Preference for Efficiency or Confusion? A Note on a Boundedly Rational Equilibrium Approach to Individual Contributions in a Public Good Game

Luca CORAZZINI* Marcelo TYSZLER†

February 2011

Abstract

Does the hypothesis of preference for (group) efficiency account for subjects' over-contribution in public good games or is this mostly noise? Using a boundedly rational equilibrium approach, we aim at estimating the relative importance of efficiency concerns relative to a noise argument. By using data from a VCM experiment with heterogeneous endowments and asymmetric information, we estimate a quantal response equilibrium (QRE) extension of a model in which subjects have preference for group efficiency. Under the hypothesis of homogeneous population most of the over-contribution seems to be explained by noisy behaviors. A different picture emerges when we introduce cross-subject heterogeneity in concerns for group efficiency. In this case, the majority of the subjects makes contributions that are compatible with the hypothesis of preference for (group) efficiency. A formal likelihood-ratio test strongly rejects the models not allowing for noise in contributions and homogeneous subjects for the more general QRE extension with heterogeneous preferences for (group) efficiency coupled with noise in subjects' behavior.

Keywords: Preference for (Group) Efficiency; Voluntary Contribution Mechanism; Quantal Response Equilibrium; Laboratory Experiment; Bounded Rationality; Public Good Game.

JEL classification: C92, D03, D71.

---

*Department of Economic Sciences "Marco Fanno", University of Padua, Via Del Santo 33, 35123 Padova, Italy and ISLA, Bocconi University, via Roentgen, 1, 20136, Milan, Italy. E-mail: luca.corazzini@unipd.it.

†*Corresponding author.* Center for Research in Experimental Economics and political Decision making (CREED), University of Amsterdam; Roetersstraat 11, 1018WB Amsterdam, The Netherlands. E-mail: marcelo@tyszler.com.br. Phone: +31205254109.

1 Introduction

Subjects contribute to public goods even in situations in which it is individually optimal to free-ride. Amongst the experimental paradigms, over-contribution in linear public good experiments represents one of the best documented and most studied regularities. In order to explain this evidence, social scientists have elaborated a large number of behavioral explanations that are based on refinements of the hypothesis of "other-regarding preferences": reciprocity (Sugden, 1984; Hollander, 1990; Falk and Fischbacher, 2006; Fischbacher et al., 2001), altruism and spitefulness (Levine, 1998; Andreoni 1989; Andreoni, 1990), commitment and Kantianism (Laffont, 1975; Bordignon, 1990), norm compliance (Bernasconi et al., 2010), and team-thinking (Bacharach et al., 2006; Sugden, 2003; Cookson, 2000).

Recently, the hypothesis of preference for (group) efficiency has been invoked as an additional psychological explanation for agents' attitude to freely engage in pro-social behaviors. Indeed, there is evidence showing that experimental subjects often make choices that increase group efficiency, even at the cost of sacrificing their own payoff (Charness and Rabin, 2002; Engelman and Strobel, 2004). Corazzini et al. (2010) use this behavioral hypothesis to explain evidence from linear public good experiments based on prizes (a lottery, a first price all pay auction and a voluntary contribution mechanism used as a benchmark) and characterized by endowment heterogeneity and incomplete information on the distribution of incomes. In particular, they present a simple model in which subjects bear psychological costs from contributing less than what is efficient for the group. The main theoretical prediction of their model when applied to linear public good experiments is that the equilibrium contribution of a subject is increasing in both her endowment and the weight attached to the psychological costs of (group-)inefficient contributions in the utility function. The authors show that this model is capable of accounting for over-contribution as observed in their experiment as well as evidence reported by related studies.

However, as argued by several scholars, rather than being related to subjects' kindness, over-contribution may reflect their natural propensity to make errors. There are several experimental studies (Andreoni, 1995; Palfrey and Prisbrey, 1996, 1997; Brandts and Schram, 2001; Houser and Kurzban, 2002; Goeree et al., 2002) that seek to disentangle other-regarding preferences from noisy behaviors by running *ad hoc* variants of the linear public good game. A general finding in these papers is that *"warm-glow effects and random error played both important and significant roles"* (Palfrey and Prisbrey, 1997, p. 842) in explaining over-contribution.

In a similar vein, one may wonder about the relative importance of noise and preference for efficiency in explaining the experimental evidence. In order to tackle this question, we build and estimate a quantal response equilibrium (henceforth, QRE; McKelvey and Palfrey, 1995) extension of the model presented by Corazzini et al. (2010). This boundedly rational model formally incorporates both preference for efficiency and the noise arguments. Moreover, in contrast to previous studies that aim to find the relative importance of error and other-regarding preferences, the QRE approach explictly applies an equilibrium analysis.

There are several alternative theoretical frameworks that can be used to model noisy behaviors (bounded rationality) and explain experimental evidence in strategic games. Two examples are the "level-k" model (e.g. Stahl and Wilson, 1995; Ho et al., 1998; Stahl and Haruvy, 2008) and (reinforcement) learning models (e.g. Erev and Roth, 1998). In the "level-k" model of iterated dominance, "level-0" subjects choose an action randomly and with equal probability over the set of possible pure strategies while "level-k" subjects choose the action that represents the best response against level-$(k-1)$ subjects. Level-k models have been used to explain experimental results in games in which other-regarding preferences do not play any role, such as p-Beauty contests and other constant sum game. Since in public good games there is a strictly dominant strategy of no contribution, unless other-regarding preferences are explicitly assumed, "level-k" models do not apply. Similar arguments apply to learning models . In the basic setting, each subject takes her initial choice randomly and with equal probability over the set of possible strategies. As repetition takes place, strategies that turn out to be more profitable are chosen with higher probability. Thus, unless other-regarding preferences are explicitly incorporated into the utility function, repetition leads to the Nash Equilibrium of no contribution.

The QRE approach has the advantage that even in the absence of other-regarding preferences it can account for over-contribution in equilibrium. Moreover, we can use the model to assess the relative importance of noise and efficiency concerns.

For our QRE approach, we start from a benchmark model in which the population is homogeneous in both concerns for (group) efficiency and the noise parameter. Then, we allow for heterogeneity across subjects by assuming the population to be partitioned into subgroups with the same noise parameters but distinct in the preference for (group) efficiency.

In line with our theoretical setting, we use data from the VCM sessions of Corazzini et al. (2010) to compare estimates from the model not accounting for noise in subjects' behaviors with those from the QRE extension. For the QRE model with a homogeneous population, we find that subjects' over-contribution is entirely explained by noise in behaviors, with the estimated parameter of concerns for (group) efficiency being zero. A formal likelihood ratio test strongly rejects the specification not allowing for randomness in contributions in favor of the more general QRE model. A different picture emerges in the QRE model with heterogeneous subjects. In the model with two sub-groups, the probability of a subject being associated with a strictly positive degree of preference for (group) efficiency is approximately one third. This probability increases to 59% when we add a third subgroup characterized by an even higher efficiency concern. A formal likelihood ratio test confirms the supremacy of the QRE model with three subgroups over the other specifications. These results are robust to learning processes over repetitions. Indeed, estimates remain qualitatively unchanged when we replicate our analysis on the last 25% of the experimental rounds. The rest of this paper is structured as follows. In section 2, we describe the experimental setting of Corazzini et al. (2010). In section 3, we present the QRE extension of the model based on the preference for (group) efficiency hypothesis. Section 4 reports results from our statistical analysis. Section 5 concludes.

2 The Experiment

Our statistical analysis is based on the experimental results reported by Corazzini et al. (2010). More specifically, we use data from three sessions of a voluntary contribution mechanism with endowment heterogeneity and incomplete information. Each session consisted of 20 rounds and involved 16 subjects. At the beginning of each session, each subject was randomly and anonymously assigned an endowment of either 120, 160, 200, or 240 tokens. The endowment assigned at the beginning was kept constant throughout the 20 rounds of the experiment and this was common knowledge. The experiment was run in a strangers condition (Andreoni, 1988) such that, at the beginning of each round, subjects were randomly and anonymously rematched in groups of four players. Thus, in each round subjects made their choices under incomplete information on the distribution of the endowments in their group. In each round, every subject had to allocate her endowment between an individual and a group account. While subjects allocated tokens to the accounts, payoff were expressed in points. The individual account implied a private benefit such that for each token a subject allocated to the individual account, she received two points. On the other hand, tokens in the group account generated monetary returns to each of the group members. In particular, each subject received one point for each token allocated by her, or by any other member of her group to the group account. Thus, the marginal per capita return used in the experiment was 0.5. At the beginning of each round, the experimenter exogenously allocated 120 tokens to the group account, independently of subjects' choices, thus implying 120 extra points for each group member. At the end of each round, subjects received information about their payoffs. Points were converted to euros using an exchange rate of 1000 points per euro. Subjects, mainly undergraduate students of economics, earned 12.25 euros on average for sessions lasting about 50 minutes. The experiment took place in May 2006 in the Experimental Economics Laboratory of the University of Milan Bicocca and was computerized using the z-Tree software (Fischbacher, 2007).

The main features of anonymity and random rematching introduced by Corazzini et al. (2010) in their experimental setting, narrow the relevance of some"traditional" behavioral hypotheses used to explain subjects over-contribution. For instance, they preclude subjects' possibility to reciprocate group members (un)kind contributions (Rabin, 1993). Moreover, under these conditions, subjects with preferences for equality cannot make compensating contributions to reduce (dis)advantageous inequality (Fehr and Schmidt, 1999, Bolton Ockenfels, 2000). Rather, the hypothesis of preference for (group) efficiency as a particular form of warm-glow (Andreoni 1989; Andreoni, 1990) appears a more plausible justification.

3 Theoretical Predictions and Estimation Procedure

Consider a finite set of subjects $P = \{1, 2, ..., p\}$. In a generic round, subject $i \in P$, with endowment $w_i \in N^+$ contributes g_i to the group account, with $g_i \in N^+$ and $0 \leq g_i \leq w_i$. The monetary payoff of subject i who contributes g_i in a round is given by

$$\pi_i(w_i, g_i) = 2(w_i - g_i) + 120 + g_i + G_{-i}, \tag{1}$$

where G_{-i} is the sum of the contributions of group members other than i in that round. Given equation (1), if subjects' utility only depends on the monetary payoff, zero contribution is the unique Nash equilibrium of each round. In order to explain the positive contributions observed in their experiment, Corazzini et al. (2010) assume that subjects suffer psychological costs if they contribute less than what is optimal for the group. In particular, psychological costs are introduced as a convex quadratic function of the difference between a subject's endowment (i.e. the social optimum) and her contribution. In the VCM, player i's (psychological) utility function is given by:

$$u_i(w_i, g_i, \alpha_i) = \pi_i(w_i, g_i) - \alpha_i \frac{(w_i - g_i)^2}{w_i} \tag{2}$$

where α_i is a non-negative and finite parameter measuring the weight attached to the psychological costs, $\frac{(w_i-g_i)^2}{w_i}$, in the utility function. Notice that psychological costs are increasing in the difference between a subject's endowment and her contribution. Under these assumptions, in each round, there is a unique Nash equilibrium in which individual i contributes:

$$g_i^{NE} = \frac{2\alpha_i - 1}{2\alpha_i} w_i \tag{3}$$

The higher the value of α_i, the higher the equilibrium contribution of subject i is. The average relative contribution, g_i/w_i, observed by Corazzini et al. (2010) in their VCM sessions is 22%. By calibrating equation (3) accordingly the authors find an average $\alpha = 0.64$.

Following McKelvey and Palfrey (1995), we introduce noisy decision-making and consider a Logit Quantal Response extension of (2). In particular, we assume subjects to choose their contributions randomly according to a logistic quantal response function. Namely, for a given endowment, w_i, and contributions of the other group members, G_{-i}, the probability that subject i contributes g_i is given by

$$q_i(w_i, g_i, \alpha_i, \mu) = \frac{\exp\left\{\frac{u_i(w_i, g_i, \alpha_i)}{\mu}\right\}}{\sum\limits_{g_j=0}^{w_i} \exp\left\{\frac{u_i(w_i, g_j, \alpha_i)}{\mu}\right\}} \tag{4}$$

where $\mu \in \Re_+$ is a noise parameter reflecting a subject's capacity of noticing differences in expected payoffs.

Therefore each subject i is associated with a w_i-dimenstional vector $\underline{\mathbf{q}}_i(w_i, \underline{\mathbf{g}}_i, \alpha_i, \mu)$ containing a value of $q_i(w_i, g_i, \alpha_i, \mu)$ for each possible contribution level $g_i \in \underline{\mathbf{g}}_i \equiv \{0, \ldots, w_i\}$. Let $\left\{\underline{\mathbf{q}}_i(w_i, \underline{\mathbf{g}}_i, \alpha_i, \mu)\right\}_{i \in P}$ be the system including $\underline{\mathbf{q}}_i(w_i, \underline{\mathbf{g}}_i, \alpha_i, \mu)$, $\forall i \in P$. Notice that since others' contribution, G_{-i}, enters the r.h.s of the system, others' q_i will also enter the r.h.s. A fixed point of $\left\{\underline{\mathbf{q}}_i(w_i, \underline{\mathbf{g}}_i, \alpha_i, \mu)\right\}_{i \in P}$ is, hence, a Quantal Response Equilibrium (QRE), $\left\{\underline{\mathbf{q}}_i^{QRE}(w_i, \underline{\mathbf{g}}_i, \alpha_i, \mu)\right\}_{i \in P}$.

In equilibrium, the noise parameter μ reflects the dispersion of subjects' contributions around the Nash prediction expressed by equation (3). The higher μ, the higher the dispersion of contributions. As μ tends to infinity, contributions are randomly drawn from a uniform distribution defined over $[0, w_i]$. On the other hand, if μ is equal to 0, the equilibrium contribution collapses to the Nash equilibrium.[1]

[1]More specifically, for each subject i equilibrium contributions converge to $q_i(w_i, g_i^{NE}, \alpha_i, 0) = 1$

In this framework, we use data from Corazzini et al. (2010) to estimate α and μ, jointly. We proceed as follows. Our initial analysis is conducted by using all rounds ($n = 20$) and assuming the population to be homogeneous in both α and μ. This gives us a benchmark that can be directly compared with the results reported by Corazzini et al. (2010). In our estimation procedure, we use a likelihood function that assumes each subject's contributions to be drawn from a multinomial distribution. That is:

$$L_i(w_i, \underline{\mathbf{g}}_i, \alpha, \mu) = \frac{n!}{\prod\limits_{g_j=0}^{w_i} n(g_j)!} \prod_{g_k=0}^{w_i} q_i^{QRE}(w_i, g_k, \alpha, \mu)^{n(g_k)} \tag{5}$$

where $n(g_j)$ is the number of times subject i contributed g_j over the n rounds of the experiment, and similarly for $n(g_k)$. The contribution of each person to the log-likelihood is the log of expression (5). The Maximum Likelihood procedure consists of finding the non-negative values of μ and α (and corresponding QRE) that maximize the summation of the log-likelihood function evaluated at the experimental data. In other words, we calculate the multinomial probability of the observed data by restricting the theoretical probabilities to QRE probabilities only.

We then extend our analysis to allow for cross-subject heterogeneity. In particular, we generalize the QRE model above by assuming the population to be partitioned into S subgroups that are characterized by the same μ but different α. In this case, the likelihood function becomes:

$$L_i(w_i, \underline{\mathbf{g}}_i, \alpha_1, \alpha_2, ..., \alpha_S, \gamma_1, \gamma_2, ..., \gamma_S, \mu) = \sum_{s=1}^{S} \gamma_s \frac{n!}{\prod\limits_{g_j=0}^{w_i} n(g_j)!} \prod_{g_k=0}^{w_i} q_i^{QRE}(w_i, g_k, \alpha_s, \mu)^{n(g_k)} \tag{6}$$

where $\gamma_1, \gamma_2, ..., \gamma_S$, with $\sum\limits_{s=1}^{S} \gamma_s = 1$, are the probabilities for agent i belonging to the sub-group associated with $\alpha_1, \alpha_2, ..., \alpha_S$, respectively. This allows us to estimate the value of μ for the whole population, the value of $\alpha_1, \alpha_2, ..., \alpha_S$ for the S subgroups and the corresponding probabilities, $\gamma_1, \gamma_2, ..., \gamma_S$. For identification purposes we impose that $\alpha_s \leq \alpha_{s+1}$. The introduction of one group at a time accompanied by a corresponding likelihood-ratio test allows us to determine the number of $\alpha-$groups that can be statistically identified from the original data. In the following statistical

and $q_i(w_i, g_i, \alpha_i, 0) = 0, \forall g_i \neq g_i^{NE}$.

analysis, estimates account for potential dependency of subject's contributions across rounds. Confidence intervals at the 0.01 level are provided using the inversion of the likelihood-ratio statistic, subject to parameter constraints, in line with Cook and Weisberg (1990), Cox and Hinkley (1974) and Murphy(1995).

4 Results

Using data from the 20 rounds of the experiment, table (1) reports average contributions (both by endowment type and overall) observed in the experiment, average contributions as predicted by the model not accounting for noise in subjects' contributions and estimates as well as average contributions from different parameterizations of the Logit Quantal Response extension of the model. In particular, specification (1) refers to a version of the model in which both α and μ are constrained to be equal to benchmark values based on Corazzini et al. (2010). Under this parameterization, α is fixed to the value computed by calibrating equation (3) on the original experimental data, 0.64, while μ is constrained to 1[2].

[Table 1 about here]

As shown by the table, specification (1) closely replicates predictions of the original model presented by Corazzini et al. (2010) not accounting for noise in subjects' contributions. In specification (2), α is fixed to 0.64, while μ is estimated by using equation (4). The value of μ increases substantially with respect to the benchmark value used in specification (1). A likelihood-ratio test strongly rejects specification (1) that imposes restrictions on the values of both α and μ in favor of specification (2) in which μ can freely vary on $\Re_+$ ($LR = 10460.33; \Pr\{\chi^2(1) > LR\} < 0.01$). However, if we compare the predicted average contributions of the two specifications, we find that specification (1) better approximates the original experimental data. This is because a higher value of the noise parameter spread the distributions of contributions around the mean. Therefore even with mean contributions further from the data (induced by the fixed value of α) the spread induced by the noise parameter in specification (2) produces a better fit. This highlights the importance of taking into account not only the average (point) predictions, but also the spread around it. It also suggests that allowing α to vary can improve fit.

[2] Appendix A shows the Maximum Likelihood estimation value of α when we vary μ. It is possible to see that for a large range of values of μ this value is close to 0.64. We choose $\mu = 1$ as a sufficiently low value in which the estimated α is close 0.64 and thus provide a noisy version of the base model which can be used for statistical tests.

In specification (3), α and μ are jointly estimated using equation (5), subject to $\alpha \geq 0$. If both parameters can freely vary over $\Re_+$, α reduces to zero and μ reaches a value that is higher than what was obtained in specification (2). As confirmed by a likelihood-ratio test, specification (3) fits the experimental data better than both specification (1) ($LR = 11086.54; \Pr\{\chi^2(2) > LR\} < 0.01$) and specification (2) ($LR = 626.21; \Pr\{\chi^2(1) > LR\} < 0.01$). Thus, under the maintained assumption of homogeneity our estimates suggest that contributions are better explained by randomness in subjects' behavior rather than by concerns for efficiency.

In order to control for learning effects, we replicate our analysis using the last five rounds only.

[Table 2 about here]

Consistent with a learning argument, in both specifications (2) and (3), the values of μ are substantially lower than the corresponding estimates in table (1). Thus, repetition reduces randomness in subjects' contributions. The main results presented above are confirmed by our analysis on the last five periods. Looking at specification (3), in the model with no constraints on the parameters, the estimated value of α again drops to 0. Also, according to a likelihood-ratio test, specification (3) explains the data better than both specifications (1) ($LR = 1578.83; \Pr\{\chi^2(2) > LR\} < 0.01$) and (2) ($LR = 203.85; \Pr\{\chi^2(1) > LR\} < 0.01$)

These results seem to reject the preference for (group) efficiency hypothesis in favor of pure randomness in subjects' contributions. A different picture emerges when we allow for cross-subject heterogeneity, however. In table (3) we drop the assumed homogeneity. In particular, we consider two models with heterogeneous subjects: the first assumes the population to be partitioned into two sub-groups ($S = 2$) and the second into three subgroups ($S = 3$).[3] As before, we conduct our analysis both by including all rounds of the experiment and by focusing on the last five repetitions only.

[Table 3 about here]

We find strong evidence of heterogeneity. Focusing on the analysis across all rounds, according to the model with two sub-groups, a subject is associated with $\alpha_1 = 0$ with probability 0.66 and with $\alpha_2 = 0.53$ with probability 0.34. Results

[3] We have also estimated a model with $S = 4$. However, adding a fourth sub-group does not significantly improve the goodness of fit of the model compared to the specification with $S = 3$.In particular, with $S = 4$, the point estimates for the model with all periods are: $\mu = 21.81$, $\alpha_1 = 0$, $\alpha_2 = 0.38$, $\alpha_3 = 0.61$, $\alpha_4 = 1.04$, $\gamma_1 = 0.39$, $\gamma_2 = 0.42$, $\gamma_3 = 0.09$.

are even sharper in the model with three subgroups: in this case $\alpha_1 = 0$ and the two other $\alpha-$parameters are strictly positive: $\alpha_2 = 0.43$ and $\alpha_3 = 1.04$. Subjects are associated with these values with probabilities 0.41, 0.50 and 0.09, respectively. Thus, in the more parsimonious model, the majority of subjects contribute in a way that is compatible with the preference for (group) efficiency hypothesis. These proportions are in line with findings of previous studies (Andreoni, 1995; Houser and Kurzban, 2002; Brandts and Schram, 2001) in which, aside from confusion, social preferences explain the behavior of about half of the experimental population.

Allowing for heterogeneity across subjects reduces the estimated randomness in subjects' contributions: the value of μ reduces from 41.59 in specification (3) of the model with homogeneous population, to 28.50 and 22.14 in the model with two and three subgroups, respectively. According to a likelihood-ratio test, both the models with $S = 2$ and $S = 3$ fit the data better than the (unconstrained) specification of the model with homogeneous subjects (for the model with $S = 2$, $LR = 117.25; \Pr\{\chi^2(2) > LR\} < 0.01$; whereas for the model with $S = 3$, $LR = 174.66; \Pr\{\chi^2(4) > LR\} < 0.01$). Moreover, adding an additional subgroup to the model with $S = 2$, significantly increases the goodness of fit of the specification ($LR = 57.42; \Pr\{\chi^2(2) > LR\} < 0.01$). As before, all these results remain qualitatively unchanged when we control for learning processes and we focus on the last 5 experimental rounds.

In order to check for the robustness of our results in table (3), we have also estimated additional specifications accounting for heterogeneity in both concerns for (group) efficiency and noise in subjects' behaviors. Although the log-likelihood of the model with both sources of heterogeneity significantly improves in statistical terms, the estimated values of the $\alpha-$parameters remain qualitatively the same of those reported in the third column of table (3).

5 Conclusions

Is over-contribution in linear public good experiments explained by subjects' preference for (group) efficiency or does it rather simply reflect their natural attitude to make errors? In order to answer this fundamental question, we have built and estimated a quantal response equilibrium model in which, in choosing their contributions, subjects are influenced by both a genuine concern for (group) efficiency and a random noise in their behavior.

In line with other studies, we find that both concerns for (group) efficiency and noise in behaviors play an important role in determining subjects' contributions. However, assessing which of these two behavioral hypotheses is more relevant in explaining contributions strongly depends on the degree of cross-subject heterogeneity admitted by the model. Indeed, by estimating a model with homogeneous subjects, the parameter capturing concerns for (group) efficiency vanishes while noise in behavior entirely accounts for over-contribution. A different picture emerges when we allow the subjects to be heterogeneous in their concerns for efficiency. By estimating a model in which the population is partitioned into three subgroups that differ in the degree of concerns for efficiency, we find that the most of the subjects contribute in a way that is compatible with the preference for (group) efficiency hypothesis. A formal likelihood-ratio test confirms the supremacy of the QRE model with three subgroups over the other specifications.

Previous studies (Andreoni, 1995; Palfrey and Prisbrey, 1996, 1997; Brandts and Schram, 2001; Houser and Kurzban, 2002; Goeree et al. 2002) tried to disantangle the effects of noise from other-regarding preferences mainly by manipulating the experimental design. Our approach adds a theorethical foundation in the form of an equilibrium analysis. In contrast to studies which focus mostly on (direct) altruism, we follow Corazzini et al. (2010) and allow for preference for efficiency. Our results are in line with the literature in the sense that we also conclude that a combination of noise and social concerns play a role. Our results, however, are directly supported by a sound theorethical framework proven valid in similar settings (e.g., Goeree and Holt, 2005).

Recent studies (Fischbacher and Gächter, 2006; Erlei, 2008) have emphasized the importance of admitting heterogeneity in social preferences in order to better explain experimental evidence. In this paper we show that neglecting heterogeneity in subjects' social preferences may lead to erroneous conclusions on the relative importance of the love for (group) efficiency hypothesis with respect to the confusion argument. Indeed, as revealed by our analysis, the coupling of cross-subject heterogeneity in concerns for (group) efficiency with noise in the decision process seems to be the relevant connection to better explain subjects' contributions.

6 Acknowledgements

We thank Arthur Schram, Jens Grosser and participants to the 2010 Credexea Meeting at the University of Amsterdam for useful comments and suggestions.

Table 1. Homogeneous population (all rounds)

	Data	*CFS*	(1) $\overline{\mu}, \overline{\alpha}$	(2) $\mu, \overline{\alpha}$	(3) μ, α
μ	–	–	1	21.83	41.59
				[19.69; 24.34]	[39.11; 44.34]
α	–	0.64	0.64	0.64	0
					[0; 0.01]
(Predicted) Avg. Contributions					
Overall endowments	37.91	39.38	39.41	60.24	37.91
$w_i = 120$	34.02	26.25	26.34	44.84	34.12
$w_i = 160$	24.53	35.00	35.03	55.68	37.67
$w_i = 200$	47.50	43.75	43.76	65.57	39.48
$w_i = 240$	45.57	52.50	52.50	74.86	40.36
$\log ll$			-8713.95	-3483.79	-3170.69
Obs.	960	960	960	960	960

This table reports average contributions as well as estimates and predictions from various specifications of the model based on the efficiency concerns assumption using all 20 rounds of the experiment. CFS refers to the specification not accounting for noise in subjects' contributions while (1), (2) and (3) are Logit Quantal Response extensions of the model. In (1) α and μ are constrained to 0.64 and 1, respectively. In (2), the value on α is set to 0.64, while μ is estimated through equation (5). Finally, (3) refers to the unconstrained model in which both α and μ are estimated through equation (5). The table also reports, for each specification, the corresponding log-likelihood. Confidence Intervals are computed using a inversion of the likelihood-ratio statistic, at the 0.01 level, subject to parameter constraints.

Table 2. Homogeneous population (last 5 rounds)

	Data	*CFS*	(1) $\overline{\mu}, \overline{\alpha}$	(2) $\mu, \overline{\alpha}$	(3) μ, α
μ		–	1	11.63	26.91
				[9.80; 13.95]	[24.14; 30.17]
α		0.59	0.59	0.59	0
					[0; 0.03]
(Predicted) Avg. Contributions					
Overall endowments	25.94	26.39	26.91	44.78	25.94
$w_i = 120$	21.13	17.59	18.44	34.55	25.05
$w_i = 160$	17.63	23.46	24.04	41.68	26.01
$w_i = 200$	35.28	29.32	29.71	48.30	26.30
$w_i = 240$	29.70	35.19	35.45	54.60	26.39
$\log ll$			−1675.03	−987.54	−885.62
Obs.	240	240	240	240	240

This table reports average contributions as well as estimates and predictions from various specifications of the model based on the efficiency concerns assumption using the last 5 rounds of the experiment only. The same remarks as in table (1) apply.

Table 3. Heterogeneous Subjects (all and last 5 rounds)

	$\mu, \alpha_1, \alpha_2 (n = 20)$	$\mu, \alpha_1, \alpha_2 (n = 5)$	$\mu, \alpha_1, \alpha_2, \alpha_3 (n = 20)$	$\mu, \alpha_1, \alpha_2, \alpha_3 (n = 5)$
μ	28.50	15.07	22.14	14.25
	[25.88; 31.26]	[12.90; 17.64]	[20.56; 23.95]	[12.04; 16.85]
α_1	0	0	0	0
	[0; 0.01]	[0; 0.02]	[0; 0.01]	[0; 0.02]
α_2	0.53	0.54	0.43	0.48
	[0.46; 0.60]	[0.47; 0.61]	[0.39; 0.46]	[0.40; 0.56]
α_3			1.04	0.76
			[0.92; 1.16]	[0.53; 1.01]
γ_1	0.66	0.63	0.41	0.59
	[0.53; 0.78]	[0.50; 0.75]	[0.33; 0.46]	[0.49; 0.64]
γ_2			0.50	0.34
			[0.43; 0.55]	[0.23; 0.40]
(Predicted) Avg. Contributions				
Overall endowments	37.15	25.57	38.51	25.72
$w_i = 120$	32.06	22.17	31.97	22.07
$w_i = 160$	36.04	24.63	36.78	24.68
$w_i = 200$	39.01	26.77	40.83	26.99
$w_i = 240$	41.48	28.73	44.45	29.13
$\log ll$	−3112.06	−865.75	−3083.35	−865.16
$Obs.$	960	240	960	240

This table reports estimates and predictions from two specifications of the model with efficiency concerns accounting for cross-subject heterogeneity in the value of α. The analysis is conducted both by including all experimental rounds and by focusing on the last five repetitions only. Parameters are estimated through equation (6). Given the linear restriction $\sum_{s=1}^{S} \gamma_s = 1$, we only report estimates of $\gamma_1, \gamma_2, ..., \gamma_{S-1}$.Confidence Intervals are computed using a inversion of the likelihood-ratio statistic, at the 0.01 level, subject to parameter constraints.

References

Andreoni, J. (1988). Why free ride? Strategies and learning in public goods experiments. Journal of Public Economics, 37(3), 291-304.

Andreoni, J. (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence. Journal of Political Economy, 97(6), 1447-1458.

Andreoni, J. (1990) Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving? Economic Journal, 100(401), 464-477.

Andreoni, J. (1995). Cooperation in Public-Goods Experiments: Kindness or Confusion? American Economic Review, 85(4), 891-904.

Bacharach, M. (2006). Beyond Individual Choice: Teams and Frames in Game Theory. Edited by N. Gold and R. Sugden. Princeton University Press.

Bernasconi, M, Corazzini, L, Marenzi, A. (2010). 'Expressive' Obligations in Public Good Games: Crowding-in and Crowding-out Effects. Working Papers 2010_04, University of Venice "Ca' Foscari", Department of Economics.

Bolton, G.E. and Ockenfels, A. (2000). A Theory of Equity, Reciprocity and Competition, American Economic Review, 90, 166-193.

Bordignon, M. (1990). Was Kant Right? Voluntary Provision of Public Goods under the Principle of Unconditional Commitment. Economic Notes, 0, 342-372.

Brandts, J, Schram, A. (2001). Cooperation and Noise in Public Goods Experiments: Applying the Contribution Function Approach. Journal of Public Economics, 79(2), 399-427.

Charness, G, Rabin, M. (2002). Understanding Social Preferences With Simple Tests. Quarterly Journal of Economics, 117(3), 817-869.

Cook, R.D., and Weisberg, S. (1990). Confidence Curves in Nonlinear Regression. Journal of the American Statistical Association, 85:544-551

Cookson,R. (2000). Framing Effects in Public Goods Experiments. Experimental Economics, 55(3), 55-79.

Cox, D.R., and Hinkley, D.V. (1974). Theoretical Statistics. London: Chapman and Hall.

Corazzini, L, Faravelli, M, Stanca, L. (2010). A Prize to Give For: an Experiment on Public Good Funding Mechanisms, The Economic Journal, 120:944-967.

Engelmann, D, Strobel, M. (2004). Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Reply. American Economic Review, 94(4), pp. 857-869.

Erev, I, Roth, A.E. (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. American Economic Review, 88(4), 848-881.

Erlei, M. (2008). Heterogeneous Social Preferences. Journal of Economic Behavior and Organization, 65, 436-457.

Falk, A, Fischbacher, U. (2006). A Theory of Reciprocity. Games and Economic Behavior, 54(2), 293-315.

Fehr, E. and Schmidt, K.M. (1999). A Theory Of Fairness, Competition, and Cooperation, Quarterly Journal of Economics, 114, 817-868.

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

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

Fischbacher, U, Gachter, S. (2006). Heterogeneous Social Preferences and the Dynamics of Free Riding in Public Goods. CeDEx Discussion Paper 2006-01.

Goeree, J.K, Holt, C.A. (2005). An explanation of Anomalous Behavior in Models of Political Participation. American Political Science Review, 99(2), 201-213.

Goeree, J.K, Holt, C.A, Laury, S.K. (2002). Private costs and public benefits: unraveling the effects of altruism and noisy behavior. Journal of Public Economics, 83(2), 255-276.

Ho, T.H, Camerer, C, Weigelt, K. (1998). Iterated Dominance and Iterated Best Response in Experimental “p-Beauty Contests”. American Economic Review, 88(4), 947-69.

Hollander, H. (1990). A Social Exchange Approach to Voluntary Contribution. American Economic Review, 80(5), 1157-1167.

Houser, D, Kurzban, R. (2002). Revisiting Kindness and Confusion in Public Goods Experiments. American Economic Review, 92(4),1062-1069.

Laffont, J.J. (1975). Macroeconomic Constraints, Economic Efficiency and Ethics: An Introduction to Kantian Economics. Economica, 42, 430-437.

Levine, D.K. (1998). Modeling Altruism and Spitefulness in Experiment. Review of Economic Dynamics, 1(3), 593-622.

McKelvey, R.D, Palfrey, T.R. (1995). Quantal Response Equilibria for Normal Form Games. Games and Economic Behavior, 10(1), 6-38.

Murphy, S. A. (1995). Likelihood Ratio-Based Confidence Intervals in Survival Analysis. Journal of the American Statistical Association Vol. 90, No. 432, pp. 1399-1405

Palfrey, T.R, Prisbrey, J.E. (1996). Altruism, reputation and noise in linear public goods experiments. Journal of Public Economics, 61(3), 409-427.

Palfrey, T.R, Prisbrey, J.E. (1997). Anomalous Behavior in Public Goods Experiments: How Much and Why? American Economic Review, 87(5), 829-46.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics, American Economic Review, 83(5): 1281-1302.

Rogers, B.W.,Palfrey T.R., Camerer, C.F. (2009). Heterogeneous quantal response equilibrium and cognitive hierarchies, Journal of Economic Theory, 144(4), 1440-1467.

Stahl, D.O, Haruvy, E. (2008). Level-n bounded rationality in two-player two-stage games. Journal of Economic Behavior and Organization, 65(1), 41-61.

Stahl, D.O, Wilson, P.W. (1995). On Players' Models of Other Players: Theory and Experimental Evidence. Games and Economic Behavior, 10(1), 218-254.

Sugden, R. (1984). Reciprocity: the Supply of Public Goods Through Voluntary Contributions. Economic Journal, 94(376), 772-787.

Sugden, R. (2003). The Logic of Team Reasoning. Philosophical Explorations, 6, 165-181.

Appendix A

Table (A1) shows the Maximum Likelihood value of α and the log-likelihood according to equation (5) as μ decreases from 1000 to 0.4. As shown by the table, for high values of μ the estimated value of α is 0. When μ is equal to 10, the estimated value of alpha is 0.50. Moreover, for μ lower than 2.00, the estimated value of α is 0.61. For the specification tests presented in section (4), we set $\mu = 1$. This is a sufficiently low value of μ in order to generate a noisy version of the base model. Two arguments indicates why this choice is valid. First, for a range of values including $\mu = 1$, the estimated α is stable. Moreover, since the log-likelihood of a model with $\alpha = 0.61$ and $\mu = 1$ is higher than that corresponding to a model with μ=0.4 (and similarly for $\alpha = 0.64$), the choice any μ lower than 1 for the benchmark value would only reinforce the results of section (4). More specifically, both likelihood-ratio statistics comparing specifications (1) with specifications (2) and (3) of tables (1) and (2) would increase.

Table A1

μ	α	log-likelihood
1000.00	0	−3637.64
500.00	0	−3591.79
333.33	0	−3548.7
250.00	0	−3508.34
200.00	0	−3470.67
166.67	0	−3435.64
142.86	0	−3403.19
125.00	0	−3373.25
111.11	0	−3345.76
100.00	0	−3320.64
90.91	0	−3297.8
83.33	0	−3277.17
76.92	0	−3258.65
71.43	0	−3242.16
66.67	0	−3227.61
62.50	0	−3214.93
58.82	0	−3204.01
55.56	0	−3194.79
52.63	0	−3187.18
50.00	0	−3181.1
40.00	0	−3171.22
30.30	0.14	−3192.52
20.00	0.33	−3247.22
10.00	0.50	−3444.42
9.09	0.52	−3488.26
8.00	0.53	−3555.57
7.04	0.55	−3634.45
5.99	0.56	−3753.67
5.00	0.57	−3916.89
4.00	0.58	−4173.35
3.00	0.59	−4615.92
2.00	0.60	−5547.35
1.00	0.61	−8506.13
0.90	0.61	−9181.67
0.80	0.61	−10032.07
0.70	0.61	−11133.38
0.60	0.61	−12612.63
0.50	0.61	−14699.13
0.40	0.61	−17852.64

This table reports Maximum Likelihood estimates of α for selected values of μ (see equation (5)). The last column reports the corresponding log-likelihood value.