Altruistic Preferences in the Dictator Game: Replication of Andreoni and Miller (2002) in Japan

Jiayi Xue[1], Yohsuke Ohtsubo[1,*]

[1]Kobe University, 1-1 Rokkodai-cho, Nada, Kobe 657-8501, Japan

[*]Author for correspondence (yohtsubo@lit.kobe-u.ac.jp)

We conducted two replication studies of Andreoni and Miller's (2002) modified dictator game study, which revealed that participants' altruistic decisions were consistent with the notion of utility maximization. The two studies (Study 1 with small stake sizes and Study 2 with large stake sizes) included 11 modified dictator games, in which participants allocated a fixed number of tokens between themselves and their recipient. In eight of the 11 games, each token's value was different for each player. In Study 1 (N = 78), 85% of participants did not violate the generalized axiom of revealed preference (GARP) throughout the 11 games. In Study 2 (N = 58), 81% of participants did not violate GARP. These results suggest that participants' decisions were largely consistent with utility maximization. Following Andreoni and Miller's analysis, we classified all participants (except one anomalous case) into the Selfish, Leontief (egalitarian), and Perfect Substitutes (utilitarian) groups. The majority of participants were classified into either the Leontief or Prefect Substitutes groups (i.e., non-selfish groups).

Keywords

dictator game, generalized axiom of revealed preference (GARP), utility maximization

Introduction

Human societies are characterized by large-scale cooperation among genetically unrelated individuals. An evolutionary explanation of this uniquely human sociality is strong reciprocity, which consists of a predisposition to cooperate with others (i.e., other-regarding preference) and a predisposition to punish non-cooperators (Gintis, 2000). Although the first component (i.e., the other-regarding preference) contradicts a standard assumption of economics (i.e., payoff maximization), it has been empirically supported by a large number of economic game experiments (for reviews, see Camerer, 2003; Gintis et al., 2003). For example, in the dictator game, in which one player (allocator) decides how to distribute a sum

of money between them and another player (recipient), participants in the allocator role tend to give the recipient non-zero amounts, instead of keeping everything for themselves (e.g., Forsythe et al., 1994). However, this finding may be accounted for by confusion: Participants did not understand the rules of the game (Delton et al., 2011; Gintis, 2000; Gintis et al., 2003; Hagen & Hammerstein, 2006). Gintis (2011) counteracted this explanation by referring to Andreoni and Miller's (2002) study that showed that participants' generous allocations in a modified version of dictator games were far from random. In particular, participants' series of allocation decisions did not violate the generalized axiom of revealed preference (GARP), which suggests that participants might have maximized some other-regarding preferences.

To explain the logic behind Andreoni and Miller's (2002) study, we briefly describe revealed preference theory, an economic theory originated by Samuelson (1938, 1948). The core idea of this theory is that although individuals' preferences for various goods are not directly observable, they can be inferred by observing choices that individuals have made. For example, suppose that you have 1,000 Japanese yen (JPY), which is your budget constraint, and purchase apples and oranges. The price of one apple and one orange is 200 JPY and 100 JPY, respectively; following Varian's (1982) notation, these prices are denoted as p^i = (200, 100). Suppose that you purchased four apples and two oranges; these quantities are denoted as x^i = (4, 2). This purchase behavior reveals part of your preferences: you prefer this bundle of goods (x^i) to any other combination within your budget constraint of 1,000 JPY (henceforth, these less-preferred combinations are collectively denoted as x). Note that if $p^i x^i \geq p^i x$ holds, x is purchasable within the 1,000 JPY budget constraint. Therefore, $p^i x^i \geq p^i x$ implies that x^i (the chosen bundle of goods) is directly revealed preferred to other purchasable bundles, x. In the analogy of standard inequality, $p^i x^i > p^i x$ implies that x^i is strictly directly revealed preferred to x.

In the above example, x includes x^j = (2, 4) because $p^i x^i$ = 1,000 ≥ $p^i x^j$ = 800. More precisely, the choice of x^i implies that x^i is directly revealed preferred to x^j. In other words, if you have temporarily stable preferences for apples and oranges, you should not purchase x^j (the less-preferred option) when x^i (the preferred option) is purchasable. For example, if you purchase x^j when prices are p^j = (100, 200), this implies that x^j is directly revealed preferred to x^i (because $p^j x^j$ = 1,000 > $p^j x^i$ = 800). These two revealed preferences are mutually contradictory to each other (see Supplementary Material for graphical explanations): Intuitively, the choice of x^i with the price of p^i implies that you prefer "more apples" to "more oranges," while the choice of x^j with the price of p^j implies that you prefer "more oranges" to "more apples." Therefore, these choices are not expected if you make two choices based on a single utility function. In fact, these two choices violate the so-called weak axiom of revealed preference (WARP),

which is defined below (e.g., Varian, 1982).

Weak Axiom of Revealed Preference (WARP): If x^i is directly revealed preferred to x^j ($x^i \neq x^j$), then x^j is not directly revealed preferred to x^i.

The notion of "directly revealed preference" can be extended to "indirectly revealed preferences." Suppose that you have made a series of decisions: $x^i, x^j, x^k, \ldots x^m$. If $p^i x^i \geq p^i x^j, p^j x^j \geq p^j x^k \ldots p^m x^m \geq p^m x$ holds for the sequence of decisions, then x^i is indirectly revealed preferred to x. Based on the notion of indirectly revealed preference, more stringent axioms can be defined (see Andreoni & Miller, 2002; Varian, 1982).

Strong Axiom of Revealed Preference (SARP): If x^i is indirectly revealed preferred to x^j ($x^i \neq x^j$), then x^j is not directly revealed preferred to x^i.

Generalized Axiom of Revealed Preference (GARP): If x^i is indirectly revealed preferred to x^j, then x^j is not strictly directly revealed preferred to x^i.

Of the three axioms, GARP is "a necessary and sufficient condition for data to be consistent with utility maximization" (Varian, 1982, p. 948). Therefore, when we evaluate our experimental data, particular emphasis is placed on GARP violations.

Revealed preference theory can be applied to the allocation data in a modified version of the dictator game. In this version, the allocator decides how to distribute a certain number of tokens (m) between themselves and the recipient, and the monetary value of each token differs between the two players. For example, the allocator may be endowed with $m = 10$ tokens, each of which is worth 1 JPY for the allocator and 2 JPY for the recipient. If the allocator keeps five tokens and gives five tokens to the recipient, the allocator and the recipient will receive 5 JPY and 10 JPY, respectively. This situation is in fact analogous

to the situation in which one decides how to allocate their budget on apples and oranges. The allocator needs to decide how to allocate the budget (m) between the self (allocator) and the other person (recipient). In this example, each token allows the allocator to "purchase" 1 JPY for the self or 2 JPY for the other. More precisely, the payoff structure of this game can be summarized as $m = p_s \pi_s + p_o \pi_o$, where π_s and π_o (analogous to the quantities of two goods) represent the monetary rewards for the self and the other, respectively, and p_s and p_o represent the prices for the self and the other to "purchase" 1 JPY with the token, respectively (these prices are reciprocals of the monetary values of one token). In the above example, $m = 10$, $\pi_s = 5$, $\pi_o = 10$, $p_s = 1$, and $p_o = 1/2$. Owing to this mathematical equivalence, this modified dictator game allows the researchers to evaluate whether participants' allocation data is consistent with utility maximization using the aforementioned axioms. In other words, if participants make allocation decisions to maximize other-regarding utility (or any other types of utility), their decisions do not violate GARP. This can be considered as counterevidence to the explanation that participants who make altruistic choices are merely confused (Gintis, 2011).

Andreoni and Miller (2002) used 11 modified dictator games (although some of their participants engaged only in the first eight of the 11 games). Their incentive structures were comparable with the 11 games used in our study (see Table 1). Andreoni and Miller showed that only 10% of their participants (18 of 176) exhibited at least one violation of GARP. They then assessed the severity of these violations using Afriat's (1972) Critical Cost Efficiency Index (CCEI). This index (denoted as e) is used to relax the budget constraint by multiplying $p^i x^i$ by e ($0 \leq e \leq 1$)—more severe violations require smaller e values to be eliminated from the data. Following Varian's (1994) proposal, Andreoni and Miller used $e = .95$ as the threshold of severe violations and found that only three participants' violations were severe. This result suggests that participants' altruistic choices in the dictator game are

Table 1. Modified dictator game used in Study 1 (top panel) and Study 2 (bottom panel).

Game	Total Tokens Allocated	Hold Value for the Self	Pass Value for the Other	Relative Price of Giving	Average Tokens Passed (*SD*)
			Study 1		
1	150	1	2	0.5	53.3 (31.3)
2	80	1	3	0.33	24.4 (20.1)
3	150	2	1	2	72.6 (41.6)
4	120	1	2	0.5	43.1 (29.3)
5	80	3	1	3	42.6 (25.7)
6	120	1	1	1	51.0 (22.1)
7	200	1	1	1	84.7 (37.0)
8	120	2	1	2	58.7 (33.8)
9	160	1	1	1	68.3 (29.0)
10	80	4	1	4	42.5 (28.3)
11	80	1	4	0.25	25.2 (24.1)
			Study 2		
1	75	10	20	0.5	24.1 (20.1)
2	40	10	30	0.33	11.2 (10.8)
3	75	20	10	2	26.0 (21.7)
4	60	10	20	0.5	17.8 (13.7)
5	40	30	10	3	15.0 (13.0)
6	60	10	10	1	21.6 (12.8)
7	100	10	10	1	33.2 (21.7)
8	60	20	10	2	21.4 (17.3)
9	80	10	10	1	28.5 (16.5)
10	40	40	10	4	14.2 (13.9)
11	40	10	40	0.25	12.3 (12.3)

not due to confusion; rather, they make altruistic decisions to maximize some utility function. Andreoni and Miller then classified participants' utility function into three types: Selfish preferences of $U(\pi_s, \pi_o) = \pi_s$ (participants of this type are only concerned about their own monetary rewards); Leontief preferences of $U(\pi_s, \pi_o) = \min\{\pi_s, \pi_o\}$ (participants of this type care about the well-being of a relatively worse off player—either the self or the other—and thus endorse outcome equality); and Perfect Substitutes preferences of $U(\pi_s, \pi_o) = \pi_s + \pi_o$ (participants of this type are concerned about the sum of the two players' final rewards and can be considered utilitarian). In Andreoni and Miller's study, 47.2%, 30.4%, and 22.4% of participants' preferences were classified as Selfish, Leontief, and Perfect Substitutes, respectively. Although the consistency in the dictator game allocations was conceptually replicated by other researchers (e.g., Fisman et al., 2007), to our knowledge, it has not been tested in Japan. Therefore, the primary purpose of this study was to replicate Andreoni and Miller's study in Japan. Although we conducted two separate studies (Studies 1 and 2), we report them conjointly because they only differed in stake sizes.

Methods

Participants were 80 and 59 undergraduate students for Studies 1 and 2, respectively. After eliminating participants who did not follow the instructions, we included 78 (39 females; $M_{AGE} \pm SD = 20.12 \pm 1.77$ years) and 58 (28 females; $M_{AGE} \pm SD = 20.28 \pm 1.18$ years) participants, respectively, in the subsequent analyses. The experiment was conducted in a small group setting (each session included 6 to 16 participants) as a part of larger data collection. One of the experimenters explained the rules of the modified dictator game and told participants that they would be anonymously paired with one of the other participants in the same session. The experimenter also explained that they would receive monetary rewards based on the result of one randomly chosen game.

Participants then received the questionnaire listing the 11 games (the questionnaires used in Studies 1 and 2 are available at https://osf.io/4ykwx/) and made the 11 allocation decisions. For each game, participants decided how many tokens they would give to the self and how many to the other (note that in the Japanese questionnaire, we used "point" instead of "token"). As shown in Table 1, the stake sizes were substantially different between the two studies: the maximum amount of reward in Study 1 was 320 JPY (if the participant allocated all 80 tokens to the other in game 10 or to the self in game 11), while the maximum reward in Study 2 was 1,600 JPY, which was almost comparable with Andreoni and Miller's (2002) stake size (if we assume 1.00 USD = 100 JPY).

Results

Using the standard algorithm proposed by Varian (1982), we first counted the numbers of WARP, SARP, and GARP violations in each participant's allocation decisions (analytic codes in the R Markdown HTML format are available at https://osf.io/4ykwx/). Because we had 11 games, there were a maximum of 55 violations (= $_{11}C_2$) in each participant's data. The results are reported in Table 2 in the same format as Andreoni and Miller's (2002) Table 2. In Study 1, there were 12 participants (15%) whose

Table 2. Observed WARP, SARP, and GARP violations and CCEI to eliminate GARP violations.

| Participant ID | Number of Violations | | | CCEI |
	WARP	SARP	GARP	
		Study 1		
4	1	5	5	1*
11	2	2	2	1*
15	2	3	3	0.937
21	1	1	1	1*
22	2	9	7	1*
24	4	4	4	1*
49	3	11	9	0.937
56	1	1	1	1*
68	1	1	1	0.953
73	1	1	1	0.953
76	1	1	1	0.953
77	1	3	3	0.969
		Study 2		
12	2	10	7	0.833
16	3	4	4	1*
18	3	38	37	0.833
23	1	1	1	1*
33	2	4	3	1*
35	3	4	4	0.937
41	1	3	2	1*
47	1	1	1	1*
55	2	3	2	1*
57	6	20	20	0.833
58	2	5	5	0.937

Notes. The possible range of the number of violations is [0, 55]. "1*" in the CCEI column means that the participant's violations were eliminated by the minimal change in *e*. Shaded rows indicate severe GARP violations (associated with CCEI smaller than 0.95).

allocation decisions violated the three axioms at least once. However, only two participants' GARP violations were severe (i.e., the necessary CCEI to eliminate those violations was smaller than 0.95). In Study 2, there were 11 participants (19%) whose decisions violated the three axioms at least once; five of them committed severe violations. The frequency of GARP violations was slightly higher in Study 2 than in Study 1 and Andreoni and Miller's study. However, it was still much lower than the expected numbers of violations under the assumption of random choice. For example, Andreoni and Miller's simulation showed that a random population of 50,000 committed an average of 4.39 violations of WARP, 17.62 of SARP, and 17.28 of GARP. In Study 1, the average number of violations (± *SD*) were 0.26 ± 0.71 (WARP), 0.54 ± 1.79 (SARP), and 0.49 ± 1.53 (GARP), which were all significantly different from the above simulation results by one-sample *t*-tests: $t(77) = -51.38$ (WARP), -84.17 (SARP), and -97.14 (GARP), all *p*s < .001. In Study 2, the average numbers of violations were 0.45 ± 1.11 (WARP), 1.60 ± 5.76 (SARP), and 1.48 ± 5.57 (GARP), and were significantly different from the simulation results: $t(57) = -27.02$ (WARP), -21.20 (SARP), and -21.61 (GARP), all *p*s < .001. Therefore, the participants' allocation decisions were far different from random decisions, suggesting that they made their decisions based on certain preferences.

Following Andreoni and Miller (2002), we then classified our participants' preferences into three types: Selfish, Leontief (egalitarian), and Perfect Substitutes (utilitarian). We first counted the number of participants whose allocation decisions perfectly matched the prototypical allocations of each type (see Supplementary Material for the prototypical allocations for each preference type). The results are summarized in Table 3, which uses the same format as Andreoni and Miller's Table 3. In Study 1, the number of participants whose allocation decisions exactly matched prototypical utility functions (indicated as "strong fit" in Table 3) were 7 (Selfish), 21 (Leontief), and 4 (Perfect Substitutes), which comprise 41% of Study 1 participants (cf. 43% in Andreoni and Miller's study). We then calculated Euclidean distances between each participant's decisions to the three prototypical decisions. Comparing the three distance scores, we assigned each participant to the closest class (these participants are denoted as "weak fit" in Table 3). However, there was one participant who always gave all the tokens to the recipient and thus did not fit any of the three utility functions. After removing this anomalous participant and combining the remaining strong- and weak-fit participants, 15.6% of participants were classified as Selfish, 72.7% as Leontief,

and 11.7% as Perfect Substitutes. In Study 2, there were 6 (Selfish), 10 (Leontief), and 2 (Perfect Substitutes) strong-fit participants. Combining the strong- and weak-fit participants, 32.8% were classified as Selfish, 50.0% as Leontief, and 17.2% as Perfect Substitutes.

For an exploratory purpose, we tested the difference in the distributions of the three utility types between the two studies. An omnibus 3×2 χ^2-test revealed that they are significantly different, $\chi^2(df = 2) = 7.69$, $p = .021$. Using the chisq.posthoc.test package of R (Ebbert, 2019), we conducted post hoc tests with the Bonferroni adjustment. The results showed that the frequency of the Leontief type was significantly lower in Study 2 than in Study 1 ($p = .041$), while the frequency of the other two types was not significantly different between the two studies ($p = .113$ and 1.000 for Selfish and Perfect Substitutes, respectively).

In sum, despite the significant decrease in the Leontief type (i.e., egalitarian) in the high-stake study (Study 2), the majority of participants were assigned to either one of the two non-selfish groups (i.e., Leontief or Perfect Substitutes) in both studies. In other words, the majority of participants made a series of non-selfish allocations. More importantly, most of these non-selfish allocations did not violate GARP, which implies that they tried to maximize some other-regarding utility.

Discussion

We conducted two replications of Andreoni and Miller's (2002) modified dictator game experiment. In both studies, we confirmed that allocation decisions (i.e., altruistic decisions in an experimental game setting) of Japanese participants tended to not violate GARP, which closely replicated Andreoni and Miller's results. This is counter to the argument that participants behaved in an altruistic manner because of confusion. However, as Delton et al. (2011) argue, if participants misapply a decision rule that is adaptive in ecologically valid settings to economic games, the above pattern is still attributable to "confusion" of the situation. Moreover, some simple heuristics, such as the equality heuristic (Messick, 1993), also cause the observed pattern. Further studies are needed to conclude that participants do in fact maximize their utility.

A significant difference was observed in the frequency of egalitarian participants (i.e., Leontief type) between Study 1 (72.7%) and Study 2 (50.0%): Participants were less egalitarian when the stake size was larger (Study 2). It is also noteworthy that Japanese participants were more frequently classified as Leontief than a group of American economic students who partook in Andreoni and Miller's

Table 3. Participant classification by prototypical utility function.

Utility Function	Fit		Total
	Strong	Weak	
		Study 1	Anomalous preference excluded
Selfish	7	5	12 (15.4%) (15.6%)
Leontief	21	35	56 (71.8%) (72.7%)
Perfect Substitutes	4	5	9 (11.5%) (11.7%)
Anomaly (Altruistic)	1	0	1 (1.3%)
		Study 2	
Selfish	6	13	19 (32.8%)
Leontief	10	19	29 (50.0%)
Perfect Substitutes	2	8	10 (17.2%)

(2002) study (Leontief frequency = 30.4%). However, our studies were not designed to test these differences. In future research, these factors (stake size, culture, academic major) must be experimentally manipulated to determine which factor(s) are responsible for the observed differences.

Acknowledgement

Supplementary Material
Electronic supplementary material is available online. Data in the CSV format and analytic codes in the R Markdown HTML format are available in the Open Science Framework (OSF): https://osf.io/4ykwx/.

References

Afriat, S. N. (1972). Efficiency estimation of production functions. International Economic Review, 13, 568-598. https://dx.doi.org/10.2307/2525845

Andreoni, J. & Miller, J. (2002). Giving according to GARP: an experimental test of the consistency of preferences for altruism. Econometrica, 70, 737-753. https://doi.org/10.1111/1468-0262.00302

Camerer, C. F. (2003). Behavioral game theory: experiments in strategic interaction. Sage.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. Proceedings of the National Academy of Sciences of the United States of America, 108, 13335-13340. https://doi.org/10.1073/pnas.1102131108

Ebbert, D. (2019). A post hoc analysis for Pearson's Chi-squared test for count data. https://cran.r-project.org/web/packages/chisq.posthoc.test/index.html

Fisman, R., Kariv, S., & Markovits, D. (2007). Individual preferences for giving. American Economic Review, 97, 1858-1876. https://dx.doi.org/10.1257/aer.97.5.1858

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. Games and Economic Behavior, 6, 347-369. https://doi.org/10.1006/game.1994.1021

Gintis, H. (2000). Strong reciprocity and human sociality. Journal of Theoretical Biology, 206, 169-179. https://doi.org/10.1006/jtbi.2000.2111

Gintis, H. (2011). Gene-culture coevolution and the nature of human sociality. Proceedings of the Royal Society B: Biological Sciences, 366, 878-888. https://doi.org/10.1098/rstb.2010.0310

Gintis, H., Bowles, S., Boyd, R., & Fehr, E. (2003). Explaining altruistic behavior in humans. Evolution and Human Behavior, 24, 153-172. https://doi.org/10.1016/S1090-5138(02)00157-5

Hagen, E. H., & Hammerstein, P. (2006). Game theory and human evolution: a critique of some recent interpretations of experimental games. Theoretical Population Biology, 69, 339-348. https://doi.org/10.1016/j.tpb.2005.09.005

Messick, D. M. (1993). Equality as a decision heuristic. In B. A. Mellers & J. Baron (Eds.), Psychological perspectives on justice (pp. 11-31). Cambridge University Press. https://doi.org/10.1017/CBO9780511552069.003

Samuelson, P. A. (1938). A note on the pure theory of consumers' behaviour. Economica, 5, 61-71. https://dx.doi.org/10.2307/2548836

Samuelson, P. A. (1948). Consumption theory in terms of revealed preference. Economica, 15, 243-253. https://dx.doi.org/10.2307/2549561

Varian, H. R. (1982). The nonparametric approach to demand analysis. Econometrica, 50, 945-973. https://dx.doi.org/10.2307/1912771

Varian, H. R. (1994). Goodness-of-fit for revealed preference tests. University Library of Munich, Germany. https://ideas.repec.org/p/wpa/wuwpem/9401001.html