# Giving according to GARP: A replication exercise

Esteban Degetau

2023-12-05

## Introduction

*Giving according to GARP* is an experiment designed to test whether altruistic behavior is consistent with a utility maximization process in the limited context of experimental economics (Andreoni and Miller 2002). If altruistic behavior is indeed rationalizable, then we can say it is driven by incentives through standard economic models. In the original experiment, subjects were given tokens that could be kept or given to a randomly chosen and unknown fellow subject. Tokens were then transformed into monetary payoffs at varying relative prices of giving. The experiment found that subjects' choices were indeed consistent with a utility maximization process through the complance of revealed axiom preferences.

I replicate Andreoni and Miller (2002) experiment with a sample of 15 students from the class of Applied Microeconomics at graduate level at Instituto Tecnológico Autónomo de México (ITAM). The most important change to the experimental design is that tokens were transformed into additional grade points in the midterm exam for the class. This change was made to test whether the incentive to give is different when the payoff is not monetary, but rather a significant increase in the grade for the class. I find that subjects in my sample were strictly more altruistic than those in Andreoni and Miller (2002). Whether this is the case because I used additional midterm grade points to incentivize subjects' behavior or because of other idiosyncratic characteristics such as different culture is yet unclear.

The rest of this paper is organized as follows. Section 2 describes the template for analysis. Section 3 describes the experimental design. Section 4 checks whether subjects' choices are consistent with a utility maximization process. Section 5 classifies subjects' preferences over altruism. Section 6 estimates parameters for CES utility functions. Section 7 concludes.

## Template for analysis

Andreoni and Miller (2002) seek to test whether choices over altruism are consistent with rationality and thus with economic models. Let $\pi_i$ represent payoffs for person $i$. For simplicity, choices made by person $s$ only affect her payoff $\pi_s$ and that of one other person $o$, $\pi_o$. Either persons are absolutely selfish, and have preferences $U_s = \pi_s$ or they care about the welfare of $o$, through the more general utility function:

$$U_s = u_s(\pi_s, \pi_o) \tag{1}$$

Whether experimental subjects maximize such a selfless utility function is the main concern of Andreoni and Miller (2002) and of my replication study.

## Experimental design

Andreoni and Miller (2002) use a modified version of the Dictator Game in Forsythe et al. (1994). From a given endowment $m$, a dictator must choose tokens to keep to herself $\pi_s$ or to give to a randomly chosen and unknown fellow subject $\pi_o$. Where the original Dictator Game does not allow for for varying relative prices of giving and keeping tokens, Andreoni and Miller (2002) exploit such variations to test whether subjects' decisions are rationalizable. Thus, a dictator will face a budget constraint $p_s\pi_s + p_o\pi_o = m$.

Different from Andreoni and Miller (2002), the key feature in my replication study is that tokens are transformed into payoffs as additional grade points in a midterm exam for the class of Applied Microeconomics at graduate level. Indeed, I test the hypothesis of whether the allocation of additional grade points produce different preferences over altruism among students, vis-a-vis a significant monetary payoff as in the original study.

Note that, as is the case in Andreoni and Miller (2002), payoffs in my replication study translate into additional *consumption* in subjects' utility function, as subjects need not study as hard for the final exam if they manage to get substantive additional points. Indeed, the experiment was conducted during the last lecture of the semester, when students knew their midterm grade, but were yet to decide the amount of effort they would put into studying for the final exam. Hence, additional grade points in the exam can enter students' utility functions as chenges in consumption.[1]

Tokens were transformed into payoff points at the rate given by the prices $(p_s, p_o)$. For example, at prices $(3, 1)$, 40 tokens were transformed into 120 points if all were held by the dictator, and into only 40 points if all were given to other player. Payoff points were then tranformed

[1] Subjects' differences in initial miterm grade may cause them to behave differently in the experiment, and thus my estimates may not trully provide insights on individuals' absolute and abstract preferences over altruisim, but merely preferences over risk from their final grade.

Table 1: Allocation choices

| Round | Tokens | Hold value | Pass value | Relative price of giving | Av. Tokens passed (Andreoni) | Av. Tokens passed (Replication) |
|---|---|---|---|---|---|---|
| 1 | 40 | 3 | 1 | 3.0 | 8.0 | 8.4 |
| 2 | 40 | 1 | 3 | 0.3 | 12.8 | 17.5 |
| 3 | 60 | 2 | 1 | 2.0 | 12.7 | 15.5 |
| 4 | 60 | 1 | 2 | 0.5 | 19.4 | 25.7 |
| 5 | 75 | 2 | 1 | 2.0 | 15.5 | 20.5 |
| 6 | 75 | 1 | 2 | 0.5 | 22.7 | 37.7 |
| 7 | 60 | 1 | 1 | 1.0 | 14.6 | 22.3 |
| 8 | 100 | 1 | 1 | 1.0 | 23.0 | 35.7 |

into additional grade points at a rate of 0.01. So if kept, the 40 tokens increased the grade by 1.2 percentage points. If given, the 40 tokens increased the grade by 0.4 percentage points.

Table 1 shows the rounds played in my replication study. Eight rounds were played, with varying token endowments and varying relative price of giving. The table also shows some results of my game and the comparison to Andreoni and Miller (2002). On average, subjects' total payoffs amounted to 7.7 additional percentage points to their midterm grade. Note how subjects in my sample are consistently more altruistic across game profiles vis-a-vis the original sample.

Figure 1 shows the budget constraints that subjects faced on every round of the experiment. Although rounds are numbered sequentially, every subject saw them randomized differently, following Andreoni and Miller (2002). Note how the varying slopes drawn will provide variation to test whether subjects comply with revealed preference axioms.

Figure 2 clearly shows subjects' reaction to the incentive to give. As the relative price of giving increases, the share of tokens that subjects passed to their fellow classmates decreased, consistent with utility maximizing individuals. Note how subjects in my sample were more altruistic than those in Andreoni and Miller (2002) for low prices, but converge as the price of giving increases.

## Checking rationality

Subjects were indeed more altruistic when splitting additional grade points than they were when splitting monetary payoffs. But did subjects choose following a utility maximizing process? i.e. Are subjects' choices rationalizable? Table 2 used the computational methods developed by Surana (2021) to compute the number of violations of revealed preference axioms. I found that all subjects in my sample were strictly rational across the eight rounds of the game.

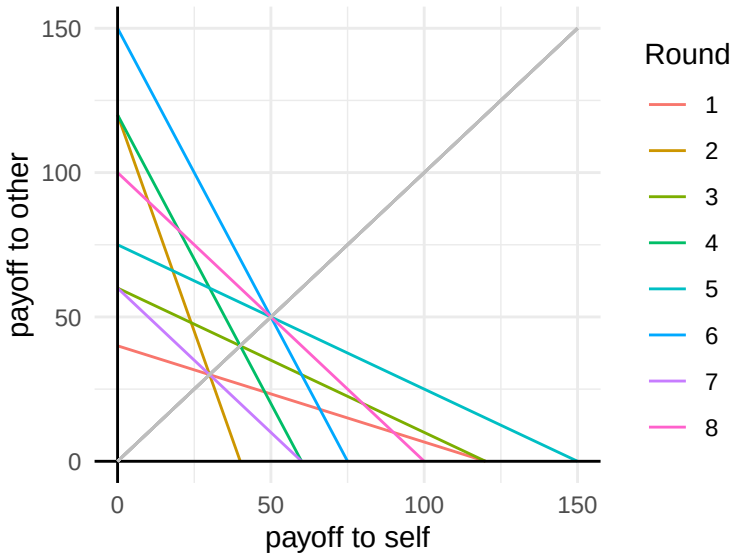


Figure 1: Budget constraints

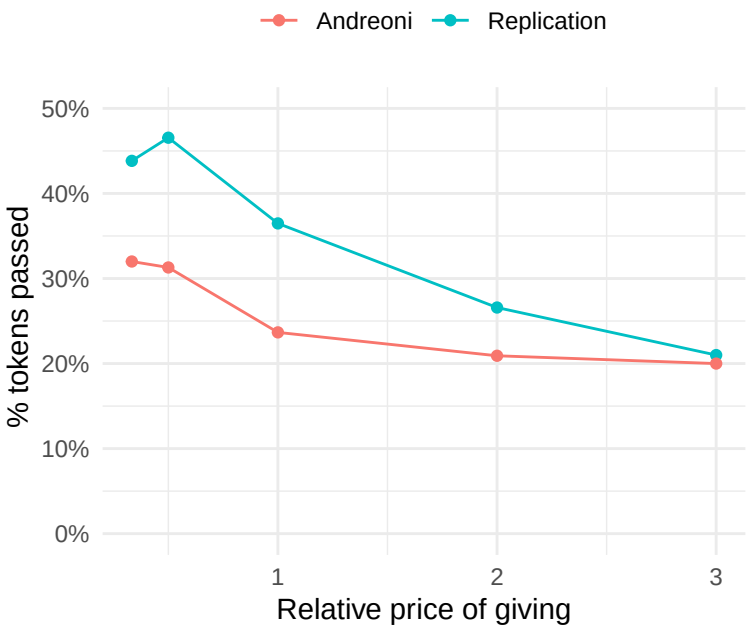


Figure 2: Tokens passed as a function of the relative price of giving

Table 2: Violations of Revealed Preferences

| Subject | WARP | SARP | GARP | CCEI |
|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 1 |
| 2 | 0 | 0 | 0 | 1 |
| 3 | 0 | 0 | 0 | 1 |
| 4 | 0 | 0 | 0 | 1 |
| 5 | 0 | 0 | 0 | 1 |
| 6 | 0 | 0 | 0 | 1 |
| 7 | 0 | 0 | 0 | 1 |
| 8 | 0 | 0 | 0 | 1 |
| 9 | 0 | 0 | 0 | 1 |
| 10 | 0 | 0 | 0 | 1 |
| 11 | 0 | 0 | 0 | 1 |
| 12 | 0 | 0 | 0 | 1 |
| 13 | 0 | 0 | 0 | 1 |
| 14 | 0 | 0 | 0 | 1 |
| 15 | 0 | 0 | 0 | 1 |

[a] Computed using Surana (2021).

These results can be verified by the reader in Figure 4, that shows the choices made by every subject.

Although it is surprising that none of my experimental subjects had a revealed preference axiom violation[2], the small sample I collected may not provide enough individual variation to accurately replicate this precise result.

## Individual preferences

Now that we know that subjects' choices are consistent with a utility maximization process, we can classify subjects' preferences over altruism. Andreoni and Miller (2002) classify their subjects into three prototypical utility functions: selfish, Leontief, and perfect substitutes. The selfish type is characterized by a utility function $U_s = \pi_s$, i.e. the dictator only cares about her own payoff. The Leontief type is characterized by a utility function $U_s = \min(\pi_s, \pi_o)$, i.e. the dictator cares about the payoff of the other player only as much as her own. This utility representation is consistent with a Rawlsian theory of justice. The perfect substitutes type is characterized by a utility function $U_s = \pi_s + \pi_o$, i.e. the dictator cares about the payoff of the

[2] Andreoni and Miller (2002) found 10% of subjects had an axiom violation, but only 2% had a severe enough violation (below the 0.95 CCEI threshold).

Table 3: Subject classification by prototypical utility function

| Utility function | Strict | Weak | Total | % |
|---|---|---|---|---|
| Selfish | 1 | 2 | 3 | 20 |
| Leontief | 0 | 6 | 6 | 40 |
| Perfect Substitutes | 1 | 5 | 6 | 40 |

other player as much as her own. This utility representation is consistent with a utilitarian theory of justice, and would maximize the points given to all players as a social planner.

Following Andreoni and Miller (2002), I tested whether subjects' choices were strictly consistent with one of the three prototypical utility functions. Table 3 shows the results from stric and weak classification exercises. It is surprising that most subjects did not exhibit strict preferences. This is different from the original paper. Where 43 percent of the subjects in the original sample were strictly matched to one of the three types; only 13 in my sample were.

Next, I replicated Andreoni and Miller (2002) weak classification into our three prototypical types. Although the authors do not provide the exact procedure to classify subjects into weak types, I followed the same intuition as in the strict classification. I computed the distance between the share of tokens passed by each subject in each round and the share of tokens passed by the three prototypical types. Then I classified subjects into the type that minimized the sum of distances across all rounds.

Table 3 shows the results from strict and weak classification exercises. It is surprising that most subjects did not exhibit strict preferences. This is different from the original paper, where 43 percent of the subjects were strictly matched to one of the three types. This is likely due to the small sample size of my replication exercise. Consistent with Andreoni and Miller (2002), most subjects were not classified as selfish. Figure 3 shows an example from each of the three weak types.

## Prediction

In this section, I seek to replicate the estimation of parameters for the CES utility function. Follwing Andreoni and Miller (2002), I find different parameters for each of the weak types found in Table 3. A CES utility maximizing agent solves the following problem:

$$\max_{\pi_s,\pi_o} \quad (a\pi_s^\rho + (1-a)\pi_o^\rho)^{\frac{1}{\rho}} \qquad \text{s.t. } p_s\pi_s + p_o\pi_o = m$$

where the share parameter $a$ denotes selfishness, $\rho$ captures the convexity of preferences through the elasticity of substitution $\sigma = 1/(\rho - 1)$. Note that budgets can also be written in terms of the price of the self-payoff, so $\pi_s + (p_o/p_s)\pi_o = m/p_s$ or simply $\pi_s + p\pi_o = m'$

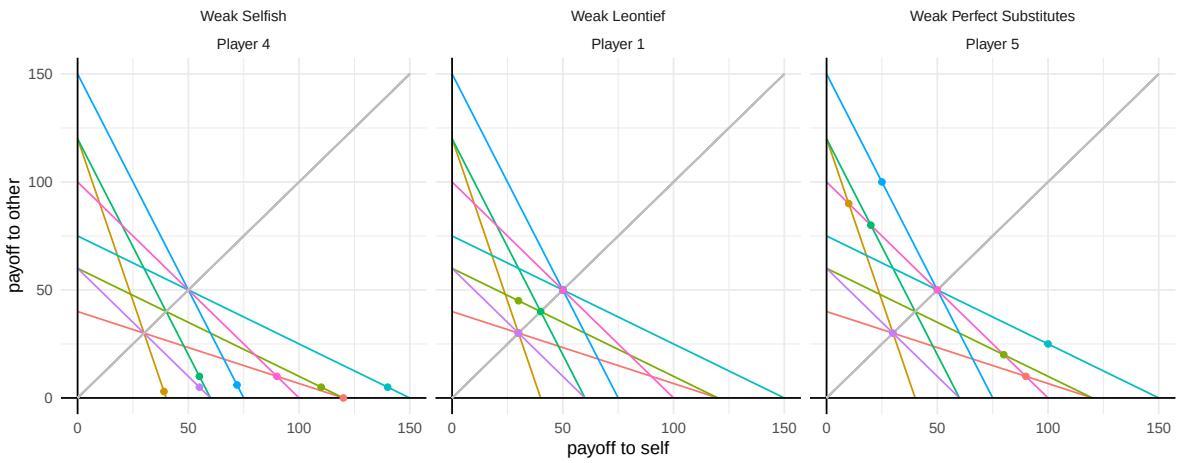


Figure 3: Choices from weak types

where $p = p_o/p_s$ and $m' = m/p_s$. The maximization problem yields the demand expression for self-payoff as a fraction of the budget:

$$\pi_s(p, m') = \frac{A}{p^r + A} m' \tag{2}$$

Where $A = \left(\frac{a}{1-a}\right)^{\frac{1}{1-\rho}}$ and $r = \frac{-\rho}{1-\rho}$.

Note that Andreoni and Miller (2002) are not explicit about the estimation strategy. They mention that they use a two-sided censored Tobit model, and that the dependent variable mutated as the share of tokens passed to avoid heteroskedacity. However, they do not provide the exact specification of the model. I derive the estimation model as follows:

Let $c = \frac{\pi_s}{m'}$ denote the share of the budget spent on the self-payoff. Then

$$\begin{aligned} c &= \frac{A}{p^r + A} \\ cp^r + cA &= A \\ cp^r &= A(1 - c) \end{aligned}$$

Taking logs on both sides yields

$$\log(c) = \log(A) - r\log(p) + \log(1 - c)$$

Table 4: Estimates of parameters for CES Utility functions for the three weak types

| | Weak Selfish | Weak Leontief | Weak Perfect Substitutes |
|---|---|---|---|
| $A = (\frac{a}{1-a})^{\frac{1}{1-\rho}}$ | 1.349 | 1.837 | 1.305 |
| | (0.095) | (0.013) | (0.015) |
| $r = \frac{-\rho}{1-\rho}$ | -0.657 | 0.032 | -0.272 |
| | (0.308) | (0.046) | (0.066) |
| $a$ | 0.545 | 0.652 | 0.552 |
| $\rho$ | 0.397 | -0.033 | 0.214 |
| $\sigma$ | -1.657 | -0.968 | -1.272 |
| Cases | 16 | 48 | 40 |
| Players | 2 | 6 | 5 |

[a] Delta method standard errors in parentheses.

Then we can write

$$\log(c_i) = \beta_0 + \beta_1 \log(p) + \epsilon_i \tag{3}$$

where $\beta_0 = \log(A)$ and $\beta_1 = -r$. Because $c$ takes values at 0, I approximate with $\sinh^{-1} c \approx \log c$. Then I estimate the Tobit model, censored at $0 \leq c \leq 1$.

Table 4 shows estimated values of the CES parameters for each of the three weak types found in Table 3. Estimates for parameter $A$ are all statistically different from zero, but the estimates for $r$ are not. This is not surprising, since the replication excericise has a very small sample compared to the original paper.

Although my estimates for the slefishness parameter $a$ are really close to the ones reported in Andreoni and Miller (2002), my estimates did not find the highest $a$ for the weak selfish type. I did find weak Leontief types to have the lowest (in absolute terms) elasticity of substitution $\sigma$, consistent with the original paper.

## Conclusions

Although the sample I collected was very small, some conclusions can be drawn from my replication study. First, subjects in my sample were strictly more altrusitic than those in Andreoni and Miller (2002). Whether this is the case because I used additional midterm grade points to incentivize subjects' behavior or because of other idiosyncratic characteristics such as different culture is a very interesting question for further studies. Second, choices from all the subjects were consistent with a utility maximization process, and thus confirms Andreoni and Miller (2002) finding that altruistic behavior is indeed rationalizable. Third,

very few subjects (13%) exhibited a behavior strictly consistent with a prototypical utility function such as selfish, Leontief or perfect substitutes. Fourth, despite the small sample, parameter estimation of CES utility functions across the three profiles drew curves consistent with microeconomi theory.

## References


Aguilar Esteva, Arturo. n.d. "Notas de Microeconometría." https://bookdown.org/viclzrz/notasmicro/.

Andreoni, James, and John Miller. 2002. "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica* 70 (2): 737–53. https://www.jstor.org/stable/2692289.

Forsythe, Robert, Joel L. Horowitz, N. E. Savin, and Martin Sefton. 1994. "Fairness in Simple Bargaining Experiments." *Games and Economic Behavior* 6 (3): 347–69. https://doi.org/10.1006/game.1994.1021.

Henningsen, Arne. 2022. *censReg: Censored Regression (Tobit) Models*. https://cran.r-project.org/web/packages/censReg/index.html.

Rawls, John. 1999. *A Theory of Justice*. 2nd edition. Cambridge, Mass: Belknap Press: An Imprint of Harvard University Press.

Surana, Khushboo. 2021. *Revpref: Tools for Computational Revealed Preference Analysis*. https://cran.r-project.org/web/packages/revpref/index.html.

Varian, Hal R. 1991. *Goodness-of-Fit for Revealed Preference Tests*. Department of Economics, University of Michigan Ann Arbor.

———. 1992. *Microeconomic Analysis*. Vol. 3. Norton New York.

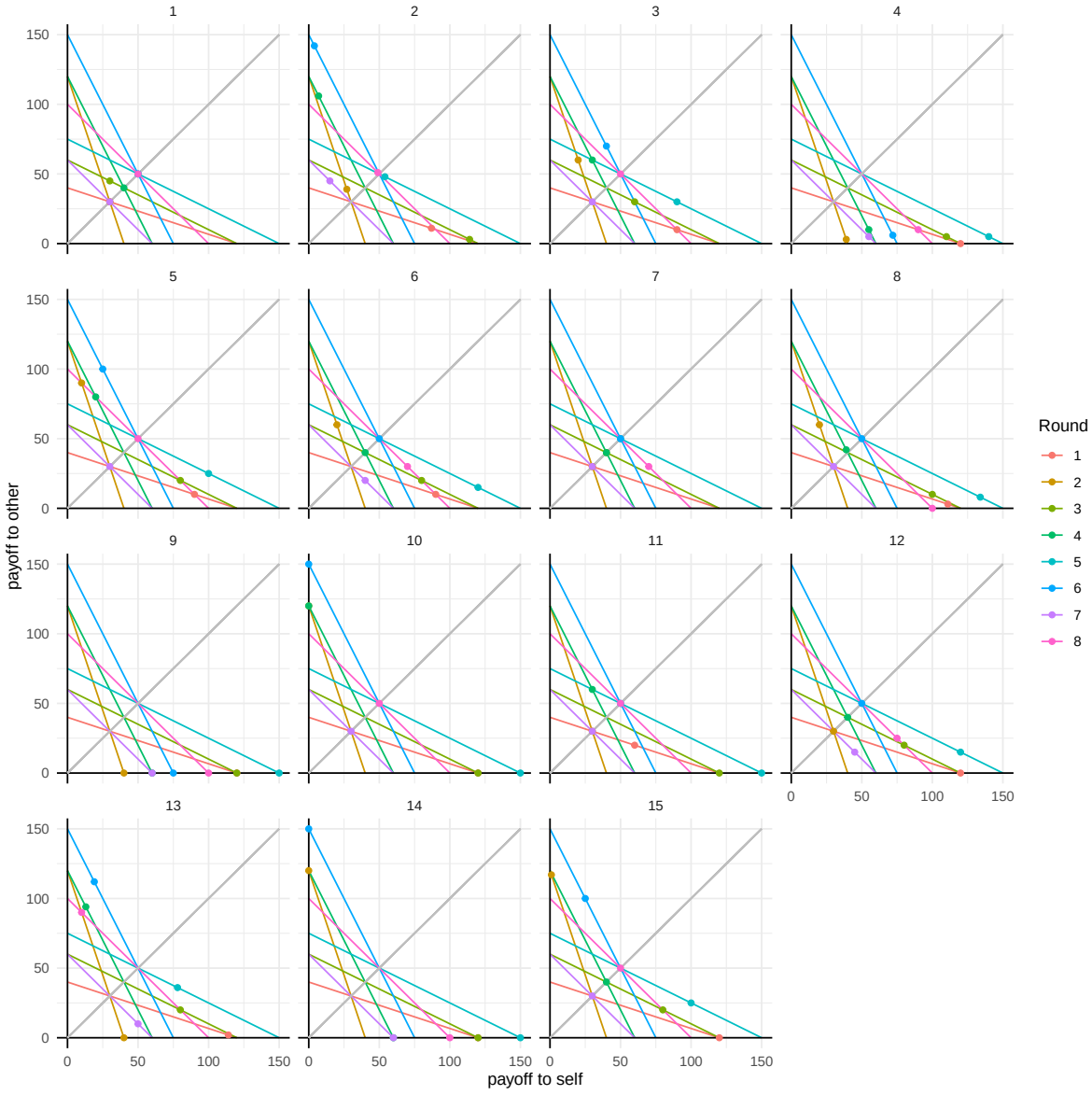


Figure 4: Choices from all the subjects