# DO WATER MANAGERS COOPERATE IN PUBLIC GOODS GAMES?

DAVID ZETLAND†

Abstract. Managers of public water companies present themselves and are seen as public servants maximizing public welfare. Because water is rarely allocated through market mechanisms, this maximization requires that managers cooperate in a bureaucratic version of a social dilemma. Members of the Metropolitan Water District of Southern California (MET, a consumer cooperative) face just such a dilemma: MET's member agencies make policies as members (setting prices, for example) that they obey as consumers.

This chapter reports the results of experiments that quantified cooperation among MET's member agency managers (MAMs) using public goods games. The results indicate that MAMs are neither relatively nor absolutely cooperative in comparison to, respectively, groups of students and a threshold efficiency consistent with maximizing social welfare. Additional results on type indicate that MAMs have a larger share of cooperators and free-riders than students, but MAMs are twice as likely to be free-riders as cooperators. MAM also appear to engage in cheap talk: Their responses to trust questions (stated preference) have no correlation with their experimental behavior (revealed preference); student preferences are correlated.

## Introduction and motivation

The Metropolitan Water District of Southern California (MET) is a public cooperative with 26 member agencies that resell MET water to over eighteen million people — making MET the largest water utility in the United States by population served and volume of treated water sold (Thomas, 2007). Most of MET's water comes from the Colorado River (via the Colorado River Aqueduct, or CRA) and Sacramento-San Joaquin Delta (via the California Aqueduct of the State Water Project). (Figure 1 shows the physical location of MET and these sources.) For most of its member agencies, MET is the sole supplier of imported water.

[Figure 1 about here]

*Date*: October 2, 2009. This chapter is derived from Chapter 5 of Zetland (2009).

† Wantrup Fellow, Department of Agricultural & Resource Economics, University of California, Berkeley, dzetland@gmail.com.

My interest in MET began with a story of rent-seeking inside the organization: In 1995, one of MET's member agencies — the San Diego County Water Authority (SDCWA) — tried to circumvent MET's self-declared monopoly on imported water by purchasing water from the Imperial Irrigation District (IID), an agency outside MET's service area that was already selling water to MET under a 1988 deal. Since SDCWA had no pipeline to IID, it wanted to use MET's CRA to **wheel** (move) the water from IID. SDCWA was prepared to pay MET the marginal cost of wheeling, but MET's other member agencies wanted SDCWA to pay the average cost of wheeling — meaning that SDCWA would subsidize *them*. MET's Board of Directors voted 25–1 to charge SDCWA the higher price, but San Diego (the nay vote) took MET to court. The dispute took several turns and only ended when the California Legislature allowed both to claim victory by contributing $235 million of taxpayer funds towards a settlement agreement (Erie, 2000).[1]

Was the IID-wheeling dispute a sign of doom or a mere transaction cost on the path to maximum efficiency? The answer to that question — 'Is MET an efficient organization?' — took an entire dissertation,[2] and this chapter formed part of that answer. This chapter examines the claim that MET can be efficient because the managers of its member agencies have social preferences that will allow them to 'cooperate their way' to efficiency.

The next section explains how the experimental quantification of cooperation fits into the larger argument that MET is inefficient and explains how the IID-wheeling dispute resulted from MET's cooperative structure. The sections thereafter cover the experiment, from hypotheses to results, with a discussion of their significance.

**Efficient cooperatives.** Because MET is a cooperative, it regulates itself in making decisions on pricing, hiring, capital expenditures, and so on.[3]

The problem with a collective structure — as Olson (1971) notes — is its one-size-fits-all nature. Polices that apply to all members are approved by a majority vote, and these votes will be more contentious when members' characteristics diverge. From this observation has emerged a theoretical consensus that cooperatives are more efficient than organizations with outside ownership (profit-maximizing firms) if and only if the members of the cooperative share a single goal or the same ordering of goals; that is, their preferences are 'reasonably' homogenous (Hart and Moore, 1996, 1998; Meade, 2005; Herbst and Prüfer, 2007). Based on this criterion alone, MET would be inefficient as a cooperative because member agencies differ in their dependency on MET: Some rely on MET for all their water, some for none.

But heterogeneous preferences do not end the story. Hart and Moore assume that the consumer/producer cooperative allocates a *scarce* good and that cooperative members are *self-interested*. If we relax their assumptions, homogenous preferences are still sufficient but no longer necessary, and two alternative sufficient conditions for efficiency emerge:

First, the cooperative may produce an **abundant** good for its customers — a good for which one member's consumption does not affect another's. Since the cooperative need make no policy on managing that good (it is not scarce), all members can consume according to their own preferences, and MET is efficient. MET was efficient for many years because it had too much, too cheap water that it managed as a **club good**; members who joined could use as much as they wanted.[4] Abundance allowed MET to grow for many years, and Southern California prospered. When abundance ended in the 1960s, costs went up and water grew scarce — becoming a **private good**.

Second, members of the cooperative may have **social preferences** such that they include other members in their utility functions. If members have social preferences, they will make cooperative policies that put more weight on group welfare — offsetting all or part of their underlying differences — and maximize group surplus. I test for social preferences by having MAMs play public goods games in an experimental lab — an **artefactual field experiment** (Harrison and List, 2004).

My results indicate that MAMs — the population of interest — are neither relatively nor absolutely cooperative in comparison to, respectively, groups of students and a threshold efficiency consistent with maximizing social welfare. (MAM cooperation is also similar to that of two comparison groups — MET executives (METs) and executives from investor-owned water companies (CWAs).) With neither abundance nor social preferences, the conditions for efficiency revert back to those of Hart and Moore. Because MET's member agencies do not possess the necessary homogenous preferences, its form as a cooperative is inefficient.

**Cooperation games.** How do we measure social preferences or cooperation? Stated preferences (for example: 'We are public servants who work together to serve the community.') may be cheap talk, but revealed preferences (lawsuits, contentious votes, and so on) may be normal friction attendant to a surplus-maximizing outcome. Experiments allow us to quantify cooperation as the fraction of maximum social welfare achieved. This quantification makes it possible to compare the cooperation levels of different groups.

I test cooperation using a **public goods game** (PGG) that emphasizes the tension between contributing to the public account (helping others at a cost to oneself) or a private account. In each run, each player (in a group of $n$) splits his endowment ($e$) between the group's public account and his private account. Each subject earns one token for each token in his private account plus the number of tokens equal to total contributions to the public account times the Marginal Contribution Ratio ($MCR \in (\frac{1}{n}, 1)$).[5] Given that all players benefit from contributions to the public account, the rational strategy is to contribute nothing; in the resulting Nash equilibrium, all subjects contribute nothing and receive $e$ (Ledyard, 1995). The social-welfare maximizing strategy is for all subjects to contribute their total endowment, so each receives $MCR(n * e) > e$. I define group **efficiency** as average subject earnings (from public and private accounts) divided by maximum possible earnings (Plott, 1982). Thus, a group of free-riders would have an efficiency of 50 percent ($e/MCR(n * e)$) and a group of cooperators would have an efficiency of 100 percent.

But how much cooperation is enough? Is cooperation relative to another group all we need to know or is some measure of absolute cooperation more important? And how does one define absolute? I determine if MAMs are absolutely cooperative using an out-of-sample method that uses experimental results from 'average people' (undergraduate students), who are classified as **cooperators** (contributing a lot to the public account), **free-riders** (contributing little or nothing), or **reciprocators** (contributing more when others do). With these typing results, I define the threshold of **absolute cooperation** as the minimum efficiency of *any* group of cooperators and free-riders.[6] I elaborate on this method under 'experimental design.'

**External validity.** The most common criticism of experiments is that they are not **externally valid**, which means that the subjects or design of the experiment resemble neither the population nor situation of interest outside the lab. Subject validity does not matter here, since the experimental subjects (water managers) are the target population, and students only serve as a comparison group. It is more important that the situation in the lab map to the situation outside the lab. Although it is hard to believe that people will change completely outside the lab, the framing and/or rendering of 'reality' in lab experiments can make results useless outside the lab.

A small literature links lab behavior to everyday life. In Mestelman and Feeny (1988), subjects who are common property advocates (mostly human ecologists and anthropologists) give more to the public account than students do. Cadsby and Maynes (1998) find — using a multiple equilibrium experimental setting — that nurses focus on the social-welfare maximizing equilibrium while students in economics or business focus on the individual-welfare maximizing equilibrium. Cooper et al. (1999) find that managers playing production games in the lab learn faster and are more strategic when context matches their everyday experience. Cardenas (2000) finds that experimental subjects with income from common-pool resources (wood-cutting or fishing) reach higher levels of cooperation faster than subjects with income from private goods (farming). In a famous paper on games that anthropologists conducted in many cultures, Henrich et al. (2001, pp. 76-77) conclude that 'the degree of cooperation, sharing, and punishment exhibited by experimental subjects closely corresponds to templates for these behaviors in the subjects' daily lives.' Alatas et al. (2006) find that Indonesian public servants are less corrupt than students in a lab experiment with context: Public servants — despite experiencing more corruption at work — offered fewer bribes. They passed on the opportunity to earn twenty times their hourly earnings 'to reduce corruption and social costs.' Students, on the other hand, 'made their decisions to maximize their payoffs' [pp. 17-18]. Palacios-Huerta and Volij (2008) report that soccer players transfer their skills from penalty kicks to zero-sum games in the lab, beating students on efficiency.[7] Herrmann et al. (2008) measure anti-social punishment in 16 countries, finding a negative relationship between *punishing* cooperation and GDP/capita.[8] They hypothesize that free-riders get revenge by punishing those who want cooperation; in the resulting fight, overall cooperation falls. Although their results are driven by culture, there is no reason why they could not apply to the culture of organizations or industries — explaining how internecine battles can weaken an organization.

An even smaller literature examines the impact of outside relationships on PGG results in the lab. Zelmer (2003), in a meta-analysis of PGGs, finds that previous friendship does not affect contributions. This result has exceptions — or perhaps clarifications: Peters et al. (2004) report that people playing with members of their family contribute more than when they are playing with strangers; Haan et al. (2006) report that teens contribute more in groups of friends than in groups of classmates.

In summary, significant evidence supports the relevance of outside conditions (occupation, experience, relationships) to lab results.[9] This result is important if we are to believe that water manager cooperation in the PGG reflects their cooperation in the office.

## Experimental design and hypotheses

This section describes the PGG design, method of typing subjects from their game decisions, calculation of efficiency, hypotheses tested, logistics of experimental sessions, and characteristics of participants.

In most PGGs, subjects play simultaneously, which creates some uncertainty over who or what one is reacting to and who or what others will react to in the next round. If, for example, two players in a group of four change their contribution in equal but opposite ways, the other players will not know if the first two changed their behavior or did nothing at all. Likewise, players changing their behavior may do so in the hope that others' actions will offset or reinforce their own.

Players will act differently in a sequential setting when they know their actions occur in isolation. From the experimenter's perspective, this sequential isolation is useful because it clarifies what the player saw and did — allowing the experimenter to type each player based on his behavior. The experiments here use a sequential contribution PGG inspired by the design of Kurzban and Houser (2005), or KH.[10]

The PGG described below uses the following terminology: A **round** is an individual decision; a **period** passes when all players in a group have had one round; a **run** lasts as long as the maximum number of rounds (the **round limit**), and every **session** had five runs. Simultaneous decisions made in period zero did not count toward the round limit, but runs would end in mid-period if the round-limit was not divisible by four. The round limits (from KH) were used in all sessions: The first run took 16 rounds (4 periods), the second was 7 rounds (2 periods/mid-period end), then 23 (6 periods/mid-period end), 32 and 32 rounds (8 periods/each).[11] The PGG had the following steps:[12]

(1) Subjects were randomly placed in groups of four or five at the beginning of each run.
(2) In period zero, all subjects had 20 seconds to make a simultaneous initial contribution to the public account from their 50-token endowments. Participants understood that their remaining tokens were provisionally (and in the final round, permanently) allocated to their private accounts.

Contributions were final only when the run ended. Subjects knew they had at least one opportunity to confirm/change their period zero contribution. Period zero contributions were non-binding cheap talk, but decisions after period zero were payoff relevant because the run could end at any point.

(3) After period zero, contributions were sequential. While the others waited, one participant per group saw his prior contribution, the average contribution of others, and the total in the public account (see Figure 2); he then changed or confirmed his contribution within the ten second duration of his round.

[Figure 2 about here]

(4) Each group's public account was updated, the next round began, and the next member of the group could change/confirm his contribution. Rounds and periods ran without signal or interruption.

(5) This updating continued for an unknown, random number of rounds until the run ended, contributions were finalized, and subjects saw their payoffs from public and private contributions.[13] Each subject received one token for each token in his private account and 0.5 token ($MCR = 0.5$) for each token in the group account.

(6) In each session, there were five runs of the game, each ending after KH's quasi-random number of rounds. Participants were randomly shuffled into groups and randomly ordered at the beginning of each run and played in the same order for that run. They knew they were in new groups, but they did not know the round limit or number of runs in the session.

**Determining individual type.** Each subject made one to eight decisions in each of the five runs — simultaneous, period zero contributions are ignored — with some players making more decisions than others.[14] Thus each player had 26–28 {own contribution, average contribution of others} observation pairs. To type players, each player's observations are used to estimate this equation:

$$x_{igt} = \alpha_i + \beta_i \bar{x}_{igt} + \epsilon_{igt}, \tag{1}$$

where $x_{igt}$ is the contribution of person $i$ of group (run) $g$ in round $t$, $\bar{x}_{igt}$ is the average contribution of other group members observed by $i$ in round $t$ of run $g$; $\alpha_i$ and $\beta_i$ are individual-specific parameters to be estimated, and $\epsilon_{igt}$ is a mean-zero disturbance term ($\sim N(0, \sigma_i^2)$) that controls for group effects ($g$) and trend effects ($t$).

Each individual's type depends on $\alpha_i$ and $\beta_i$ values estimated in an individual OLS regression of Equation (1).[15] Given $\hat{\alpha}$ and $\hat{\beta}$, KH's classification rules for type are as follows:

**Cooperators:** $\hat{\alpha} \geq 25$ and $\hat{\beta} \geq 0$: A cooperator's estimated contribution is non-decreasing in the average contribution of others and is always at least 25 (of 50) tokens.[16]

**Free-riders:** $\hat{\beta} \geq 0$ and $\hat{\alpha} + \hat{\beta}(50) < 25$: A free-rider's estimated contribution is non-decreasing in the average contribution of others but stays below 25 tokens.

**Reciprocators:** $\hat{\alpha} < 25$, $\hat{\beta} \geq 0$ and $\hat{\alpha} + \hat{\beta}(50) \geq 25$: A reciprocator's estimated contribution is increasing in the average contribution of others, below 25 tokens when the average contribution of others is zero, and at least 25 tokens when the average contribution of others is 50.

**No Type:** $\hat{\beta}_i < 0$. Players who give less when others give more are classified as 'no type' and ignored in further analysis.

Figure 3 shows datapoints for subjects who were classified as cooperators and free-riders; Figure 4 shows subjects classified as reciprocators. Each panel shows all data for one player in one session; the fitted line matches regression output and takes the form of $\hat{x}_i = \hat{\alpha}_i + \hat{\beta}_i \hat{\bar{x}}_i$. Note that each dot is a $(\bar{x}_{igt},\, x_{igt})$ pair that records the average contributions of others to the public account (independent variable on x-axis) and how much that subject puts in the public account (dependent variable on y-axis).

[Figure 3 about here]

[Figure 4 about here]

KH's method of classifying subjects with linear OLS point estimates is fast and easy to use (typing only requires a few observations), but some worry that it ignores potentially important factors. First, point estimates ignore the error structure of $\epsilon_{igt}$; second, type may not fit a linear profile; and third, Tobit is more appropriate for typing censored observations. After I report the results, I discuss how these effects do not have a significant impact on typing. With Occam's Razor in hand, I use KH's method instead of more complicated classification methods discussed in El-Gamal and Grether (1995); Houser et al. (2004); Houser and Winter (2004), for example.

**Determining efficiency.** Efficiency is quantified as the percentage of maximum social welfare achieved. In PGGs, both surplus and profit are maximized when players put all tokens into the group account. With an MCR of 0.5, a group of four and endowments of 50 tokens each, each player receives $0.5(4 * 50) = 100$ tokens — or double his endowment.[17]

Efficiency is 100 percent in this case and 50 percent if no players contribute to the public account. Efficiency for a session is the average efficiency for all runs, and run efficiency is the average efficiency for all groups.

Note the connection between types and efficiency: If cooperators dominate, group contributions and efficiency are higher; if free-riders dominate, they are lower. (Reciprocators react to others.) The order of play by types (cooperator in round one, free-rider in round two, for example) will have a minimal effect on efficiency if there are enough rounds or if randomization provides sufficient variation in type order.[18]

**Hypotheses.** The PGG allows us to test the following hypotheses:

- $\mathbf{H}_0^1$ **(Relative Cooperation):** MAMs and students are equally cooperative; this means that their efficiencies are not statistically different.
- $\mathbf{H}_0^2$ **(Absolute Cooperation):** MAMs achieve cooperation at a level greater than or equal to 80 percent.

The threshold for absolute cooperation is calculated from the reaction functions of average student cooperators and reciprocators. (These functions — $\hat{x}_c = 28.9 + 0.34\hat{\bar{x}}$ for cooperators and $\hat{x}_r = 6.1 + 0.79\hat{\bar{x}}$ for reciprocators — are depicted in Figure 6 on page 33.) If these representative students were put in a group of one cooperator and three reciprocators, they would achieve an efficiency of 87 percent.[19] Since it seems reasonable to say that such a group is 'cooperative,' I use this figure as a reference in setting the benchmark for **absolute cooperation** at 80 percent, using a lower number to be conservative. If any group — no matter its mix of types — should achieve the same or better cooperation, it also seems reasonable to call it 'absolutely cooperative.'

**Logistics.** Nine sessions with UC Davis students (UG1–UG9) established a baseline for comparison to three water manager sessions. All sessions took place in late 2006. Student sessions took place at a computer lab at UC Davis; water manager sessions took place at two locations.[20] Table 1 gives summary descriptive statistics for each session.

Each session began after subjects signed legal consent/disclosure forms and received their anonymous participant number. Participants heard directions (see Appendix A) and played five runs of the PGG. After each run, they were reshuffled into new groups of four or five participants. After playing, subjects answered questions that provided demographic information and values for the Trust Index (see below); Appendix B has the questionnaire. Finally, each player received an anonymous cash payment in proportion to his performance. The average payment to students was about $15; for water managers, the figure was tripled. Total session length was less than 90 minutes.

The **Trust Index** comes from the answers to four yes/no questions: 'People generally do the right thing;' 'I find it better to accept others for what they say and appear to be;' 'I am doubtful of others until I know they can be trusted;' and 'I almost always believe what people tell me.'[21] The yes answers to these questions are added (+1,+1,-1,+1) to get individual TI values $\in [-1, 3]$. If stated preferences match revealed preferences, the TI should be higher for cooperators and lower for free-riders.

[Table 1 about here]

## Results

This section has results from all experimental sessions except UG1, which was interrupted by a computer crash. Besides the main results on group efficiency and player type, there is an analysis of the correlation between TI values from the questionnaire and types from the experiments.[22]

Figure 5 shows mean efficiency (cooperation) for each subject pool.[23] The student average of 67 percent efficiency is indicated by a horizontal line with ±2.2 percent error bars marking the 95% confidence interval).[24] These results allow us to evaluate the null hypotheses:

[Figure 5 about here]

**$H_0^1$ (Relative Cooperation):** MAMs and students are equally cooperative; this means that their efficiencies are not statistically different. **Fail to Reject:** A two-sided t-test comparing MAMs to students fails to indicate any difference in efficiency (p-value 0.92).[25]

**$H_0^2$ (Absolute Cooperation):** MAMs achieve cooperation at a level greater than or equal to 80 percent. **Reject:** The 95 percent confidence interval for MAM efficiency ranges from 59 to 76 percent.

Thus we can conclude that MAMs are 'average' in their cooperation and that this cooperation is lower than we would expect if MAMs came from a population of reciprocators and cooperators. These results indicate that MAMs cannot solve collective-action problems in the lab. Applicability outside the lab is discussed below.

To further understand our efficiency results, it helps to look at the shares of cooperators, free-riders and reciprocators.[26] Table 2 shows the share of each type and efficiency for each group. Figure 6 shows the average estimated contribution profile for types among students.

[Table 2 about here]

[Figure 6 about here]

An alternative measure of 'propensity to cooperate' is the Trust Index from the questionnaire. Table 3 shows the relation between values for stated preference and values for revealed preference: For students, average TI scores for each type are significantly different at the 5 percent level; for managers, they overlap even at the 10 percent level. The result for students (cooperators are more trusting than free-riders) is consistent with the literature addressing that relationship (Gächter et al., 2004; Gächter and Thöni, 2005; Nowak and Sigmund, 2005).[27]

[Table 3 about here]

## Discussion

In this section, I discuss the method of typing and application of these results to everyday activities of water managers outside the lab. Table 4, which sets the stage for the discussion of typing methods, shows the share of types calculated by the OLS method described above ('all coefficients') and share of types resulting from other OLS methods are discussed next.

[Table 4 about here]

**Using point estimates to determine type.** OLS regressions give estimates of $\hat{\alpha}$ and $\hat{\beta}$, and some coefficients are statistically insignificant. KH use all point estimates in their typing, and that method is used here. Using all estimates means ignoring error structures that are typically addressed using robust standard errors or clustering effects, but this cost to accuracy is more than compensated by avoiding an even great problem — bias in typing.

Bias is introduced when insignificant estimates are discarded (set to zero) because zero values of $\hat{\beta}$ are associated with free-riders. Put differently, using only significant coefficients increases the number of subjects classified as free-riders. We can see this effect in Table 4, where the share of free-riders rises from 12 to 27 percent in the 'significant coefficients only' scenario.

**Estimating types with a quadratic form.** KH's typing method forces each player's actions to fit a linear form, and it is not hard to imagine that some players may play a different strategy such as increasing contributions up to a certain point and then decreasing them. Fischbacher et al. (2001) find that 14 percent of subjects have that hump-shaped contribution profile.

Allowing for quadratic variation in Equation (1) gives us:

$$x_{i,g,t} = \alpha_i + \beta_i \bar{x}_{i,g,t} + \gamma_i (\bar{x}_{i,g,t} - \bar{x}_i)^2 + \epsilon_{i,g,t}, \quad (2)$$

where $\bar{x}_i$ is the average contribution of others for all rounds and $\gamma_i$ is an additional parameter to be estimated. $\bar{x}_i$ is used in $(\bar{x}_{i,g,t} - \bar{x}_i)^2$ to increase variation in the quadratic relationship and reduce problems with collinearity.

The estimating of Equation 2 with data from 144 students results in the classification of six students as 'no type' (meaningful negative $\beta$ and $\gamma$ coefficients). Of the 138 remaining students, 14 percent (as in Fischbacher et al. (2001)) are humped types (positive $\beta$ and negative $\gamma$ coefficients); 4 percent are cooperators; 10 percent are free-riders; and 72 percent are reciprocators. Figure 7 shows the estimated contribution profile for average types when using a quadratic classification.

[Figure 7 about here]

How do we interpret these results? The main shift (compared to KH's linear method) is from reciprocators to hump-shaped. In Table 4, we see that the share of reciprocators in OLS falls from 84 percent to 72 percent in OLS-quadratic, which has 14 percent hump-shaped types. This result implies that some subjects typed as reciprocators using a linear approximation contribute fewer tokens when the average contribution to the public account is above 28–30 tokens.

I do not use quadratic estimates because classification of types is more arbitrary; without statistical significance as a filter, most subjects have *some* value for $\gamma$, which introduces confusion as to who is a free-rider, reciprocator or cooperator. Since quadratic estimates merely reclassify some reciprocators as hump-shaped types, I do not use them.

**Estimating types with a Tobit model.** Finally, there is the much larger issue of contributions that are censored at upper and lower boundaries in the estimation model. Since an OLS estimate of the relationship between censored values of $x_{i,g,t}$ and $\bar{x}_{i,g,t}$ will produce inconsistent estimates, a Tobit model would probably be more accurate. Tobit is not used because it would require new definitions of types, and types defined under such a scheme would not be compatible with types defined with KH's scheme. Since a Tobit typing scheme is beyond the scope of this work, it is left for future efforts.

**Applying these results to MET.** MAMs are neither relatively more cooperative than the benchmark group of students nor absolutely cooperative compared to an out-of-sample benchmark of cooperation among a mixed group of 'representative' cooperators and reciprocators.

Since this lab result contradicts claims that MAMs are 'cooperative enough' to make MET efficient, it seems reasonable to assume that MET's institutions and incentive structures should be designed to encourage cooperative behavior (or outcomes) among managers who place greater weight on self interest than group welfare.

On a more basic level are three questions: Can the actions of an individual in an experiment be used to 'type' that individual's organization? If we assume they do, then which individuals represent their member agency's type — managers or directors? Finally, If we measure actions in the lab, do these results apply outside the lab?

The first answer is no: An organization's 'behavior' is the outcome of a complex negotiation among all its members, and every member agency is subject to the opinion and action of many parties (see note 1). The basic conclusion is that no person is an accurate representative of his organization.

Taking this caveat as given, we come to the second best option: If we *had* to choose one person to represent each member agency, who should we choose? While directors are the formal representatives of member agencies on the Board, and their votes determine the policies that MET implements, I am confident that MAMs are more representative of member agencies and that interactions among MAMs more accurately capture the relations among member agencies and how cooperative MET is. First, directors no longer dominate MET or run it to maximize regional benefits — insiders, Hundley Jr. (1992), and O'Connor (1998b) all agree that the introduction of term limits in 1974 reduced directors' institutional knowledge and increased politicization of their appointments by member agencies. These effects reinforce each other: Many directors represent their agencies instead of the region and treat their position as an intermediate step on the way to higher office. Second, this effect increases the relative power of member agency staff — and the general manager (McDermott, 1998). Last, MAMs work with each other to design and implement policies. Thus, If we want to choose any group of people whose relations were going to affect cooperation between member agencies, it would be the MAMs.[28]

Finally, do these results apply outside the lab? Do they apply to everyday coordination, cooperation and conflict at MET? Yes they do. Consider the free-rider strategy in a PGG: Contribute nothing and benefit from the contributions of others. Perhaps the closest match to the PGG is the capital projects game: Frequent water buyers pay (via sales revenue) for storage projects that irregular water buyers use in drought conditions.[29] Because storage is a public good (all member agencies benefit from it), freeriding agencies get the benefit without the cost.[30]Water managers call this strategy **cost shifting**, and the battles over who pays and who benefits are prominent in MET's history. These quotations characterize a problem that has existed for over fifty years:

> Since all members [of MET] are not of equal size and since all do not use District water equally, there is a strong tendency for the cities with a strong demand for water, but with little assessed property value, to pursue policies in making water prices as low as possible and to let the bulk of the costs be carried by taxes.
>
> — Milliman (1956, p. 491)

> [Some] agencies appear to want MET to develop costly backup capacity — or insurance — for their local supply strategies, while seeking to shift the cost from these benefits on to Metropolitan and other agencies and consumers.
>
> — Blue Ribbon Task Force (1994, p. 23)

Although infrastructure spending and PGGs have similarities, their decision processes differ. Experiments use a non-cooperative structure (no communication or binding agreements), but member agencies talk and make agreements — actions that characterize a cooperative game. Are managers really playing a cooperative game? According to Ostrom et al. (1994), communication and contracts mean little without external enforcement or trust. MET is a self-regulated cooperative with rules and agreements that are enforced through internal administrative codes. As far as trust is concerned, both insiders and outsiders (PriceWaterhouseCoopers, 1998; O'Connor, 1998a,b; Erie, 2006) observe low trust among member agencies and their representatives. Although MAMs did not participate in explicit trust experiments, they had the opportunity to express their level of trust when answering the questions behind the Trust Index. Unfortunately, the absence of any correlation between what managers said and did (no difference across types in Table 3) leads one to conclude that their answers are not inconsistent with cheap talk.

Even if we assume the situation at MET is non-cooperative, perhaps MAMs can talk their way to a more efficient outcome?[31] Communication will increase cooperation. But how high will it get? Could managers achieve an absolute level of cooperation that is efficient enough? Perhaps, but results from this experiment should temper our optimism.

## Summary

This chapter describes a cooperation experiment in which member agency managers (MAMs) and several comparison groups (students, other managers) decided how much of a public good they wanted to produce. The results indicate that MAMs are neither relatively nor absolutely cooperative in comparison to, respectively, groups of students and a threshold efficiency consistent with maximizing social welfare. Although a significant share of water managers behaved in ways consistent with being cooperative types, twice as many behaved as free-riders. These results are subject to critiques, but even a skeptical reader would find them hard to dismiss ignore. Although MET is a cooperative, its executives and managers from its member agencies do not behave in ways consistent with cooperative attitudes and preferences. Generalizing this finding obliges us to consider the possibility that public employees may favor their private interests over public welfare. In this case, the institutions surrounding public organizations (monitoring, checks and balances, decentralization, competition, and so on) should be structured to handle selfish — not just social — preferences.

## Notes

[1]Organizations do not take actions or have preferences. People within organizations have preferences, and their preferences are reconciled in some decision model that often results in an action attributed to the organization that may be neither rational not consistent; see Allison (1969) and Fehr (2005) for example. For convenience, however, let us assume — following McFadden (1975) — that the people in member agencies act as if they are part of a rational, monolithic entity with a single set of preferences in a single decisionmaking unit. From this assumption comes statements such as 'LADWP decides' or 'water managers are cooperative and thus so are their agencies.'

[2]The answer is no. In Zetland (2009), I explain how institutions formed in MET's early years are more appropriate for managing abundant water. As scarcity has increased, MET has failed to update its rules for pricing, voting, cost allocation, and so on. As a result, MET's efficiency is lowered by conflict among member agencies, growing cross-subsidies, and misallocation of water, among other things.

[3]The State Water Resources Control Board regulates water quality. The California Public Utilities Commission regulates investor-owned private utilities.

[4]Alternatively, MET had excess supply; these words are interchangeable in MET's situation.

[5]The MCR describes the marginal benefit to *all* players from an individual's contribution of a unit to the public account.

[6]A player's type (or preference) is the same as her strategy. Thus, a free-rider pursues a strategy of defection or selfishness.

[7]Levitt et al. (2007) report that experienced poker players and US soccer placers do not play so efficiently. This example highlights another advantage of experiments — that they can be replicated.

[8]Their analysis also supports a positive correlation between cooperation and GDP per capita, but punishing cooperation — instead of free-riding — is a novel finding.

[9]In an admirable reversal of this idea, Cardenas (2004) returned to play games in communities in which he had conducted cooperation experiments. In new sessions, he observed

learning and diffusion — both experienced and amateur players played with the understanding that 'both trust and cooperation could be sustained and would be profitable' [p. 27].

[10]The main difference between this treatment and KH's treatment was that subjects did not see the average contribution of others in KH's treatment. This difference is important, because KH assumed that players 'knew' that average from the information they saw. Although irrelevant here (all subjects saw the average), the impact of displaying the average is explored in Zetland (2008).

[11]In mixed groups of four and five, four-player groups waited while five-player groups finished, but fifth player decisions did not count towards the round limit. (Because subjects were already waiting for others in their group, additional waiting did not affect the flow of the game.)

[12]See Appendix A for a copy of instructions. These experiments ran on z-Tree (Fischbacher, 2007).

[13]Average contributions were biased upwards by the limited number of rounds and $1/n$ probability of any given run ending after a player's contribution decision. Since this bias applied to both students and water managers, it is ignored.

[14]When the number of rounds divided by the number of players is not an integer, the number of observations is not equal; for example, seven rounds for four players means that three players had two rounds (updates), and one had a single update.

[15]Since KH's classification method uses point estimates for individuals' parameter values (that is, ignoring statistical significance; see the Discussion), the error term is ignored. This method therefore ignores both group effects and trend effects. The impact of the former is minimized by shuffling players between groups. Although I ignore trend effects, there is the possibility that learning across the five runs is relevant. When I retype student players using data from runs 2–5, 10/144 (7 percent) of players change from one type to another: The share of cooperators rises by 2 percent and the shares of free-riders and reciprocators fall by 1 percent/each. Since these changes are small (the share of reciprocators goes from 79 to 78 percent), I also ignore trend.

[16]KH used 25/50 tokens (50 percent) as a cut-off between free-riders who never give more than half their endowment (contributing less than 25 even when others average 50) and cooperators who always give more than half (contributing more than 25 even when others average zero). In ultimatum games (one player decides how to split an endowment and his partner decides to reject — leaving both with nothing — or accept the split), the modal offer is fifty/fifty; see Camerer and Thaler (1995).

[17]Players in group of five had the same MCR, which meant a player's maximum payoff was 125 tokens. Although this higher payoff increases the incentive to contribute to the public good, the calculation of efficiency takes that effect into account by lowering minimum efficiency (zero contribution to the public good) from 50 to 40 percent.

[18]Say we have three reciprocators and one free-rider. If all donate 50 in (non-binding, simultaneous) round zero, the total is 200. If the free-rider gives zero in round one, the total drops to 150. The next reciprocator will see an average of 33 — $(150 - 50)/3$ since the average excludes own contribution — and give 33 in round two; the next one will see an average of 28 and give 28. Total contributions to the public account will continue to deteriorate until the run ends or they hit zero. Now imagine that a reciprocator goes in round one and a free-rider in round two: The same deterioration will occur with a one-round lag. With enough rounds, therefore, final efficiency will be the same. Typing (based on reactions to others' contributions) is not affected by type ordering.

[19]A group of five with one cooperator and four reciprocators reaches 84 percent. A group of four cooperators and one reciprocator reaches 91 percent efficiency.

[20]The CWA session (executives, consultants and staff attending the annual meeting of the investor-owned utilities' California Water Association) took place in San Francisco; the MAM session (Member Agency Managers attending their monthly meeting) and MET session (MET executives and senior staff) took place at MET's headquarters in Los Angeles.

[21]Pre-testing included 20 questions from sources exploring Machiavellianism (Gunnthorsdottir et al., 2002), and four with the largest variance were retained.

[22]Although we have reason to believe the demographic characteristics affect cooperation in these games (see List (2004), for example), a multinomial logit examination of student characteristics on type — reported in Zetland (2008) — does not indicate that gender, age or major were significantly correlated with type. (Trust Index was significant; see Note 27.) The same regression fails to find any significant correlations between demographic characteristics and type for the 34 (of 42) water managers assigned a type. That result may be explained by data problems (few observations, missing variables) as much as the absence of a significant correlation. These problems cannot be addressed by a pooled-data regression because pooling is likely to suffer from additional problems with missing variables.

[23]Average efficiency is calculated using the efficiency from average group efficiency within each of the five runs in each session. MAM efficiency is thus calculated from 15 observations given by five runs of three groups; students have 175 observations because their sessions had four to five groups in each run.

[24]The ordering of the results for different groups of water managers matches our intuition of cooperation rising with familiarity. MAMs have explicitly cooperative relations; METs work together but may be competitive; and CWAs know each other professionally. That said, the differences between these single observations of the target populations are not statistically-robust.

[25]The same test also fails to reject different means for CWAs (p-value 0.39) and METs (p-value 0.35) versus the student mean.

[26]136 of 144 students and 34 of 42 water managers are classified within these three types.

[27]A multinomial logit examination (not reported here) of the relationship between type (dependent variable) and individual characteristics shows that cooperators have significantly greater TI values than reciprocators. Although the TI coefficient for free-riders is negative, it is not statistically significant (p-value 0.18).

[28]These concerns do not mean that it would not be a good idea to run cooperation experiments with directors, only that the results of those sessions would have been less important than the MAM results reported here.

[29]This case occurs in 1987–1991 drought: In 1990, LADWP takes over six-times its 1986 delivery while SDCWA — closer to its base allocation — takes only 27 percent more water.

[30]The *existence* of storage is really a club good, it is a public good for member agencies in the MET 'club.'

[31]Ostrom et al. (1994) find that communication — without binding agreements — in PGGs increases efficiency by 42–80 percent. Other mechanisms (punishment, incremental commitments or voting) can also increase efficiency; see examples in Fehr and Schmidt (1999), Kurzban et al. (2001) and Kroll et al. (2007).

## References


Alatas, V., Cameron, L., Chaudhuri, A., Erkal, N., and Gangadharan, L. (2006). Subject Pool Effects in a Corruption Experiment: A Comparison of Indonesian Public Servants and Indonesian Students. University of Melborne Research Paper 975.

Allison, G. T. (1969). Conceptual Models and the Cuban Missile Crisis. *American Political Science Review*, 63:689–718.

Blue Ribbon Task Force (1994). Final Report, Metropolitan Water District of Southern California.

Cadsby, C. B. and Maynes, E. (1998). Choosing Between a Socially Efficient and Free-Riding Equilibrium: Nurses versus Economics and Business Students. *Journal of Economic Behavior & Organization*, 37(2):183–192.

Camerer, C. and Thaler, R. H. (1995). Anomalies: Ultimatums, Dictators and Manners. *Journal of Economic Perspectives*, 9(2):209–219.

Cardenas, J.-C. (2000). How Do Groups Solve Local Commons Dilemmas? Lessons from Experimental Economics in the Field. *Environment, Development and Sustainability*, 2:303–322.

Cardenas, J.-C. (2004). Bringing the Lab to the Field: More than Changing Subjects. *Working Paper*. www.aeaweb.org/annual_mtg_papers/2005/0107_0800_0304.pdf.

Cooper, D. J., Kagel, J. H., Lo, W., and Gu, Q. L. (1999). Gaming against Managers in Incentive Systems: Experimental Results with Chinese Students and Chinese Managers. *American Economic Review*, 89(4):781–804.

El-Gamal, M. and Grether, D. M. (1995). Are People Bayesian? Uncovering Behavioral Strategies. *Journal of the American Statistical Association*, 90(432):1137–1145.

Erie, S. P. (2000). Mulholland's Gifts: Further Reflections upon Southern California Water Subsidies and Growth. *California Western Law Review*, 37(1):147–60.

Erie, S. P. (2006). *Beyond Chinatown: The Metropolitan Water District, Growth, and the Environment in Southern California*. Stanford University Press, Stanford, CA.

Fehr, E. and Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, 114(3):817–868.

Fehr, Ernst; Tyran, J.-R. (2005). Individual Irrationality and Aggregate Outcomes. *Journal of Economic Perspectives*, 19(4):43–66.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics*, 10(2):171–178.

Fischbacher, U., Gachter, S., and Fehr, E. (2001). Are People Conditionally Cooperative? Evidence from a Public Goods Experiment. *Economics Letters*, 71(3):397–404.

Gächter, S., Herrmann, B., and Thonï, C. (2004). Trust, Voluntary Cooperation, and Socio-Economic Background: Survey and Experimental Evidence. *Journal of Economic Behavior & Organization*, 55(4):505–531.

Gächter, S. and Thöni, C. (2005). Social Learning and Voluntary Cooperation among Like-Minded People. *Journal of the European Economic Association*, 3:303–14.

Gunnthorsdottir, A., McCabe, K., and Smith, V. (2002). Using the Machiavellianism Instrument to Predict Trustworthiness in a Bargaining Game. *Journal of Economic Psychology*, 23(1):49–66.

Haan, M., Kooreman, P., and Riemersma, T. (2006). Friendship in a Public Good Experiment. Working Paper. http://www.eco.rug.nl/~haanma/drop.pdf.

Harrison, G. W. and List, J. A. (2004). Field Experiments. *Journal of Economic Literature*, 42.

Hart, O. and Moore, J. (1996). The Governance of Exchanges: Members' Cooperatives versus Outside Ownership. *Oxford Review of Economic Policy*, 12(4):53–69.

Hart, O. and Moore, J. (1998). Cooperatives vs. Outside Ownership. NBER Working Paper 6421.

Henrich, J. et al. (2001). In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *American Economic Review*, 91(2):73–78.

Herbst, P. and Prüfer, J. (2007). Firms, Nonprofits, and Cooperatives: A Theory of Organizational Choice. CentER Discussion Paper 2007-7.

Herrmann, B., Thoni, C., and Gachter, S. (2008). Antisocial Punishment Across Societies. *Science*, 319(5868):1362–1367.

Houser, D., Keane, M., and McCabe, K. (2004). Behavior in a Dynamic Decision Problem: An Analysis of Experimental Evidence Using a Bayesian Type Classification Algorithm. *Econometrica*, 72(3):781–822.

Houser, D. and Winter, J. (2004). How Do Behavioral Assumptions Affect Structural Inference? Evidence from a Laboratory Experiment. *Journal of Business & Economic Statistics*, 22(1):64–79.

Hundley Jr., N. (1992). *The Great Thirst: Californians and Water, 1770s–1990s*. University of California Press, Berkeley.

Kroll, S., Cherry, T. L., and Shogren, J. F. (2007). Voting, Punishment, and Public Goods. *Economic Inquiry*, 45(3):557–570.

Kurzban, R. and Houser, D. (2005). An Experimental Investigation of Cooperative Types in Human Groups: A Complement to Evolutionary Theory and Simulations. *Proceedings of the National Academy of Sciences*, 102(5):1803–1807.

Kurzban, R., McCabe, K., Smith, V., and Wilson, B. (2001). Incremental Commitment and Reciprocity in a Real Time Public Goods Game. *Personality and Social Psychology Bulletin*, 27:1662–1673.

Ledyard, J. O. (1995). Public Goods: A Survey of Experimental Research. In Kagel, J. H. and Roth, A. E., editors, *Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ.

Levitt, S. D., List, J. A., and Reiley, D. H. (2007). What Happens in the Field Stays in the Field: Professionals Do Not Play Minimax in Laboratory Experiments. Working Paper. `http://www.u.arizona.edu/~dreiley/papers/ProfessionalsMinimax.pdf`.

List, J. (2004). Young, Selfish and Male: Field evidence of social preferences. *Economic Journal*, 114(492):121–149.

McDermott, T. (1998). Knee-Deep disputes for "water buffaloes;" the power of the Metropolitan Water District long went unquestioned. Now it is mired in inertia, ineptitude. *Los Angeles Times*, 1 Nov.

McFadden, D. (1975). The Revealed Preferences of a Government Bureaucracy: Theory. *Bell Journal of Economics*, 6(2):401–416.

Meade, R. (2005). Ownership vs. Regulation in Electricity Reform: The Role of Governance. SSRN Working Paper 807144.

Mestelman, S. and Feeny, D. (1988). Does Ideology Matter?: Anecdotal Experimental Evidence on the Voluntary Provision of Public Goods. *Public Choice*, 57:281–286.

Milliman, J. W. (1956). *The History, Organization and Economic Problems of the Metropolitan Water District of Southern California*. PhD Dissertation, University of California, Los Angeles (Economics).

Nowak, M. A. and Sigmund, K. (2005). Evolution of Indirect Reciprocity. *Nature*, 437(7063):1291–1298.

O'Connor, D. E. (1998a). Governance of the Metropolitan Water District of Southern California: An Overview of the Issues. Legislative Analysis CRB-98-013, California State Library.

O'Connor, D. E. (1998b). Governance of the Metropolitan Water District of Southern California: Options for Change. Legislative Analysis CRB-98-018, California State Library.

Olson, M. (1971). *The Logic of Collective Action*. Harvard University Press, Cambridge, MA.

Ostrom, E., Gardner, R., and Walker, J. (1994). *Rules, Games, and Common-Pool Resources*. Ann Arbor Books, Ann Arbor, MI.

Palacios-Huerta, I. and Volij, O. (2008). Experientia Docet: Professionals Play Minimax in Laboratory Experiments. *Econometrica*, 76(1):71–115.

Peters, H. E., Unur, A. S., Clark, J., and Schulze, W. D. (2004). Free-Riding and the Provision of Public Goods in the Family: A Laboratory Experiment. *International Economic Review*, 45(1):283–299.

Plott, C. R. (1982). Industrial Organization Theory and Experimental Economics. *Journal of Economic Literature*, 20(4):1485–1527.

PriceWaterhouseCoopers (1998). Situational Analysis of the Metropolitan Water District of Southern California, Metropolitan Water District of Southern California. Presented to the MWD Board of Directors 17 Sep.

Thomas, B. G. (2007). Personal Communication. 4 Sep.

Zelmer, J. (2003). Linear Public Goods Experiments: A Meta-Analysis. *Experimental Economics*, 6(3):299–310.

Zetland, D. (2008). Focal Points in Public Goods Games: Explicit Information Increases Reciprocation. SSRN Working Paper 1122144.

Zetland, D. (2009). *Conflict and Cooperation Within an Organization: A Case Study of the Metropolitan Water District of Southern California*. VDM Verlag, Saarbrueken.

## Appendix A. Instructions

This is a game of group and individual investment behavior.

- You are in a group of 4 with 3 others, chosen at random. (If you are in a GROUP OF FIVE, you will find out during the game.)
- You have an endowment of 50 tokens to invest. Others have the same endowment.
- You invest your tokens in the Individual Exchange and the Group Exchange.
- Your earnings depend on how you and your group invest tokens.
- 50 tokens = $0.75 `[for students; $3.00 for Water managers]`

Every token you invest in the **Individual Exchange** returns one token in earnings **to you** only.

Every token you invest in the **Group Exchange** returns 0.5 tokens in earnings **to every member of your group, including yourself**. *It does not matter* who invests in the Group Exchange — everyone gets a return from every token invested in the Group Exchange, whether or not they invested.

*Your task is to maximize your earnings* by choosing how many of your tokens to invest in the Group Exchange. (Remaining tokens go to the Individual Exchange.) For example:

| | 1 | 2 | 3 |
|---|---|---|---|
| Your Group Exchange investment | 0 | 50 | 30 |
| Your Individual Exchange investment | **50** | **0** | **20** |
| *If others' Group Exchange investments total* | 90 | 110 | 0 |
| *. . . total Group Exchange investment is . . .* | $0 + 90 = 90$ | $50 + 110 = 160$ | $30 + 0 = 30$ |
| . . . and everyone's Group Exchange return is | $90/2 =$ **45** | $160/2 =$ **80** | $30/2 =$ **15** |
| Your total earnings (in tokens) are | $50 + 45 = 95$ | $0 + 80 = 80$ | $20 + 15 = 35$ |

**Game timing.**

(1) All members of your group start with a simultaneous investment in the Group Exchange (**Round 1**). Click 'Continue' after you enter your choice. *You only have 20 seconds to click.* A countdown clock is in the top-right corner of your screen.

(2) In **Round 2 and thereafter**, you will (one person at a time) see the number of people in your group (either 4 or 5), the TOTAL investment in the Group Exchange and the average investment of others in your group. You will change or confirm your Group Exchange investment and click 'Continue.' *You only have 10 seconds to click.* If you take too long, your choice does not change.

(3) The opportunity to see the total and change/confirm passes from person to person in your group for an unknown, random number of rounds until the run ends, and all investments are final. You will have *at least* one opportunity to change/confirm your investment. Although you must wait while the decision passes around your group, try to pay attention so as to not to miss your turn.

(4) When each run ends, you will see your investment, the total investment in the Group Exchange, your earnings from the current run, and your cumulative earnings.

(5) When the game repeats, players are randomly reshuffled into new groups and the final round changes to a new, random number.

## Appendix B. Questionnaire

(1) Age: .......................................................... ________ years
(2) Gender: ................................................ [MALE] [FEMALE]
(3) Your educational or professional field (circle one)
  (a) Anthropology
  (b) Economics
  (c) Engineering
  (d) Law
  (e) MBA/finance
  (f) Political Science
  (g) Sociology
  (h) Other Liberal Arts (English, Communications. . . )
  (i) Other Science (Biology, Chemistry, Math. . . )
(4) Number of years in this field ................................. ________ years
(5) Number of people in your household (including you) .................. ________
(6) Have you used an internet auction (e.g., eBay) to buy /sell something? ..... [Y] [N]
(7) Have you participated in an experiment similar to this one? ................ [Y] [N]
(8) Number of people here who. . .
  (a) . . . you know and work with? ...................................... ________
  (b) . . . you know but do NOT work with? ............................ ________
  (c) . . . you do NOT know but DO work with? ........................ ________
  (d) . . . you do NOT know and do NOT work with? .................. ________
  (e) The total (group minus 1 for you) should be .............................. 19.
(9) Some questions to answer using your own interpretation:
  (a) People generally do the right thing ................................... [Y] [N]
  (b) I find it better to accept others for what they say and appear to be .... [Y] [N]
  (c) I am doubtful of others until I know they can be trusted ............... [Y] [N]
  (d) I almost always believe what people tell me ........................... [Y] [N]

**Table 1:** Subjects' descriptive statistics. 'Experience' refers to the share of participants with experience using eBay or participating in economic experiments.

| Session Name | Date | Num. Subj. | Median Age | Share Male | Education (%) Engnr | Econ | Experience (%) eBay | Expmt | Median TI |
|---|---|---|---|---|---|---|---|---|---|
| UG1 | 19 Oct | 18 | 20 | 56 | 17 | 39 | 50 | 17 | 0.67 |
| UG2 | 23 Oct | 20 | 20 | 50 | 0 | 35 | 55 | 10 | 0.90 |
| UG3 | 24 Oct | 16 | 20 | 56 | 0 | 81 | 63 | 25 | 0.69 |
| UG4 | 26 Oct am | 13 | 21 | 69 | 0 | 62 | 69 | 31 | 0.54 |
| UG5 | 26 Oct pm | 20 | 21 | 45 | 10 | 65 | 55 | 20 | 0.80 |
| UG6 | 31 Oct am | 20 | 21 | 60 | 5 | 35 | 75 | 15 | 0.65 |
| UG7 | 31 Oct pm | 16 | 21 | 44 | 13 | 31 | 75 | 25 | 1.31 |
| UG8 | 2 Nov | 19 | 20 | 58 | 11 | 63 | 68 | 32 | 0.84 |
| UG9 | 7 Nov | 20 | 21 | 60 | 20 | 45 | 60 | 15 | 0.60 |
| CWAs | 16 Nov | 13 | 55 | 85 | 31 | 46 | 38 | 0 | 1.46 |
| MAMs | 17 Nov am | 15 | 49 | 87 | 73 | 7 | 33 | 0 | 2.47 |
| METs | 17 Nov pm | 14 | 51 | 54 | 14 | 14 | 21 | 0 | 2.07 |

**Table 2:** Subject types and group efficiency

| | Shares of Type (%) Cooperator | Free-Rider | Reciprocator | Efficiency (%) | # Obs |
|---|---|---|---|---|---|
| CWAs | 10 | 40 | 50 | 64.0 | 15 |
| MAMs | 14 | 21 | 64 | 67.4 | 15 |
| METs | 10 | 20 | 70 | 64.1 | 15 |
| UGs (8 sessions) | 4 | 12 | 84 | 67.0 | 175 |

**Table 3:** Average Trust Index values

| | Water Managers | Students |
|---|---|---|
| Cooperators | 2.25 | 1.83 |
| Free-riders | 1.89 | 0.31 |
| Reciprocators | 2.05 | 0.82 |
| Group average | 2.03 | 0.81 |

**Table 4:** Subject types by estimation method

| OLS Regression | Shares of Type (%) Coop. | Free-Rider | Recip. | Hump | No Type (count) |
|---|---|---|---|---|---|
| Students | | | | | |
| all coefficients | 4 | 12 | 84 | | 8 |
| signif. coeff. only | 3 | 27 | 69 | | 1 |
| quadratic (all coeff.) | 4 | 10 | 72 | 14 | 6 |
| Water Managers | | | | | |
| all coefficients | 12 | 26 | 62 | | 8 |
| signif. coeff. only | 13 | 45 | 43 | | 2 |
| quadratic (all coeff.) | | *not estimated* | | | |

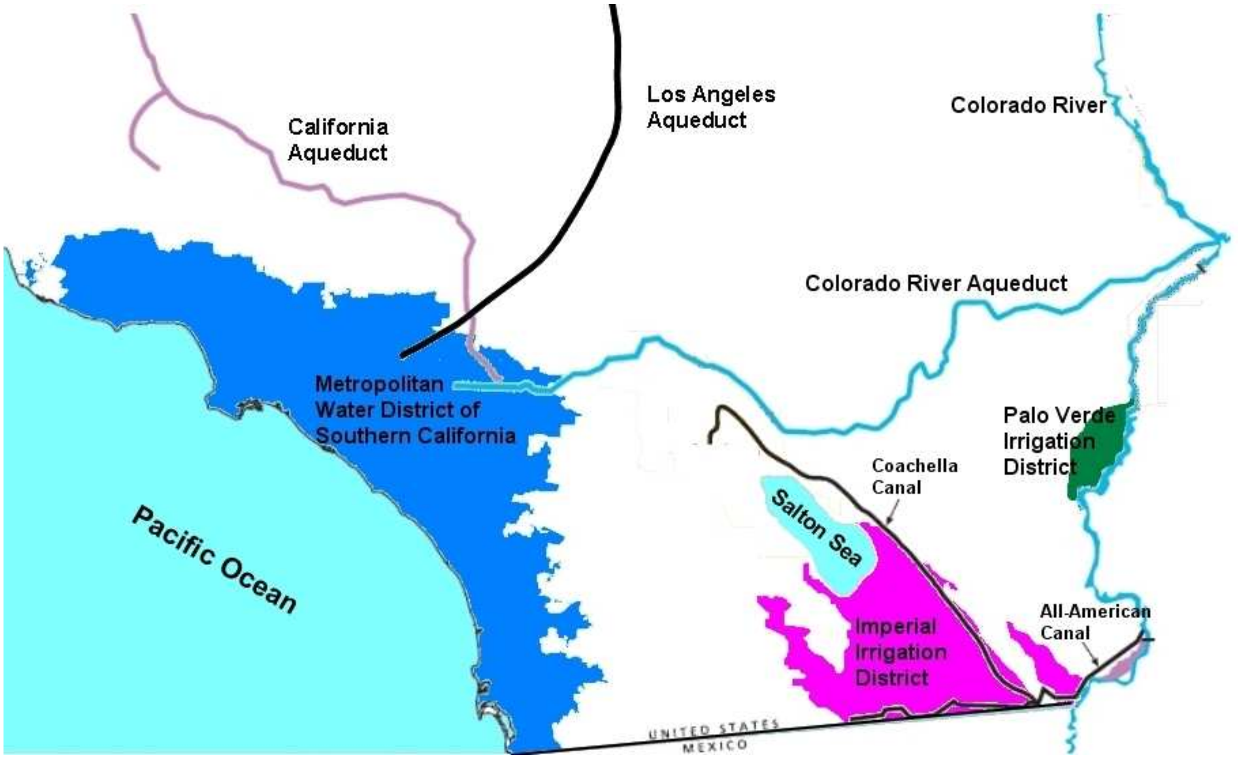


**Figure 1:** MET's service area and water sources

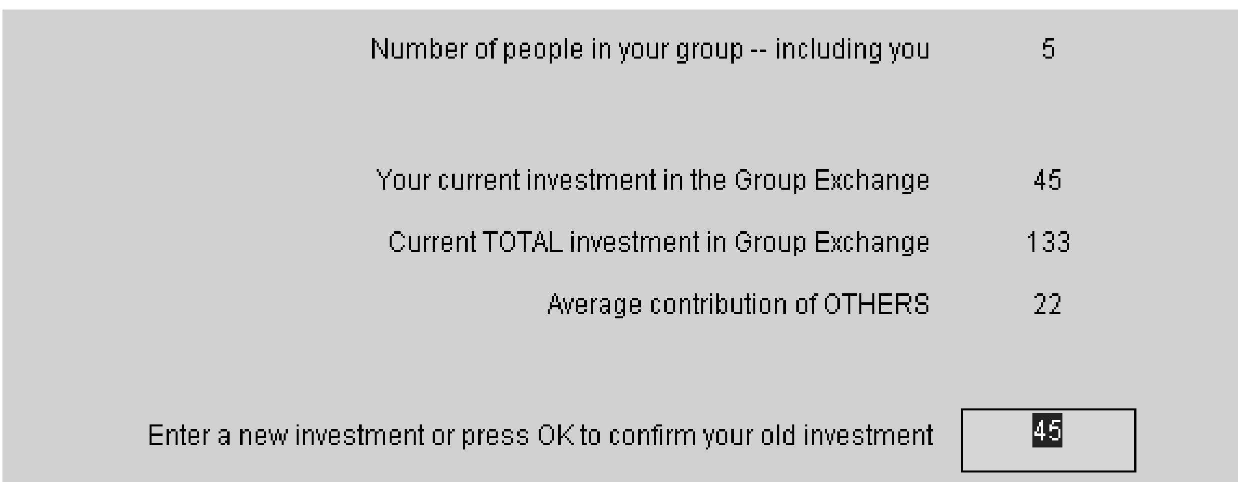


**Figure 2:** Screenshot from cooperation game

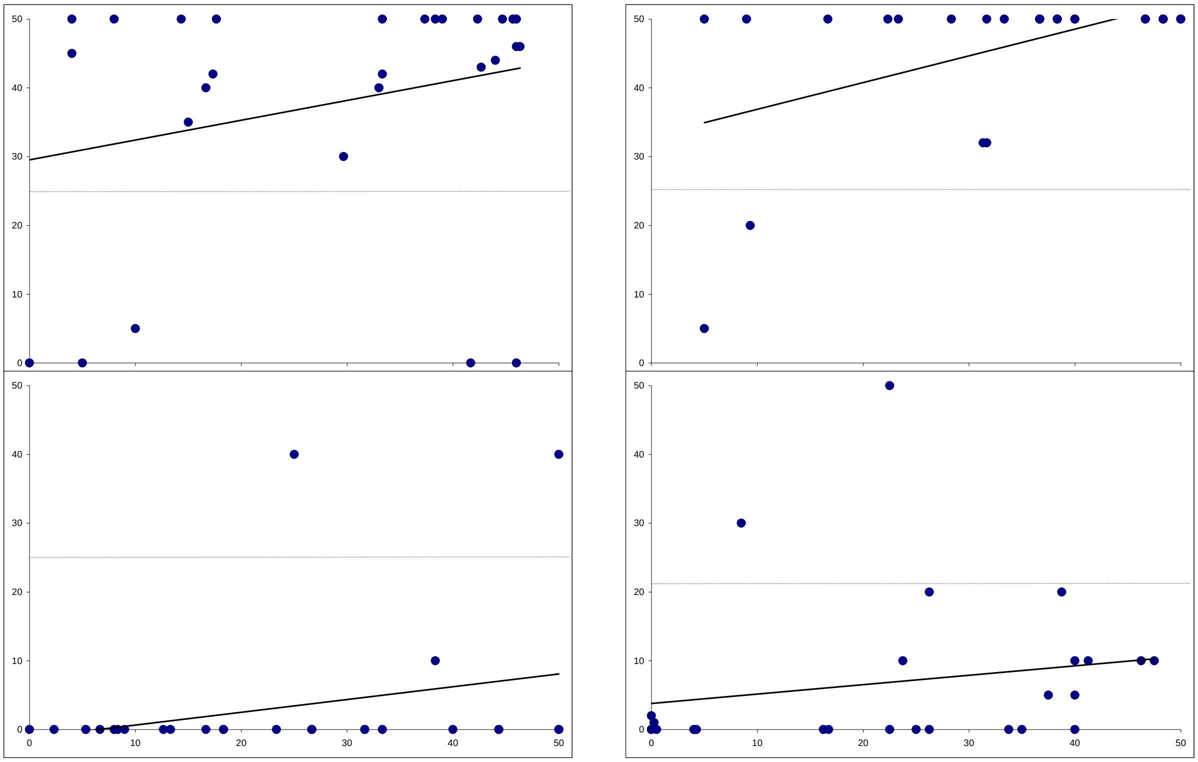


**Figure 3:** Plots for four subjects typed as cooperators (top) and free-riders (bottom). Contributions (y-axis) are in response to the average contribution of others (x-axis).

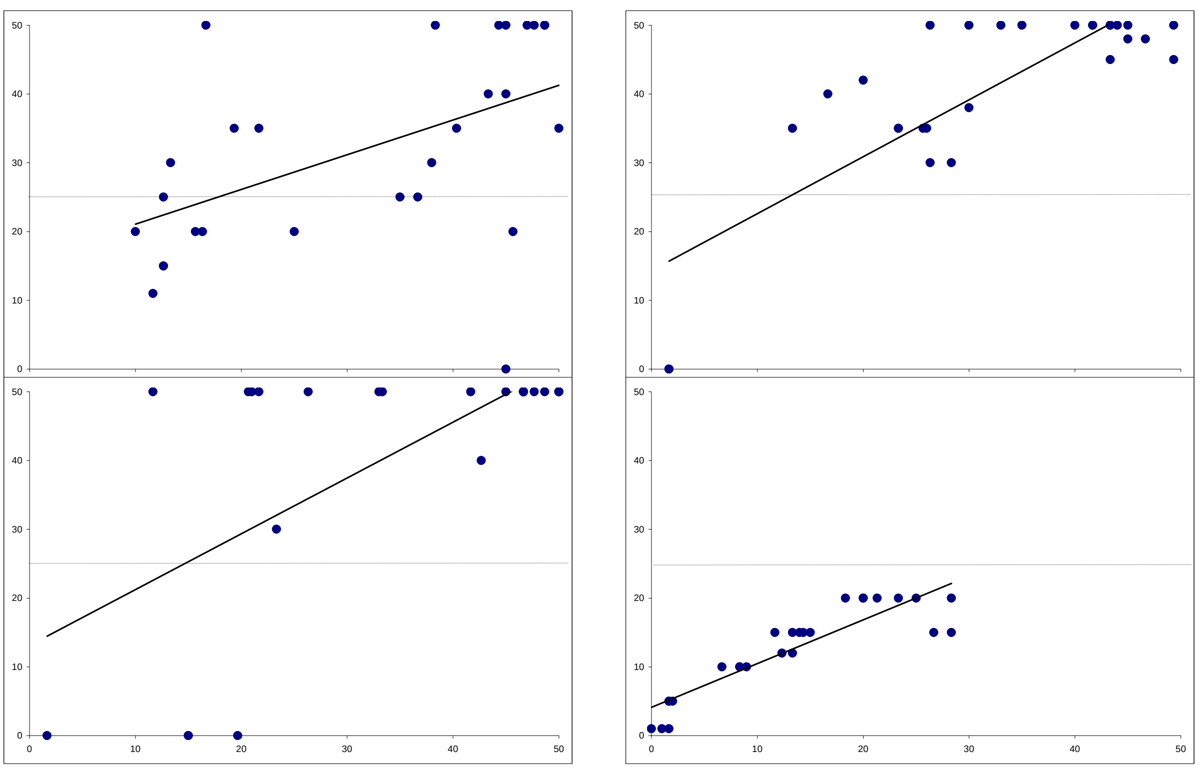


**Figure 4:** Plots for four subjects typed as reciprocators

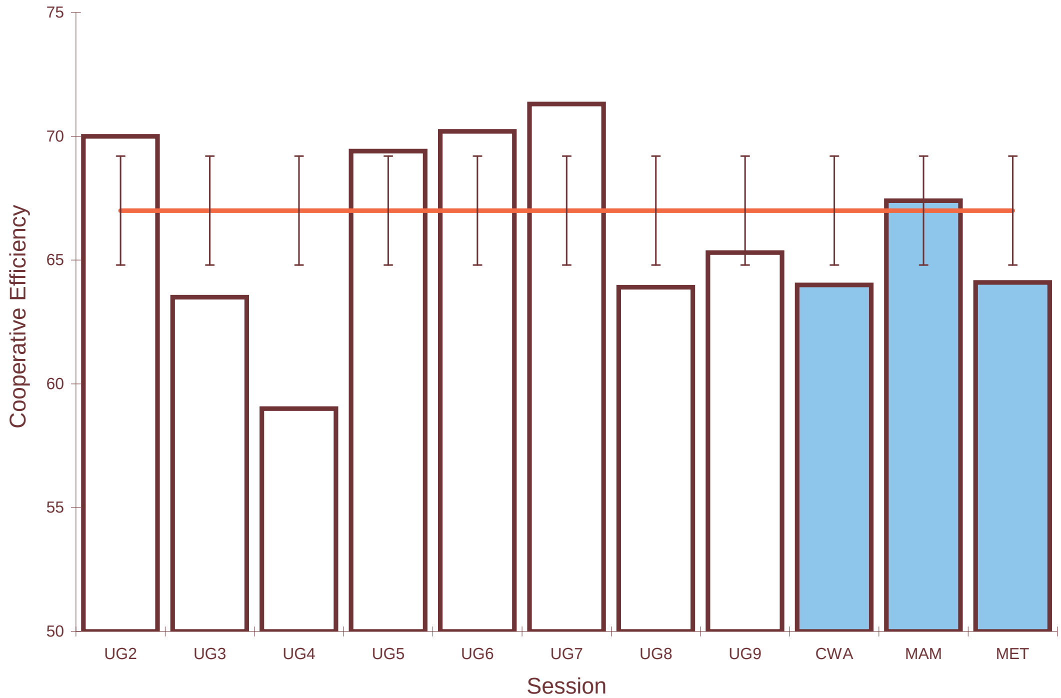


**Figure 5:** Water manager and student efficiencies (cooperation) are similar. The horizontal line indicates mean student efficiency, with error bars for the 95 percent confidence interval.

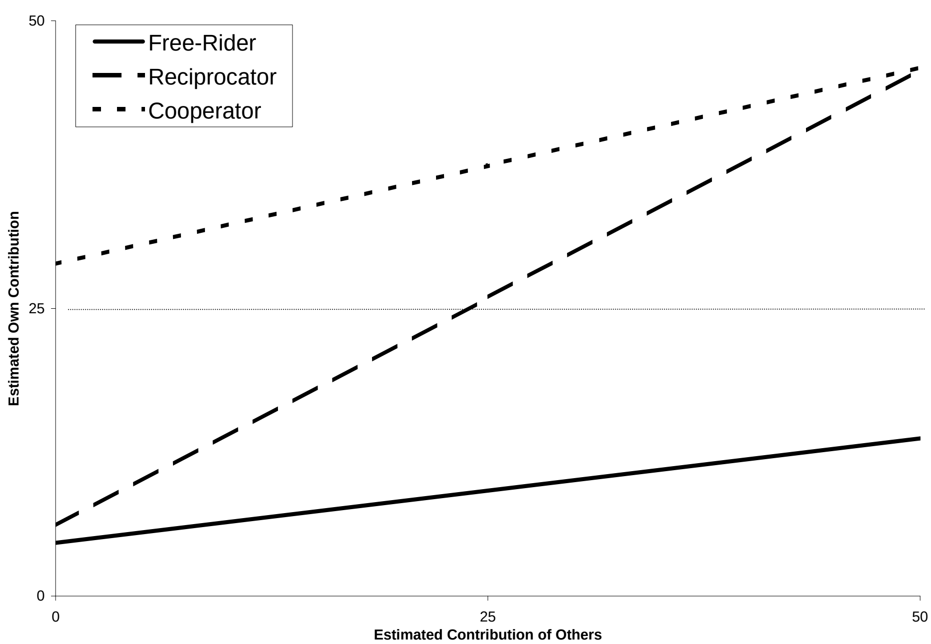


**Figure 6:** Average types from linear characterization of students.

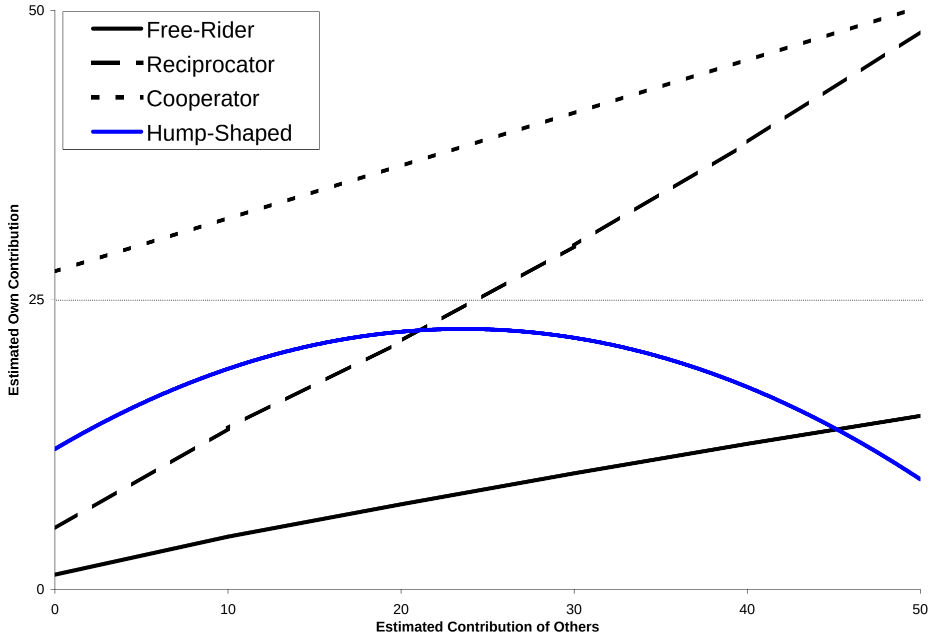


**Figure 7:** Average types from quadratic characterization of subjects.