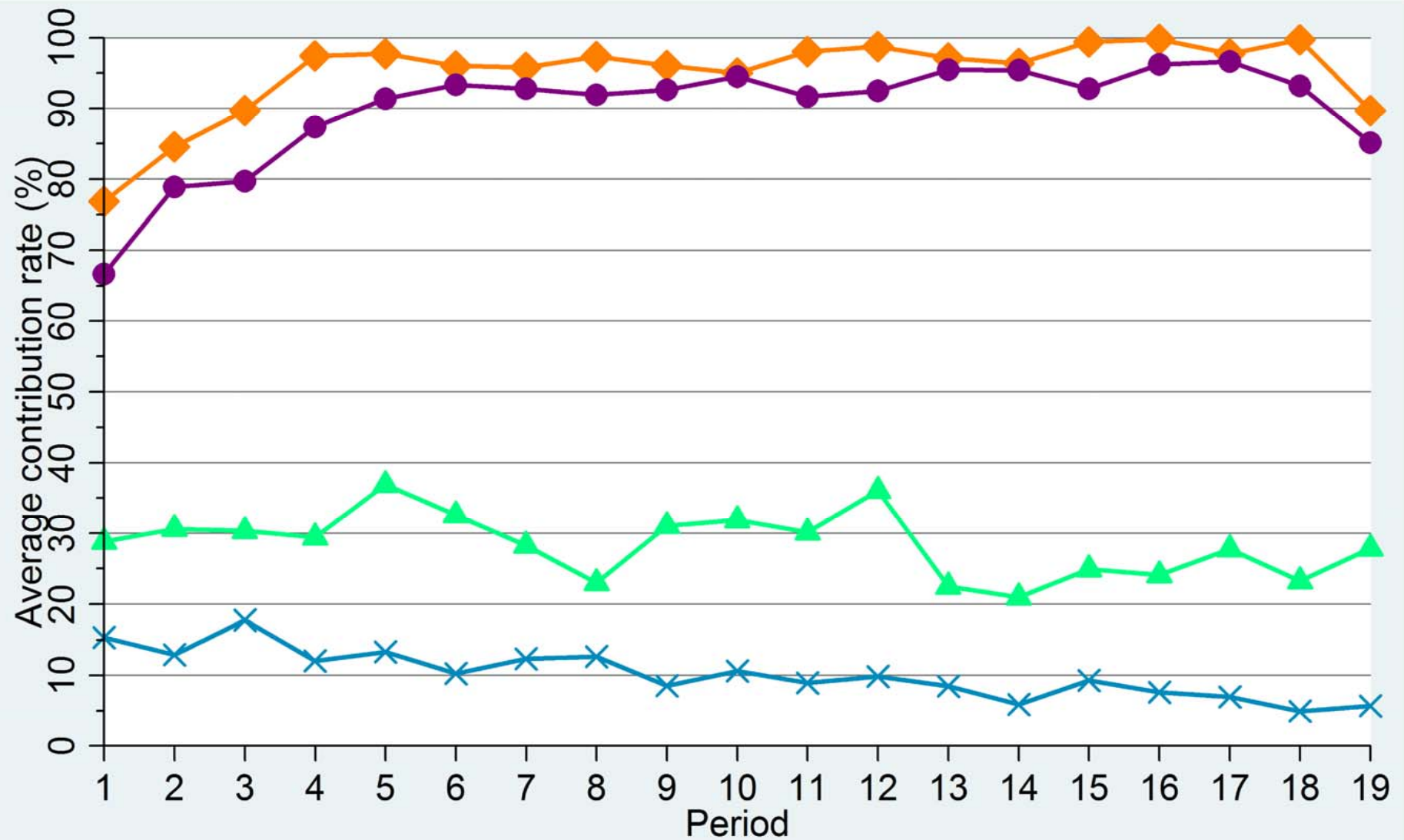
Andreoni, J., Varian, H., 1999. Preplay contracting in the prisoners' dilemma. Proc. Natl. Acad. Sci. 96 (19), 10933-10938.

Attiyeh, G., Robert, F., Isaac, R.M., 2000. Experiments with the pivot process for providing public goods. Public Choice. 102 (1), 93-112.

Cason, T.N., Saijo, T., Sjöstrom, T., Yamato, T., 2006. Secure implementation experiments: do strategy-proof mechanisms really work? Games Econ. Behav. 57 (2), 206-235.

Chen, Y., 2008. Incentive-Compatible Mechanisms for Pure Public Goods: A Survey of Experimental Literature. Plott, C.R. and Smith, V.L. (Eds.), The Handbook of Experimental Economics Results, Elsevier, 625-643.

Chen, Y., Gazzale, R., 2004. When does learning in games generate convergence to Nash equilibrium? The role of supermodularity in an experimental setting. Amer. Econ. Rev. 94 (5), 1505-1535.

Chen, Y., Plott, C.R., 1996. The Groves–Ledyard mechanism: an experimental study of institutional design. J. Public Econ. 59 (3), 335-364.

Chen, Y., Tang, F., 1998. Learning and incentive compatible mechanisms for public goods provision: an experimental study. J. Polit. Economy. 106 (3), 633-662.

Cooper, D.J., Kagel, J.H., 2005. Are two heads better than one? Team versus individual play in signaling games. Amer. Econ. Rev. 95 (3), 477-509.

Costa-Gomes, M.A., Crawford, V.P., 2006. Cognition and behavior in two-person guessing games: an experimental study. Amer. Econ. Rev. 96 (5), 1737-1768.

Falkinger, J., 1996. Efficient private provision of public goods by rewarding deviations from average. J. Public Econ. 62 (3), 413-422.

Falkinger, J., Fehr, E., Gachter, S., Winter-Ebmer, R., 2000. A simple mechanism for the efficient provision of public goods: experimental evidence. Amer. Econ. Rev. 90 (1), 247-264.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exper. Econ. 10 (2), 171-178.

Groves, T., Ledyard, J., 1977. Optimal allocation of public goods: a solution to the "free rider" problem. Econometrica. 45 (4), 783-809.Healy, P.J., Mathevet, L., 2012. Designing stable mechanisms for economic environments. Theoretical Econ. 7, 609-661.

Healy, P.J., 2006. Learning dynamics for mechanism design: an experimental comparison of public goods mechanisms. J. Econ. Theory. 129 (1), 114-149.

Huang, X., Masuda, T., Okano, Y., Saijo, T., 2013. Toward solving social dilemma: theory and experiment. Unpublished manuscript.

Kalai, E., 1981. Preplay negotiations and the prisoner's dilemma. Math. Soc. Sci. 1 (4), 375-379.

Ledyard, J.O., 1995. Public goods: a survey of experimental research, in: Kagel J., Roth, A. (Eds.), The Handbook of Experimental Economics, Princeton University Press, Princeton, pp. 111-194.

Masuda, T., 2013. A note on extensions of the minimum approval mechanism. https://sites.google.com/site/tmasudaexpecon/research.

Masuda, T., Okano, Y., Saijo, T., 2013. The minimum approval mechanism implements the efficient public good allocation theoretically and experimentally. ISER Discussion Paper No. 874.

McKelvey, R.D., Palfrey, T.R., 1995. Quantal response equilibria for normal form games. Games Econ. Behav. 10 (1), 6-38.

McKelvey, R.D., Palfrey, T.R., 1998. Quantal response equilibria for extensive form games. Exper. Econ. 1 (1), 9-41.

Renou, L., Schlag, K.H., 2011. Implementation in minimax regret equilibrium. Games Econ. Behav. 71 (2), 527-533.

Saijo,T., Tatamitani, Y., Yamato, T., 1996. Toward natural implementation. Int. Econ. Rev. , 37(4), 949-80.

Saijo, T., Yamato, T., Sjöström, T., 2007. Secure implementation. Theoretical Econ. 2 (3), 203-229.

Saijo, T., Okano, Y., Yamakawa, T., 2013. The mate choice mechanism experiment: a solution to prisoner's dilemma. Unpublished manuscript.

Varian, H., 1994. A solution to the problem of externalities when agents are well-informed. Amer. Econ. Rev. 84 (5), 1278-1293.

Walker, M., 1981. A simple incentive compatible scheme for attaining Lindahl allocations. Econometrica. 49 (1), 65-71.

Figure 1. Prisoner's dilemma game with the MCM in Saijo et al. (2013).

Figure 2. Average contributions by period, sorted by mechanism.

Legend:
- Mate Choice (28.4%)
- Mimimum Approval (94.9%)
- Simplified Minimum Approval (89.9%)
- Voluntary Contribution (10.2%)

	0	2	4	6	8	10
0	10	11.4	12.8	14.2	15.6	17
2	9.4	10.8	12.2	13.6	15	16.4
4	8.8	10.2	11.6	13	14.4	15.8
6	8.2	9.6	11	12.4	13.8	15.2
8	7.6	9	10.4	11.8	13.2	14.6
10	7	8.4	9.8	11.2	12.6	14

Table 1. Player 1's payoff table when $w = 10$ and $\alpha = 0.7$.

	y	n
y	(9.6,13.6)	(10.8,10.8)
n	(10.8,10.8)	(10.8,10.8)

Subgame (6,2)

	y	n
y	(14.6,12.6)	(13.2,13.2)
n	(13.2,13.2)	(13.2,13.2)

Subgame (8,10)

Table 2. Payoff tables in subgames (6,2) and (8,10).

	0	2	4	6	8	10
0	10	10	10	10	10	10
2	10	10.8	10.8	10.8	10.8	10.8
4	10	10.8	11.6	11.6	11.6	11.6
6	10	10.8	11.6	12.4	12.4	12.4
8	10	10.8	11.6	12.4	13.2	13.2
10	10	10.8	11.6	12.4	13.2	14

Table 3. Player 1's payoff table in the reduced normal form game under the MAM.

Treatment	SMAM	MCM	VCM
MAM	3.799**	9.450**	9.445**
SMAM		9.133**	9.275**
MCM			7.284**

Notes: ** $p < 0.01$.

Table 4. Two-tailed Mann–Whitney test statistics for the equality of average contributions, grouped by treatment.