Individual Evolutionary Learning, Other-regarding Preferences, and the Voluntary Contributions Mechanism*

Jasmina Arifovic† & John Ledyard‡

March 14, 2009

Abstract

The data from experiments with the Voluntary Contributions Mechanism suggest four stylized facts. To date, no theory explains all of these facts simultaneously. We provide a new theory by merging our Individual Evolutionary Learning model with a variation of other-regarding preferences and a distribution of types. The data generated by our model are remarkably similar to data from a variety of experiments. Further, the data generated do not seem to depend in any crucial way on the exact values of the parameters of our model. That is, we have a robust explanation for most behavior in VCM experiments.

*We thank Olena Kostyshyna and Lilong Shi for their help in generating the simulations in this paper. We would also like to thank participants at the 2008 Econometric Society Summer Meetings, Pittsburgh; 14th International Conference on Computing and Finance, 2008, Paris; 2008 North-American ESA Conference, Tucson and the participants at the seminars at University of Amsterdam, June 2008; Paris 2 University, June 2008; and University of New South Wales, August 2008.

†Simon Fraser University, arifovic@sfu.ca

‡California Institute of Technology, jledyard@caltech.hss

1 Introduction

The Voluntary Contributions Mechanism (VCM) is often used to decide how much of a public good to produce and how to fund it. There is a large body of experiments involving the VCM with a small number of well-known stylized facts. But to date, no theory fully explains these facts. Neither standard game theory nor other-regarding preferences nor learning models can, by themselves alone, explain these facts. Two studies by Cooper and Stockman (2002) and Janssen and Ahn (2006) have taken the natural step and have merged other-regarding preferences and learning models. They meet with some success in explaining the data but both the utility functions and the learning models they use suffer from some drawbacks.

In this paper, we merge an evolutionary style learning model, that includes experimentation and consequent "errors", with a distribution of heterogeneous other-regarding preferences. We produce simulated behavior that is remarkably similar to that observed in the experiments. Our model is robust in two senses. First, it relies on a small number of parameters and the precise values of those parameters do not matter very much. Second, the model generates data consistent with those of different experiments and different experimenters. Our results do raise some questions about the other-regarding preference models of Fehr-Schmidt (1999) and Charness-Rabin (2002).

The Experiments Beginning with the pioneering work of Marwell and Ames (1979), Isaac, McCue, and Plott (1980) and Kim and Walker (1981), there have been an amazing range of experiments involving the VCM. Some well-known[1] stylized facts from the many experiments with linear public goods are:

(1) Average contributions begin at around 50% of the total endowment and then decline with repetition, but not necessarily to zero.

(2) There is considerable variation in individual contributions in each repetition. Some give everything. Some give nothing. The individual contributions also show no consistent monotonic pattern over time. Some increase, some decrease, and some have a zig-zag pattern.

(3) Increases in the marginal value of the public good relative to the private good lead to an increase in the average rate of contribution. This is particularly true in later repetitions and for small groups.

(4) Increases in the size of the group lead to an increase in the the average rate of contribution. This is particularly true in later repetitions and for small values of the marginal value of the public good.

Standard Theory Standard theory provides no help in understanding these facts. In linear experiments, the prediction from game theory is that contributing zero is the dominant strategy in a one-shot game. Solving backwards implies zero contribution is also the equilibrium in games with multiple rounds. There have been many suggested modifications to the standard theory in an attempt to explain the experimental data. These include other-regarding preferences (altruism and fairness), mistakes, and learning. Holt and Laury (2002) do an excellent

[1]See Ledyard (1995), Holt and Laury (2002), and Croson (2007).

job of summarizing this literature. But it will be useful to review the few of those papers that are most closely related to our work.

Other-regarding Preferences Some have suggested that subjects bring other-regarding preferences into the lab with them which the experimenter cannot control and which cause the subjects to behave differently than predicted if one ignores these preferences.[2] Fehr and Schmidt (1999), Bolton and Ochenfels (2000), Charness and Rabin (2002), and Cox et. al. (2007, 2008) are among those who have taken this route. The idea is simple. The experimenter controls the payoff to the subject. But, each subject has their own utility function that depends on others' payoffs as well as their own, and which the experimentalist cannot control.[3] As we will show later, in the linear public good environment other-regarding preferences lead to three types of equilibrium behavior: selfishness (these give nothing in equilibrium), altruism (these give everything in equilibrium), and fair-mindedness (these give the average contribution in equilibrium). The equilibrium model captures the fact there is variation in individual contribution rates with some giving a lot (the more altruistic) and some giving little (the more selfish). It also predicts that contributions are responsive to the marginal value of the public good. But, these are static theories in a one-shot game.[4] Other-regarding preference by itself is not a sufficient explanation for the behavior observed in experiments.

Learning Some have suggested that it takes time for subjects to find the dominant strategy. The decline in contributions over periods strongly suggests that individuals begin choosing randomly (this would explain the initial average 50% contribution rates) and then learn to play better over time. Andreoni and Miller (1991) use the replicator dynamic to simulate behavior in linear public goods VCM environments. They show that the replicator dynamic will produce decay towards free riding over time as more cooperative strategies will be replaced by those that give less to the public good. Given sufficient time, the entire population will use the strategy that does the most free riding. The decay to free riding is slower the larger the group size and the larger the marginal return. But the decay is always there and is inconsistent with the higher rates of contribution seen in experiments with high marginal value of the public good. Any learning rule that search for and find better paying strategies will have this problem. Since contributing zero is a dominant strategy, it is always better paying and will eventually be found. So contributions will always converge to zero. Learning by itself is not a sufficient explanation for the behavior observed in experiments.

[2] For some, the other-regarding behavior is triggered by or framed by the experiment itself. It is not necessary for us to distinguish between the two approaches for this paper. However, it is important if one is interested in transferring the model from the VCM world to, say, a world with only markets and private goods. We discuss this further below in Section 7.

[3] We will go into more detail on the particulars of some of the utility specifications in Section 3 below.

[4] Charness and Rabin (2002) include a term in their utility function that provides a mechanism for modeling reciprocity or reactions if others "misbehave". This has the potential to lead to declining contributions over time, although for the most part they use it to define equilibrium and not dynamics. But, since we are modeling repeated games and, in that context, believe that retaliation is a strategic concept and not a basic preference concept, we prefer to leave retaliation out of the utility function.

First Attempts to Merge Only two papers have taken the obvious step and combined other-regarding preferences and learning.[5] Cooper and Stockmann (2002) merge a Bolton-Ockenfels (2002) type fairness utility function with the Reinforcement Learning model of Erev-Roth (1995). Janssen and Ahn (2006) merge a Charness-Rabin (2002) type social utility function[6] with the Experience Weighted Attraction (EWA) learning model of Camerer and Ho (1999). In both cases, the authors are reasonably successful in generating behavior similar to that observed in their respective experiments. But, in both cases there are some problems with their approaches.

One drawback is in the utility functions they use. The utility function used by Cooper and Stockman ignores altruistic motives. This is no problem for the environment they study - a 3 player minimal contributing set (MCS) game. However, as we will see below, the data from VCM experiments reject the hypothesis that there is no altruism involved. The utility function used by Janssen and Ahn has the property that individual equilibrium contributions are independent of the size of the group. However, as we will see below, the data from VCM experiments rejects the hypothesis that the size of the group does not matter.

Another drawback is in the learning models they use. Both the reinforcement learning used by Cooper and Stockman and the EWA learning used by Janssen and Ahn are ill suited to repeated games with strategy spaces that are a continuum. The models generate learning behavior that is generally much slower than that of experimental subjects.[7] This is not much of a problem for the research reported in each of the papers, although to have a successful fit Janssen and Ahn were required to estimate eight parameters for each agent in each experiment. With a small number of observations for each agent, such an estimate is mostly an ex post calibration. Because of this, it is not clear that their model and parameters would transfer to other experiments, either VCM or others, without requiring additional ex post re-estimation of the parameters.

These drawbacks lead us to believe there is something to be learned by turning to an alternative model of behavior and an alternative model of other-regarding preferences.

Our Approach In this paper, we merge our Individual Evolutionary Learning (IEL) model with a distribution of heterogeneous other-regarding preferences and produce simulated behavior that is remarkably similar to that observed in the experiments. IEL is a model we have used successfully before in a private goods setting with a call-market mechanism,[8] and in a public goods setting with the Groves-Ledyard mechanism.[9] The stage games, generated

[5] A paper by Wendel and Oppenheimer (2007) uses an agent-based model as we do, but they take a much different approach to learning and fairness. Agents' preferences include the other-regarding component that can be dampened by a 'sense' of exploitation, i.e. a sense of contributing more than the average. The sensitivity to the average contribution (that can change over time) makes preferences 'context-dependent'. Agents make probabilistic decisions given their utility functions in each time period. We prefer to maintain the approach of game theory and model utility separately from strategy. By doing so we believe the model will have broader applicability than to just VCMs.

[6] This is essentially equivalent to the Fehr-Schmidt (1999) utility function.

[7] See Arifovic and Ledyard (2004).

[8] See Arifovic and Ledyard (2007).

[9] See Arifovic and Ledyard (2008).

from the specific mechanisms in those settings, have the property that equilibria are Pareto-optimal so there is little conflict between public and private interests. We have brought IEL unchanged from those applications to the VCM experiments, to see whether it would be as successful when there is a major conflict.

In IEL, agents retain a finite set of remembered strategies. After each iteration, they update this set through experimentation and replication. Experimentation involves replacing, with low probability, some of the elements of the set with a strategy chosen at random from the entire strategy space. Experimentation introduces strategies that might otherwise never have a chance to be tried. Replication goes through the set of remembered strategies and, through a series of random paired comparisons within the set, replaces the strategies which would have provided a low payoff in the previous period with copies of those that would have yielded a higher payoff. Over time the remembered set becomes homogeneous with copies of the "best reply" strategy. To generate their strategic choices, individuals choose strategies from the set of remembered strategies, at random, proportionately to the payoff they would have received had they been played in the last round. The end result of IEL is a Markov process in strategies where the remembered set of each agent, the state space of the process, co-evolves with the sets of all other agents.

We assume that agents have other-regarding preferences. They are assumed to have a utility function that is linear in their own payoff, in the average payoff to the group and in the amount by which their payoff is less than the average payoff to the group. These three pieces reflect, respectively, a selfish preference, an altruistic preference, and a preference for fairness to self. This differs from Fehr-Schmidt (1999) and Charness-Rabin (2002) in a subtle but significant way. They use the average payoff to the rest of the group, without i, rather than the average across the whole group, with i. Although ours and theirs are linear transformations of each other, the difference turns out to be crucial.[10] For example, their specification implies that, in a VCM game, the equilibrium allocations are independent of the number of participants. This implication is not consistent with the data.

To complete the model we assume that an agent's two utility parameters, their marginal utility for altruism and their marginal disutility for unfairness, are independently and identically drawn from a probability distribution which has three parameters itself. We allow a different range of parameters than either Fehr-Schmidt (1999) or Charness-Rabin (2002). As we will see, ours is more consistent with the data.

To simulate an experiment with N individuals, we first draw N values of utility. This is like asking Nindividuals into the lab. We then use IEL with these preferences to simulate the choices of the subjects. We can repeat this as often as we want, each time simulating a new experimental trial. At the end we can compute the average rate of contribution in VCM for the IEL model with other-regarding preferences (ORP). We can then compare these to the data generated by the experiments themselves.

Summary of Findings The major finding is that our model generates behavior similar to that generated by experiments. Across a range of different experiments and different

[10]The technical reason, discussed in a later section, is that the transformation depends on N, the number of participants.

experimenters, we find that the average contributions generated by our model differ from those in the experiments by only 3.4% to 6.6%.

We also find that the precise values of the few parameters of our model do not matter very much. IEL has 3 parameters: the size of the remembered set and the rate and range of the experimentation. ORP has 3 parameters: those that determine the probability distribution on types. We show that the differences between our model and experimental data changes very little as we change the 6 parameters over a fairly wide range. This suggests that if one wants to transfer our model to other VCM experiments or to other experiments like auctions in a private goods world, there will be little need to re-calibrate the parameters. In fact, the IEL parameters we end up using here are essentially the same as those in call market experiments in Arifovic and Ledyard (2007) and Groves-Ledyard experiments in Arifovic and Ledyard (2004, 2008).

We believe we have a robust explanation for a wide range of experiments. We now turn to a fuller description of our approach and the research findings.

2 IEL and VCM

We begin with a description of a classic experiment with the VCM. We then describe the theory of behavior which we call Individual Evolutionary Learning (IEL). Finally, we test this model against the experimental data.

2.1 The VCM in Linear Environments

A linear public good environment consists of N agents, numbered $i = 1, ..., N$. Each agent has a linear payoff function $\pi^i = p^i(w^i - c^i) + y$, where $1/p^i$ is their marginal willingness to pay in the private good for a unit of the public good, w^i is their initial endowment of a private good, c^i is their contribution to the production of the public good, where $c^i \in [0, w^i]$, and y is the amount of the public good produced. The linear production function is $y = M \sum_{j=1}^{N} c^j$ where M is the marginal value of the public good.

The VCM in a linear environment creates a simple game. There are N players with strategies $c^i \in C^i = [0, w^i]$ and payoffs $\pi^i(c) = p^i(w^i - c^i) + M \sum c^j$. In what follows, we will only consider symmetric VCM problems.[11] That is, we will assume that $p^i = 1$ and $w^i = w$ for all i. For symmetric linear VCMs, it is easy to see that if $M < 1$ then each i has a dominant strategy of $c^i = 0$. It is also true that if $M > (1/N)$, then the aggregate payoff is maximized if $c^i = w, \forall i$. When $(1/N) < M < 1$, there is a tension between private and public interest that is the basis of the standard commons dilemma.

In this paper we are going to focus on the classic experimental design of Mark Isaac and James Walker that has served as the jumping off point for much of the experimental work that has gone on since. They have run a number of experiments in which subjects use the VCM

[11] We will indicate in the appropriate places how our theory applies more generally to asymmetries in p^i and in w^i.

to decide public good allocations.[12] In an experiment N individuals are gathered together. Each individual begins each period with w tokens. Each individual chooses how many of their tokens to contribute to a group exchange with a total group return of $NM \sum c^i$. After summing the c^i, the per capita individual return, $M \sum c^i$, is calculated and each individual is informed of this. Each individual also receives $w - c^i$ from their individual exchange. Thus the payment received by i for this round is $\pi^i(c) = w - c^i + M \sum c^j$.

In the experiment reported in Isaac-Walker (1988), they used a 2x2 design with N equal to 4 and 10 and M equal to 0.3 and 0.75, and structured the experiment as a repeated game, with no re-matching, played for $T = 10$ periods. We display, in Table 1, the average rate of contribution, $(\sum_t \sum_i c_t^i)/(10Nw)$ for each of the 10 periods for each of the 4 treatments in that set of experiments.[13] There are six observations per treatment, all with experienced subjects. The combined averages over all 10 periods are in Table 2. The average rate of contribution for the last three periods for each treatment is in Table 3.

Table 1
Average Rate of Contribution by period

t =	1	2	3	4	5	6	7	8	9	10
M = .3 N= 4	34.3	29.5	17.6	10.1	7.7	9.0	7.1	4.2	2.3	5.8
M = .3 N= 10	46.3	39.5	32.2	26.9	33.0	33.0	30.3	20.8	21.9	8.6
M = .75 N= 4	54.7	56.3	62.3	57.3	62.3	49.5	45	47.7	33.3	29.2
M = .75 N = 10	47.5	56.2	57.8	53	51.7	40.0	51.0	43.7	33.8	29.8

Source: IW(1988)

Table 2
Average Rate of Contribution - all 10 periods

	N = 4	N = 10
M = 0.3	12.8	29.2
M = 0.75	49.8	46.5

Source: IW(1988)

[12]See for example, Isaac, Walker and Thomas (1984), Isaac and Walker (1988), and Isaac, Walker and Williams (1994).

[13]We thank Mark Isaac and James Walker for providing us with all of their data from the 1984, 1988, and 1994 papers.

Table 3
Average Rate of Contribution - last 3 periods

	N = 4	N = 10
M = 0.3	4.1	17.1
M =0.75	36.7	35.8

Source: IW(1988)

These IW(1988) data are consistent with and representative of the four stylized facts listed in the introduction. But what behavior are the subjects really following? The standard full rationality, game theory prediction is that subjects will use their dominant strategy, $c^i = 0$. This is obviously inconsistent with these data. So we need a new theory.

2.2 Individual Evolutionary Learning

IEL is based on an evolutionary process which is individual, and not social. In IEL, each agent is assumed to carry a collection of possible strategies in their memory. These remembered strategies are continually evaluated and the better ones are used with higher probability. IEL is particularly well-suited to repeated games with large strategy spaces such as subsets of the real line.

An experiment can often be modeled as a repeated game (G, R). The repeated game has a stage game G and a number of rounds, T. The idea is that G will be played for T rounds. In $G = \{N, X, \pi, r\}$, N is the number of subjects, X^i is the action space of i, $\pi^i(x^1, ..., x^N)$ is i's payoff if the joint strategy choice is x, and $r^i(x_t)$ describes the information reported to subject i at the end of a round. These are all controlled by the experimenter. In round t, each subject chooses $x^i_t \in X^i$. At the end of round t, subject i will be told the information $r^i(x_t)$ about what happened. Then the next round will be played. A behavioral model must explain how the sequence of choices for i, $(x^i_1, x^i_2, ..., x^i_R)$ is made, given what i knows at each round t.

The primary variables of our behavioral model are a finite set of remembered strategies for each agent i at each round t, $A^i_t \subset X^i$ and a probability measure, π^i_t on A^i_t. A^i_t consists of J alternatives. J is a free parameter of IEL that can be loosely thought of as a measure of the processing and/or memory capacity of the agent. In round t, each agent selects an alternative randomly from A^i_t using the probability density π^i_t on A^i_t and then chooses the action $x^i_t = a^i_t$. One can think of (A^i_t, π^i_t) as inducing a mixed strategy on X^i at t. At the end of each round t, agents are told $r(x_t)$. At the beginning of the next round $t + 1$, each agent computes a new A^i_{t+1} and π^i_{t+1}. This computation is at the heart of our behavioral model and consists of three pieces: experimentation, replication, and selection.

We begin at the end of round, t, knowing A^i_t, π^i_t, and $r^i(x_t)$.

Experimentation comes first. Experimentation introduces new alternatives that otherwise might never have a chance to be tried. This insures that a certain amount of diversity is maintained. For each $j = 1, ..., J$, with probability ρ, a new contribution is selected at random from X^i and replaces $a^i_{j,t}$. We use a normal density, conditional on X^i, for this experimentation. For each j, the mean value of the normal distribution is set equal to the value of the alternative, $a^i_{j,t}$ that is to be replaced by a 'new' idea. That is the new alternative $a \sim N(a^i_{j,t}, \sigma)$. ρ and σ are free parameters of the behavioral model that can be varied in the simulations.

Replication comes next. Replication reinforces strategies that would have been good choices in previous rounds. It allows potentially better paying strategies to replace those that might pay less. The crucial assumption here is the measure of "potentially better paying strategies". We let $u^i(a^i_{jt}|r^i(x_t))$ be the forgone utility of alternative j at t given the information $r^i(x_t)$. This measures the utility i thinks she would have gotten had she played a_j in round t. $u^i(a_j|r^i(x_t))$ is a counter-factual valuation function and must be specified for each application.

Given a forgone utility function, u^i, we can describe how replication takes place. For $j = 1, \ldots, J$, $a^i_{j,t+1}$ is chosen as follows. Pick two members of A^i_t randomly (with uniform probability) with replacement. Let these be $a^i_{k,t}$ and $a^i_{l,t}$. Then

$$a^i_{j,t+1} = \left\{ \begin{array}{c} a^i_{k,t} \\ a^i_{l,t} \end{array} \right\} \text{if} \left\{ \begin{array}{c} u^i(a^i_{kt}|r^i(x_t)) \geq u^i(a^i_{lt}|r^i(x_t)) \\ u^i(a^i_{kt}|r^i(x_t)) < u^i(a^i_{lt}|r^i(x_t)) \end{array} \right\}.$$

Replication for $t+1$ favors alternatives with a lot of replicates at t and alternatives that would have paid well at t, had they been used. So it is a process with a form of averaging over past periods. If the actual contributions of others have provided a favorable situation for an alternative $a^i_{j,t}$ on average then that alternative will tend to accumulate replicates in A^i_t, (it is fondly remembered), and thus will be more likely to be actually used. Over time, the sets A^i_t become more homogeneous as most alternatives become replicates of the best performing alternative.

Selection is last. Each contribution $a^i_{k,t+1}$ is selected with the following probability:[14]

$$\pi^i_{k,t+1} = \frac{u^i(a^i_{k,t+1}|r^i(x_t)) - \varepsilon^i_{t+1}}{\sum_{j=1}^{J}(u^i(a^i_{j,t+1}|r^i(x_t)) - \varepsilon^i_{t+1})}$$

for all $i \in \{1, \ldots, N\}$ and $k \in \{1, ..., J\}$ and where[15]

$$\varepsilon^i_{t+1} = \min_{a \in A^i_{t+1}}\{0, u^i(a|r^i(x_t))\}.$$

[14]An alternative selection model would change the probabilistic choice function to $\pi(a^k) = \frac{e^{\lambda u^i(a^k)}}{\sum_j e^{\lambda u^i(a^k)}}$. We have found (see Arifovic and Ledyard (2004)) that the behavior predicted for any λ differs very little from that generated by our proportional selection rule. This is because the set A tends to become homogeneous fairly fast, at which point the selection rule is irrelevant. We therefore use the proportional rule since it eliminates another parameter.

[15]This implies that if there are negative foregone utilities in a set, payoffs are normalized by adding a constant to each payoff that is, in absolute value, equal to the lowest payoff in the set.

We now have a complete description of the way that an IEL agent moves from A_t^i and π_t^i to A_{t+1}^i and π_{t+1}^i. The only remaining thing to pin down is the initial values, A_1^i and π_1^i.

Initialization We use the simplest possible initialization. It is very naive behavior but this has proven successful in our previous work and we see no need to change it now. We assume that things begin randomly. We let A_1^i be populated randomly with J uniform draws from X^i. We let $\pi_{k,1}^i = 1/J \quad \forall k$.

We now have a complete model of behavior for a general repeated game. The two determining components of IEL are A and $u(a|r(x))$. The three free parameters are (J, ρ, σ). We turn next to its application to Voluntary Contribution Mechanism experiments.

2.3 IEL and VCM

In applying IEL to the VCM experiments, the key choices one needs to make are the set of strategies, A, and the forgone utility function, u^i. We take a very straight-forward approach. The choice of A is easy. In the VCM experiments subjects are given an endowment, w and told to choose at each iteration how much of that to contribute to the public good. That is, their contribution $c^i \in [0, w]$. We let $A = [0, w]$. For the choice of u^i we assume that subjects are risk-neutral, expected money maximizers. That is, we assume they want to maximize, π^i. For most VCM experiments, $r^i(c) = \sum_j c^j$. This is what subjects are told after the choices c^j are made. We use the foregone utility function, $u^i(c^i|r^i(c_t)) = \pi^i(c_t/c^i) = (w - c^i) + M(c^i + r^i(c_t) - c_t^i)$ where $r^i(c_t) = \sum_j c_t^j$ for all i.

The only other choice to make when using IEL is which values of J, ρ, and σ to use. We will show, in section 5.2, that the precise values of these parameters are of little importance. Our results are insensitive to the exact values. So, in this paper we will set $J = 100, \rho = 0.033$, and $\sigma^i = w/10$. These values are similar to the numbers we have used in our other IEL papers.

For the rest of this paper, when we refer to IEL with these parameters we will identify it as IEL*.

2.3.1 An aside: Neutrality to affine transformations

We will be looking at many VCM experiments, some of which differ in their choices of parameters. The generic linear VCM payoff function is $\pi^i = p^i[(w - c^i) + M(\sum c^j)]$. We will be looking at experiments with different values of p^i and w. Can we pool these results?

Our model of behavior, IEL, as applied to the VCM above, implies that choices, however random, will be independent of the value of p. This is easy to see by noting that p enters the IEL process only at the stages of replication and selection. Affine transformations of the payoff function do not affect these choices. Because of this, we will not differentiate between experiments with different p's.[16]

[16] We should note that it is commonly believed by experimenters that p will have an effect on the data. Higher p's are expected to produce lower variance in the data, due perhaps to higher attention rates. The effect of p on the mean data is often believed to be neutral. Our model ignores these subtleties.

Our model of behavior, IEL, as applied to the VCM above is also independent of the value of w in the sense that the percentage contribution rates made by IEL, when normalized by w will be independent of w. Consider the % contribution rate, $g^i = c^i/w$, and rewrite the payoff function as $\pi = w[(1 - g^i) + M \sum g^j]$. Then changes in w are similar to changes in p which we have already seen are neutral. So IEL produces percentage contribution rates that are independent of w.[17]

2.4 Initial simulations

To see whether IEL* is an adequate model of the behavior in VCM experiments, we need to compare the choices generated by IEL* with those observed in the laboratory. We begin by doing this for the IW(1988) data.

We use IEL* to generate 40 simulated experimental sessions for some of the 4 treatments of (M, N) used by Isaac and Walker. We then compare these simulation results to the experimental data from IW(1988). In Figure 1, we show the paths of contributions (averaged over 6 sessions) over 10 periods of the IW(1988) experimental treatments with $M = 0.3$ and $N = 4$, and $M = 0.75$ and $N = 4$. We display the comparable paths for the average contributions from the IEL* simulations. The implications of these comparisons are clear. If we only considered treatments with $M = 0.3$, we might conclude that the IEL* data and the IW(1988) data are very similar. The initial contributions are around 50% and then, presumably because of learning, decline towards 0. But if we consider the treatments with $M = 0.75$, we must conclude that the behavior generated by IEL* is not that observed in the laboratory. The patterns of IEL* contributions for $M = 0.75$ are virtually indistinguishable from those for $M = 0.3$. But in the IW(1988) data, a higher M clearly elicits higher rates of contributions.

One should not be surprised by this. IEL is about learning how to participate in a game with payoffs of π^i. The choices that IEL makes will converge to the Nash equilibria of the game (X, π) which is zero contributions. If contributions for the $M = 0.75$ are declining to the equilibrium level of zero, they are going much more slowly than predicted by IEL. If IEL is to be consistent with the data, we need to have equilibrium possibilities other than zero.[18]

3 Other-regarding Preferences

As was shown in the previous section, the level of contributions observed in experiments is often higher than is consistent with selfish preferences, even allowing for a learning process. As has been suggested by many others, one way to explain the levels of contribution observed in

[17] If endowments are asymmetric in that $w^i \neq w^j$ for some i, j, then this "neutrality" still holds for proportional changes in endowments. That is, as long as the new endowments $\hat{w}^i = \lambda w^i$ for all i and some $\lambda > 0$ then the rate of contributions generated by IEL will remain the same.

[18] In their 1994 paper, Isaac, Walker and Williams allowed some of their sessions to go as long as 60 rounds, rather than stopping at 10. The rate of contribution did continue to decline past round 10. But they did this only for $M = 0.3$. We do not know of any experiments testing the hypothesis that, when $N = 4$ and $M = 0.75$, with enough rounds contributions will converge to zero.

experiments is to assume subjects bring other-regarding preferences with them to the experiment and that this is not controlled by the experimenter. We introduce such an assumption. We do this by adding a dash of altruism and fairness to some, but not all, agents.[19]

In the experiments, each agent receives a payoff $\pi^i(c^i) = w - c^i + M\sum c^j$. We take altruism to be a preference for higher values of the average payoff to all agents,[20] $\bar{\pi} = \sum \pi^i/N = w - \bar{c} + MN\bar{c}$ where $\bar{c} = \sum c^i/N$. We take (one-sided) fairness to be a disutility for being taken advantage of[21] which happens when $\bar{\pi}(c) > \pi^i(c)$. That is, i loses utility when i's payoff is below the average payoff in this group. We maintain the linearity assumptions and model subjects as having a utility function

$$u^i(c) = \pi^i(c) + \beta^i \bar{\pi}(c) - \gamma^i \max\{0, \bar{\pi}(c) - \pi^i(c)\}. \tag{1}$$

We assume that $\beta^i \geq 0$ and $\gamma^i \geq 0$.

It is clear from experimental data that there is considerable heterogeneity in the population with respect to levels of other-regarding behavior. But the experimenter can neither observe nor control for each agent's values of (β, γ). We model the heterogeneity by assuming that each subject i comes to the lab endowed with a value of (β^i, γ^i) which is distributed, independently and identically, in the population according to a distribution $F(\beta, \gamma)$. We will be more precise about the functional form of $F(\cdot)$ below.

It will be useful to know what the Nash Equilibrium levels of contributions are for any particular draw of subject parameters (β, γ). It is very easy to characterize Nash equilibrium for the VCM in symmetric environments where $w^i = w \quad \forall i$. With the linear other-regarding preferences (1), given (N, M) and heterogeneity across (β, γ), there are only three types of Nash Equilibrium strategies: free riding or selfishness ($c^i = 0$), fully contributing or altruism ($c^i = w^i$), and contributing an amount equal to the average or fair-minded behavior ($c^i = \bar{c} = (\sum_i c^i)/N$). In equilibrium,

$$c^i = \left\{ \begin{array}{c} 0 \\ \bar{c} \\ w \end{array} \right\} \text{if} \left\{ \begin{array}{l} 0 \geq \left[(M - \frac{1}{N})\beta^i + M - 1\right] \\ \gamma^i(\frac{N-1}{N}) \geq \left[(M - \frac{1}{N})\beta^i + M - 1\right] \geq 0 \\ \gamma^i(\frac{N-1}{N}) \leq \left[(M - \frac{1}{N})\beta^i + M - 1\right] \end{array} \right\} \tag{2}$$

Which strategy a particular agent uses depends on that agent's parameters (β^i, γ^i). For our utility formulation (1), the ratio $\dfrac{1-M}{M-\frac{1}{N}}$, acts as a differentiator between selfish and non-selfish behavior in equilibrium for VCM. This is also true in non-linear utility approaches to other-regarding preferences as found, for example, in Cox, Friedman and Gjerstad (2007), Cox, Friedman and Sadiraj (2008) and Cox and Sadiraj (2007).

[19] We do this is a way that differs from the approach of others such as Fehr and Schmidt (1999) and Charness and Rabin (2002). Our reasons for this are discussed in section 3.1.

[20] We considered having preferences depend on the sum of payoffs to all, $\sum \pi^j$, but that swamps everything else when N becomes larger, which does not seem to be consistent with the data.

[21] We considered having preferences depend symmetrically and 2-sidedly on $(\bar{\pi} - \pi^i)^2$, as in Ledyard(1995), but that produced behavior that is inconsistent with the data since the decline in contributions over time predicted by the model appears to be too slow under this hypothesis.

To compute the equilibrium for a particular draw of parameters, one first counts how many selfish types there are, those for whom $0 \geq (M - \frac{1}{N})\beta^i + M - 1$. These all give 0 in equilibrium. Let N_s be the number of these selfish types. Next one counts how many altruists there are, those for whom $0 \leq \gamma^i(\frac{N-1}{N}) \leq (M - \frac{1}{N})\beta^i + M - 1$. These all give w in equilibrium. Let N_a be the number of altruists. In equilibrium, the average percentage contribution is $\bar{c}/w = N_a/(N_a + N_s)$.

It is straight-forward to do comparative statics on this equilibrium. In particular, for any specific realization of the parameters, $\partial N_s/\partial M \leq 0$ and $\partial N_a/\partial M \geq 0$. So $\partial \frac{\bar{c}}{w}/\partial M \geq 0$. Further, $\partial N_s/\partial N \leq 0$ and $\partial N_a/\partial N \geq 0$. So $\partial \frac{\bar{c}}{w}/\partial N \geq 0$. If the distribution of the parameters, $F(\cdot)$ is continuous, then the comparative statics of the expected value of $\frac{\bar{c}}{w}$ all hold with strict inequality.

3.1 Relation to other ORP models

Others have proposed other-regarding preferences before us, suggesting a variety of functional forms and parameter restrictions.[22] Equation (1) is intimately related to those of Fehr-Schmidt (1999), Bolton-Ockenfels (2000), and Charness-Rabin (2002). However, the differences in functional form are to some extent irrelevant (particularly in their linear forms). All are equivalent up to linear transformations of the other where equivalent utility functions will yield identical behavior for a wide range of behavioral models including Expected Utility Maximizing and IEL.

The Fehr-Schmidt asymmetric fairness utility function is:[23]

$$u_{FS} = \pi^i - a^i \max\{\bar{\pi}_{-i} - \pi^i, 0\} - b^i \max\{\pi^i - \bar{\pi}_{-i}, 0\} \tag{3}$$

where $\bar{\pi}_{-i} = \dfrac{1}{N-1}\sum_{j \neq i} \pi^j$.

To see the equivalence notice that $(\bar{\pi}_{-i} - \pi^i) = \dfrac{N}{N-1}(\bar{\pi} - \pi^i)$ and that $\max\{\pi^i - \bar{\pi}, 0\} = (\pi^i - \bar{\pi}) + \max\{\bar{\pi} - \pi^i, 0\}$. So $u_{FS} = \pi(1 - B) + B\bar{\pi} - (A + B)\max\{\bar{\pi} - \pi^i, 0\}$, where $A = \dfrac{N}{N-1}a$ and $B = \dfrac{N}{N-1}b$. Given u_{FS}, let $\beta = \dfrac{B}{1-B}$ and $\gamma = \dfrac{A+B}{1-B}$. Then it will be true that $u_{AL} = Ku_{FS}$ where u_{AL} is our utility functional form and $K = \dfrac{1}{1-B}$. Also, given u_{AL}, we can derive $u_{FS} = \dfrac{1}{1+\beta}u_{AL}$.

Our functional form, (1), is also equivalent to that of Charness-Rabin, if one removes the

[22] Indeed, other-regarding preferences have been recognized in economics for a long time. See, e.g., Ledyard (1971).

[23] Fehr-Schmidt actually allow their utility function to depend on the individual deviations of others from i's utility. In the VCM experiments, subjects have no knowledge of the distribution of the contributions of others. They only know the mean. For that reason, we consider only deviations of the average of the others from i's utility.

retaliatory term that adjusts for bad-behavior.[24] Their utility function is:

$$u = (\rho + \theta q)\bar{\pi}_{-i} + (1 - \rho - \theta q)\pi^i \text{ if } \pi^i \geq \bar{\pi}_{-i} \text{ and}$$
$$u = (\xi + \theta q)\bar{\pi}_{-i} + (1 - \xi - \theta q)\pi^i \text{ if } \pi^i \leq \bar{\pi}_{-i}. \tag{4}$$

where $q = -1$ if the others have misbehaved and $q = 0$ otherwise. This is the same as (3) if $b = \rho$, $a = -\xi$, and $\theta = 0$.

Finally, (1) is also equivalent to a linear version of ERC from Bolton-Ockenfels. The version of that used in Cooper-Stockman (2002) is $u^i = \pi^i - \alpha \max\{0, \bar{\pi} - \pi^i\}$. Let $\beta^i = 0$ in (1).

Given the equivalence of the varied functions, it is instructive to compare and contrast our restrictions $\beta \geq 0, \gamma \geq 0$ with those of Fehr-Schmidt and Charness-Rabin. We all pretty much agree on β and $b = \frac{N-1}{N}\frac{\beta}{1+\beta}$. Our assumption that $\beta \geq 0$ implies $0 \leq b \leq \frac{N-1}{N}$ which implies their restriction that $0 \leq b < 1$. But we differ in our restrictions on γ and $a = \frac{N-1}{N}\frac{\gamma-\beta}{1+\beta}$. Fehr-Schmidt focus on difference aversion with the restriction that $a \geq b$. Charness-Rabin consider several parameter constellations but focus on social welfare preferences with the restrictions that $a + b \geq 0$ and $-1/2 \leq a < 0$. Our restriction that $\gamma \geq 0$ is equivalent to $a + b \geq 0$ when $\beta \geq 0$, since $a + b = \frac{N-1}{N}\frac{\gamma}{1+\beta}$. However, differences arise over the rest. The restriction that $a \geq b$ is the same as $\gamma \geq 2\beta$ while the restriction that $a \geq -1/2$ is the same as $\gamma \geq (1/2)\left[\beta - \frac{N}{N-1}\right]$. As we will see, these latter two are a little tenuous when confronted with data from VCM experiments.

The linear multiplier converting Fehr-Schmidt and others to (1) and vice versa does depend on N. That is not a problem from a decision theoretic point of view, since N is never a choice variable of any agent. For example, the behavior generated by IEL for u_{AL} will be the same as that for the equivalent u_{FS}. However, the fact that N is involved in the transformation does raise some issues when one carries the functional form across VCM experiments with different size groups. An example of this can be seen in the, perhaps unintended, implication of the Fehr-Schmidt, Charnes-Rabin preference model that equilibria for the VCM are independent of N. It is easy to show that, in an equilibrium of their models, an agent will be selfish if $b \leq (1 - M)$, will be fair-minded if $a \geq M - 1$ and $b \geq (1 - M)$, and will be altruistic if $a \leq M - 1$ and $b \geq (1 - M)$. So, unless the distribution of types (a, b) is dependent on N, an assumption they never suggest, their models imply that the average rate of contributions in a VCM experiment should be independent of N. But this is not consistent with the data. For example, in the IW(1988) data for $M = 0.3$, contributions are significantly higher for $N = 10$ than for $N = 4$.

We believe that the principles of altruism and fairness that lie behind the Fehr-Schmidt and Charnes-Rabin other-regaring preferences, are sound and fundamental. However, because their functional forms and parameter restrictions conflict with the data from VCM

[24] Since we are modeling repeated games and, in that context, believe that retaliation is a strategic concept and not a basic preference concept, we prefer to leave retaliation out of the utility function.

experiments, we believe a new functional form is necessary. We propose (1) as the natural alternative.

4 Merging IEL and Other-regarding Preferences

We merge IEL and other-regarding preferences by making two assumptions. First, we assume that a subject will use the linear utility (1) in evaluating their foregone utility. We can write i's foregone utility as a function of i's strategy, c^i, and the average of the others' contributions, μ^i.

$$\begin{aligned} u^i(c^i|\mu^i) = & \left[(M-1) + \beta^i(M - \frac{1}{N}) - \gamma^{*i}\frac{N-1}{N}\right] c^i \\ & + \left[M + \beta^i(M - \frac{1}{N}) + \frac{\gamma^*}{N}\right](N-1)\mu^i \\ & + (1+\beta^i)w \end{aligned} \tag{5}$$

where

$$\gamma^{*i} = \left\{ \begin{array}{c} \gamma^i \\ 0 \end{array} \right\} \text{if} \left\{ \begin{array}{c} \bar{\pi} \geq \pi^i \\ otherwise \end{array} \right\}.$$

Let $r^i(c_t) = \mu_t^i$ where μ_t^i can be computed from $\hat{c}_t = \sum_j c_t^j$ because i also knows c_t^i and $\mu^i = \dfrac{\hat{c}_t - c^i}{N-1}$. It is this function, $u^i(a|r^i(c_t))$, we will use in the replication and selection stages of the IEL simulations.

Second, we assume something specific about the distribution of types, $F(\cdot)$. Some agents, P% of the population, are purely "selfish"; that is, they have the type (0,0). The rest, (1-P)% of the population, have other-regarding preferences where the (β^i, γ^i) are distributed identically, independently, and uniformly on $[0, B] \times [0, G]$.

We have no real justification for this particular functional form of F, other than the results in section 5. We began with just a uniform distribution on β and γ but that did not fit the data at all, since it generated too few subjects who were selfish in equilibrium. So we added the atom at (0,0). One would ultimately like to have a distribution that would hold up over a lot of experiments and not just VCM experiments, but this will clearly require a lot more data.

For the rest of this paper we will refer to the model that combines Individual Evolutionary Learning with other-regarding preferences as IELORP.

4.1 Determining (P, B, G)

The next step in our analysis is to provide some estimate of the subject population parameters (P, B, G). We do this by comparing our model to the data from Isaac-Walker (1988) described earlier.

Using IELORP with the IEL* parameters, we generated a number of simulated experiments varying the treatment (M, N) and the parameter triple (P,B,G). We used the same pairs of

(M, N) as in IW(1988). We varied P from 0.1 to 1, in increments of 0.1, and did a finer local search in the increments of 0.01. We varied B and G from 0 to 60, in increments of 1. We did not vary the IEL* parameters (J, ρ, σ) since, as we will show in section 5.2, such variation is really unnecessary.

For each treatment and for each parameter triple, we conducted 40 trials. Each trial involves a draw of a new type from $F(\cdot)$ for each agent. Those types were selected randomly as follows. Each agent became selfish with probability P. If his type turned out to be selfish, then his utility was based only on his own payoff. That is, $\beta^i = \gamma^i = 0$. If the agent did not become selfish, then we drew values of β^i and γ^i uniformly and independently from the ranges $[0, B]$, and $[0, G]$ respectively.

After running all of these simulations, we followed a standard approach to determining a good choice of (P, B, G). We choose the values that best fit the IW(1988) data. In doing so we did not want to "force" the fit to be too tight. That is, we wanted the fit to be not only good but also robust so that it has some chance of serving for other experiments. To determine a loose fit, we computed the average contribution over all agents and over all simulations for all of the ten periods, $\bar{c}^{10}_{IEL}$, as well as the average contribution over the last 3 periods, $\bar{c}^{3}_{IEL}$. We then computed the squared deviations of each of these measurements from the same measurements computed for the IW(1988) experimental data. The objective was to find the minimum of the sum of these squared deviations, MSE, i.e.:

$$\text{Min} \sum_{r=1}^{R} [(\bar{c}^{10}_{exp}(r) - \bar{c}^{10}_{IEL})^2(r) + (\bar{c}^{3}_{exp}(r) - \bar{c}^{3}_{IEL})^2(r)] \tag{6}$$

where r is a particular treatment for a given (N, M) and R is a total number of treatments. For the IW(1988) data, $R = 4$. Since we want to be able to compare these results to other experiments where R will be different, we normalize the MSE and report the *normalized mean squared error*, NMSE which is given as:

$$\text{NMSE} = \sqrt{\frac{\text{MSE}}{2R}}$$

In our comparison of IEL with other-regarding preferences to the IW(1988) data, the probability distribution of $(P, B, G) = (0.48, 22, 8)$ generates the lowest value of mean squared error equal to 0.9347. The normalized measure is given by $\sqrt{0.9347/8} = 0.341$. One can get an intuitive feeling for the goodness of fit by noting that the measurements, such as $\bar{c}^{10}_{IEL}$, used to compute the NMSE belong to $[0, 10] = [0, w]$. That is they are the average contributions for an endowment of 10 tokens. So a NMSE of 0.34 means that there is an average error of only 3.4% in our fit across the 4 treatments in IW(1988).

For the rest of this paper we will refer to the particular version of IELORP with parameters $(J, \mu, \sigma) = (100, 0.033, 1)$ and $(P, B, G) = (0.48, 22, 8)$ as IELORP*.

4.2 The quality of the fit

To provide a visualization of the high quality of the fit, in Figure 2 we present the average contribution in IELORP* simulations for $N = 4$, for both values of M. We also present the time paths of the experimental sessions for the same parameter values. In comparing the IELORP* and experimental averages, keep in mind that the experimental results are the average of 6 observations while the IELORP* results are the average of 40 observations. It is thus not surprising that the IELORP* curves are *smoother* than those from the experimental data. Figure 3 presents the paths for $N = 10$.

Further, to give some sense for how an individual simulation with IELORP* behaves, we provide in Figures 4, 5, and 6 the results of single IEL simulations, one with $(N, M) = (4, 0.3)$, one with $(N, M) = (4, 0.75)$, and one with $(N, M) = (10, 0.75)$. In Figure 4 one can see that there are 4 selfish individuals with average contributions converging to 0. In Figure 5, there are 2 selfish, 1 altruistic, and 1 fair-minded agents with average contributions converging to 35%. In Figure 6, there are 5 selfish, 3 altruistic, and 2 fair-minded agents with average contributions converging to 38%. These graphs are similar to individual data generated by the experiments. For example, they show the same zig-zag behavior (at least in the early rounds).[25]

Remember the commonly believed stylized facts that we listed in the Introduction. These are:

(1) Average contributions begin at around 50% of the total endowment and then decline with repetition, but not necessarily to zero.

(2) There is considerable variation in individual contributions in each repetition.

(3) Increases in M, the marginal value of the public good, lead to an increase in the average rate of contribution particularly in later repetitions. This is especially true for small groups.

(4) Increases in N lead to an increase in the the average rate of contribution particularly in later repetitions. This is especially true for small values of the marginal value of the public good.

Examination of Figures 2 - 5 provides a visual verification that IELORP* behaves according to all four.

4.3 Implications for the distribution of types

The parameters (P, B, G) determine the distribution of types which in turn determines the expected average equilibrium rate of contribution. In section 3, following equation (2) we saw that for any given draw of individual parameters the equilibrium rate of contribution is $\frac{N_a}{N_a + N_s}$. Therefore an approximation of the average equilibrium rate of contribution in equilibrium is $C^e = \frac{P_a}{P_a + P_s}$ where P_a is the probability that an agent will be altruistic (contribute w) and P_s is the probability an agent will be selfish (contribute 0). The probability an agent will be fair-minded is $P_f = 1 - P_s - P_a$. We can easily compute these probabilities.

[25]Of course the zigzag behavior in the IELORP simulations is due to a combination of learning and random selection, generating "errors", and not to signaling behavior which some suspect to be the cause of the spikes in the experimental observations.

Let $R = \dfrac{1-M}{M-\frac{1}{N}}$ and $S = \dfrac{N-1}{NM-1}$. Then from equation (2):

$$\begin{aligned} P_s = prob\{\beta \le R\} &= P + (1-P)(R/B) \\ P_a = prob\{\beta \ge S\gamma + R\} &= (1-P)(R - .5GS)/B \qquad \text{if } B - R - SG \le 0 \\ &= (1-P)(.5)(B-R)^2/BGS \qquad \text{if } B - R - SG \ge 0. \end{aligned}$$

In Table 4 we provide the values of these probabilities for the parameters $(P, B, G) = (0.48, 22, 8)$.

Table 4
Probability of Equilibrium Types
Selfish, Altruist, Fair-minded
P = 0.48, B = 22, G = 8

	N = 4	N = 10
M = 0.3	0.81, 0.01, 0.18	0.56, 0.11, 0.33
M = 0.75	0.49, 0.37, 0.14	0.49, 0.38, 0.13

The numbers in Table 4 highlight the fact that in the IELORP model, selfishness and altruism are relative concepts. Whether a particular subject is altruistic or selfish will often depend on the situation in which they find themselves – the number of participants and the marginal benefit of their choices. Higher M really brings out the altruism in an IELORP* agent, especially when there are few participants. Higher N also brings out altruism in an IELORP* agent especially when M is low. But, once N gets bigger than 10, the probability of selfishness stays fairly constant.

In Table 5, we provide a comparison between what the IELORP* model predicts about the expected rate of contributions in equilibrium and what the IW(1988) data say about the average of the last 3 periods of the experiments.[26] The first entry in each cell is the IELORP* prediction. The second entry is from the IW(1988) data. IELORP* is providing a good equilibrium prediction about the rate of contributions to which the IW(1988) experimental groups are converging. The average difference between the prediction of IELORP* and the data is 4.26% of the endowment, w.

Table 5
Expected and Actual Average Rate of Contribution (%)
IELORP* C^e**, IW(1988) final 3 periods**
P = 0.48, B = 22, G = 8

	N = 4	N = 10
M = 0.3	1, 4	17, 17
M = 0.75	43, 37	44, 36

[26] We are implicitly assuming that the average contributions in the last 3 periods is an estimate of where the group is converging to, the equilibrium of the group.

4.4 Implications for Fehr-Schmidt (1999)

One might wonder how our distribution on types differs from that in Fehr-Schmidt (1999). There it is suggested in Table 3 on p.843 that a discrete distribution on (a,b) where the $pr(a = 0) = .3, pr(a = .5) = .3, pr(a = 1) = .3, pr(a = 4) = .1, pr(b = 0) = .3, pr(b = .25) = .3$ and $pr(b = .6) = .4$ will explain much of the data. What would these imply for the probability of each type?

First, with $a \geq 0$, the probability of an altruistic type is 0. In the Fehr-Schmidt model, an agent is altruistic in their model iff $a \leq M - 1 < 0$. So if $a \geq 0$ there can be no altruistic type. This is not consistent with the data since there are those subjects who contribute most or all of their endowment even in round 10.

Second, for their proposed distribution on types, the probability of type does not depend on N. This is not consistent with the data for $N = 4$.

Third, it is true that for the Fehr-Schmidt distribution we should expect to see different dynamics as M varies. For M = 0.3 in their model an agent is selfish if and only if $b \leq 0.7$. So the probability of selfish behavior is 1. For M = 0.75, in their model, the probability of selfishness is 0.6 and the probability of fairness is 0.4. These imply higher rates of contributions in the earlier rounds. But without altruistic agents to pull up the contributions of the fair-minded agents, the equilibrium is that all contribute 0. In later periods, the average rate of contribution goes to 0 no matter what M is. This is also not consistent with the data.

5 Sensitivity

How excited should one be about the results above? There are number of challenges one can make. In this section and the next we look at two: sensitivity and transferability.

In this section, we ask how important the particular model parameters are - how sensitive the performance is to specific values. That is, if we change a parameter value, will the contributions generated change dramatically? If the answer is yes then one might worry that new data sets from new experimental situations might require new simulations to calibrate those parameters, rendering any particular results, including the ones above, as less interesting. On the other hand if there is a significant range of values of the parameters (J, ρ, σ) and (P, B, G) over which IELORP generates data consistent with the experimental data, then one should be a bit more excited.

There are two sets of parameters of interest. For IEL there are (J, ρ, σ) and for ORP there are (P, B, G). We will take up each in turn. But before we do so, it is interesting to note that the two sets will affect the simulated behavior, the choices of c_t^i, in fundamentally different ways. The population parameters (P, B, G) are crucial in determining the type of behavior any agent will adopt. The particular parameters that agents take into a session will affect the contribution levels to which a collection of agents will converge. The utility parameters will have relatively little effect on the rate of convergence. So the distribution of the ORP parameters should affect the average of the last three contributions more than the average of all ten. The IEL parameters (J, ρ, σ), on the other hand, have little to do with determining the equilibrium and very much to do with the rate and direction of learning. We would expect changes in the IEL parameters to affect the average of ten periods of contributions but not the averages over the last three.

5.1 Changes in ORP parameters.

We consider two types of changes in the parameters of the population type distribution function.

The particular choices of B and G are not crucial at all as long as they are paired appropriately. The results of our grid search in section 4.1 indicate that there is a wide range of the values of B and G for which the value of NMSE is equal or below 0.5. This is illustrated in figure 7 which plots NMSE as function of combinations of G (given on the x-axis) and B (given on the y-axis). The extensive blue colored valley is the region where NMSE reaches its minimal values below 0.5. A good approximate relation for the B and G that are at the bottom of this valley is $G = 0.5(B - 6)$ for $B > 6$. That is, if you pick any $B^* > 6$ and we then pick $G^* = 0.5(B^* - 6)$ and run simulations, drawing (β, γ) from the distribution of $(0.48, B^*, G^*)$, we will generate data with an NMSE of less that 0.5 from the IW(1988) data.

In Table 6 we list the values of average percentage of contribution, C^e, for values of $(P, B, G) = (0.48, B, .5(B - 6))$ where B takes on a variety of values. We do this for 8 combinations of (N, M). The key fact one should take away from this table is that, for a given (N, M), there is very little variation in C^e as we vary B. So the equilibrium changes little as we vary B and, thus, the NMSE relative to any data set will vary little as we change B.

Table 6

C^e for variations in B

	B	10	22	30	40	50	60	100
M=0.3	N=4	0	1	2	3	3	4	5
	N=10	16	17	17	18	18	18	19
M=0.75	N=4	45	43	42	42	41	41	41
	N=10	46	44	43	43	42	42	42

The NMSE between IELORP* and IW(1988) is also fairly insensitive to the choice of P as long as the choices of B and G are adjusted appropriately. As displayed in Table 7, for $P \in [0.1, 0.5]$, the NMSE is between 0.34 and 0.4 which are relatively low numbers. But it must be noted that the values of B and G that combine with P to give these low NMSEs change considerably as P changes. As P increases, implying a higher probability of purely selfish types, the minimizing ratio B/G also increases, implying a higher probability of altruists relative to fair types. Remember that the average equilibrium contribution is approximately $C^e = \dfrac{P_a}{P_a + P_s}$. If this is to remain constant, as it should to keep NMSE low, then $P_a = \dfrac{C^e}{1 - C^e} P_s$.

Table 7
Variations in P

P	NMSE	B	G
0.0	0.51	16	60
0.1	0.40	23	59
0.2	0.39	34	59
0.3	0.39	27	31
0.4	0.37	40	27
0.48	0.34	22	8
0.5	0.36	21	7

Finally, one can ask what happens if we use a distribution other than the uniform distribution. We considered the Beta distribution proposed by Janssen and Ahn (2006). They jointly estimated, for each individual, EWA parameters and the parameters of Charness-Rabin other regarding preferences. Based on these individual estimations, Janssen and Ahn propose a Beta distribution that best fits the distributions of the parameters ρ and ξ from equation (4). They suggest that $\rho \sim B(2, 0.75)$ and that $\xi \sim -9 + 10 * B(3, 0.5)$. We take their suggestions for the Beta distributions of ρ, call its cumulative $G^1(\rho)$, and ξ, call its cumulative $G^2(\xi)$, and convert these to distributions on β and γ. The distribution on β is then $F^1(\beta) = G^1\left(\frac{\beta}{1+\beta}\right)$. The distribution on γ is $F^2(\gamma|\beta) = G^2\left(\frac{\beta-\gamma}{1+\beta}\right)$. Note that the distribution on γ will no longer be independent of β.[27] We ran 40 sets of IELORP simulations using F^1 and F^2. The NMSE from the IW(1988)data for those data is 2.8, a much larger number than is desirable. The model with the Janssen-Ahn Beta distributions coupled with IELORP does not fit the data very well.

To try to attain a better fit of the Beta distribution, we added an atom of probability for selfishness as we did for the uniform. So when a type is to be drawn it is first decided, with probability P, whether a subject is just selfish. If yes, then $(\beta, \gamma) = (0, 0)$. If no, then (β, γ) is drawn from (F^1, F^2). We tried different values[28] of P and obtained the best fit for P=0.32. The value of NMSE that we get in this case is 0.47. This is fairly close to that attained by our best fit with IELORP*.

To illustrate how similar the behavior generated by IELORP* is for a wide range of population type distributions, we plot some results from 40 simulations for each in Figure 8 with $(N, M) = (10, 0.3)$. There are 3 curves varying (B, G) for $P = 0.48$. B runs from 8 to 58 with the appropriate G. There is virtually no difference. The NMSEs are 0.34, 0.42, and 0.44. There is one curve for the Beta distribution with $P = 0$. The NMSE for this is 2.8, much larger than the other NMSEs. But, there is not much difference in the average rate of contribution

[27] A more faithful application of their distributions would have $F^1 = G^1 \frac{N-1}{N} \frac{\beta}{1+\beta}$ and $F^2 = G^2 \frac{N-1}{N} \frac{\beta-\gamma}{1+\beta}$. But we wanted to preserve the independence from N and so ignored the effect of N.

[28] We did not explore the effect of changing the parameters of the Beta distributions.

until $t = 7$ at which point divergence occurs. This highlights the point that the probability distribution affects the point to which contributions converge. It also indicates that in some cases behavior with an NMSE of 2.8 is not that far off of behavior with an NMSE of 0.34. Finally, to illustrate robustness with respect to P, there is one curve for P=0.3 which again lies pretty much on top of the baseline model with $P = 0.48$. The NMSE for this is 0.37.

We conclude that there is a wide range of probability distributions that generate behavior close to that of economic experiments, although the range is certainly not arbitrary.

5.2 Changes in IEL parameters

How does IELORP perform when we change the IEL parameter values? How important were our choices of (J, ρ, σ)? We re-conducted the grid search for two values o f J, 50 and 200, and two values of the rate of experimentation, ρ, equal to 0.02 and 0.067. We also examined what happens to the NMSE values when we changed the standard deviation in the experimentation process. We tried a 10 times smaller, and a 2.5 times larger value of σ along with the values of P, B, and G that gave the best fit in our baseline model. The results of these simulations are displayed in Table 8.

Table 8
Variations in (J, μ, σ)

IEL parameters	B	G	NMSE
IELORP*	22	8	0.34
$J = 50$	16	5	0.35
$J = 200$	23	8	0.35
$\rho = 0.02$	20	7	0.34
$\rho = 0.067$	23	8	0.34
$\sigma = 0.1$	22	8	0.345
$\sigma = 2.5$	22	8	0.345

The results of these parameter variations are, to us, astounding. There is virtually no change in the normalized mean square error between the IELORP* model fit to the IW(1988) data and those from the variations. The values of B and G that minimize NMSE for each of the variations are very close to the values, $(22, 8)$, that minimize NMSE using our baseline model (see Table 8). In fact, the values of (B, G) in Table 5 satisfy $G = 0.5(B - 6)$, the relationship discussed in the previous section.

This means that as long as the parameters (J, ρ, σ) are in the set $[50, 200] \times [0.02, 0.67] \times [0.1, 0.25]$ we will get the same contribution behavior from the IELORP* model.[29] Thus if there is heterogeneity among subjects with respect to computational or memory capacity, the parameter J, or with respect to the rate and extent of experimentation, the parameters (ρ, σ), then as long as that heterogeneity is within the bounds above, it should have little effect on the closeness of the fit to the data.

[29]We did not explore the limits to which these sets can be pushed. It is highly likely that a much broader set of parameters will work.

The insensitivity of IEL to the specific values of (J, ρ, σ), as long as they are in a reasonable range, in not unique to VCM experiments. We have used IEL (with standard selfish preferences) in other repeated games. Notably we have used it to study call markets (in Arifovic and Ledyard (2007)) and to study Groves-Ledyard mechanisms for public good allocations (in Arifovic and Ledyard (2004, 2008)). The parameters for IEL that provided good comparisons to the data in these papers are essentially the same as those in Table 8. For call-markets (5 buyers and 5 sellers with unit demands or supplies), the parameters were $(J, \rho, \sigma) = (100, 0.033, 5)$.[30] For the GL mechanisms (5 participants choosing actions in $[-4, 6]$), the best parameters were $(500, 0.033, 1)$ but $(200, 0.033, 1)$ yielded results similar[31] to those for $J = 500$.

6 Transferability

A second type of challenge to the validity of IELORP* as a reasonable model of behavior asks what happens if we use it that to generate simulations to compare with data from other experiments.

6.1 Andreoni (1995)

To check the transferability of our model to other data sets, we began with Andreoni's (1995) experimental data. His *Regular* treatment corresponds closely to Isaac and Walker (1988) experimental design with two exceptions. First, he used only one treatment of $(N, M) = (5, 0.5)$. Second, in the Isaac-Walker, each subject was put into a group at the beginning of a session and remained in that group through the ten rounds. In Andreoni, twenty subjects were randomly rematched in each round into groups of size $N = 5$ with $M = 0.5$ In checking our model and parameter values against his data, we modified our IEL simulations to reflect the random matching treatment. Thus, in each simulation, we construct 20 IEL agents, and, in each period of the simulation, match them randomly into the groups of $N = 5$. We used his data set on average contributions per round (also 10 rounds) to do another grid search for the values of (P, B, G) that minimize the mean squared error. We found that $(0.39, 36, 20)$ resulted in the minimum NMSE of 0.41. This is a bit off of our values of $(0.48, 22, 8)$. But the difference is really slight. Using our values, IEL generates a NMSE of 0.49, virtually the same as the minimal value. In figure 9, we present the pattern of average contributions for Andreoni's data and for IEL simulations using our parameters $(0.48, 22, 8)$.

Andreoni (1995) also reports on another treatment called Rank. In this treatment, subjects are paid not on the value of π^i, but on the basis of where their π^i ranked with others. Andreoni uses a payoff of $r^i = R^i(\pi^1, ..., \pi^N)$ where R^i depends only on the relative rank of π^i. Note that $R^i(\pi) = R^i(w - c^1, ..., w - c^N)$. That is, it is only the selfish part of the payoff that is important. For this treatment, the foregone utility is computed as follows. Let c_t^i be the contribution of i from last time. Rank and renumber the c_t^j for $j \neq i$ so that $c_t^1 \geq c_t^2 \geq ... \geq c_t^{N-1}$. If $c_t^k \geq c^i \geq c_t^{k+1}$ then $R^i(c^i|c_t) = R_{k+1}$ and the foregone

[30]In fact, we used $\sigma = 1$ but the range of possible actions was $[0, 2]$ so normalizing gives $\sigma = 2$.

[31]We refer the reader to that paper for a more precise understanding of the nature of the similarity.

utility for i is $u(c|c_t) = R^i(c^i|c_t) - \gamma\{0, \bar{R} - R^i(c^i|c_t)\}$. Let us see how agents with utility of $u^i = \pi^i + \beta\bar{\pi} - \gamma\max\{0, \bar{\pi} - \pi^i\}$ would behave in the Rank experiment. Here, $u^i = R^i(w-c) + \beta\bar{R} - \gamma\max\{0, \bar{R} - R^i\}$. Note that $\bar{R}$ is known and constant over all choices of c. The analysis is easy if $\gamma = 0$. Here $u^i(c) > u^i(c')$ iff $w - c^i > w - c'^i$. The agent only cares about $w - c^i$, so Rank should produce IELORP results equivalent to those generated in section 2.4 with the selfish IEL where $\beta = \gamma = 0$. But when $\gamma > 0$ this becomes more interesting. Now $u^i(c) = R^i(w-c) - \gamma\max\{0, \bar{R} - R^i(c)$. If i's rank payoff is higher than average, then this is just like $\gamma = 0$. However, if i's rank payoff is lower than average, then i will have stronger incentives to push c^i to zero. So Rank should produce IEL results below the selfish IEL. As we did for Andreoni's Regular Treatment, in each simulation we generated 20 IEL agents, and randomly matched them into groups of 5 in each period of the simulation. A grid search over different values of (P, B, G) resulted in the minimum NMSE of 0.39 for $(0.31, 4, 12)$. Using our parameters, $(0.48, 22, 8)$, to generate the data yields a NMSE of 0.41 for Andreoni's rank data. We plot average contributions (over 40 simulations) for $(P, B, G) = (0.48, 20, 8)$ and compare it to the experimental data from Andreoni's Rank Treatment in figure 10.

So it seems that transferring IEL with ORP creates no problems. The data generated by IEL, using exactly the same parameters that we used in the IW(1988) simulations, produces data similar to that of the Andreoni experiments in both the Regular and Rank treatments.

6.2 Other Isaac-Walker small group experiments

Isaac and Walker have conducted a number of other VCM experiments beyond the IW(1988) on which we have focused so far. We list, in Table 9, a few of these treatments along with the IW(1988) and Andreoni's experiments we have already covered. For each of the lines in the table, we ran 40 simulations with IELORP*. In the far right column are the normalized mean square errors that arise when we compare the IELORP* averages for 10 periods and the final 3 periods with those from the data sets. Some of these are low, one is not. Let us examine each in turn to see whether or not they provide support to considering IELORP* as an appropriate behavioral model.

Table 9
NMSE Values for Small Group Experiments[32,]

	Data set	N	M	cash	exper.	obser.	NMSE
1	IW 88	4, 10	0.3, 0.75	yes	yes	6	0.34
2	Andreoni (reg)	5	0.5	yes	no	8	0.49
3	Andreoni (rank)	5	0.5	yes	no	8	0.41
4	IW 84	4, 10	0.3, 0.75	yes	yes	1	0.88
5	IW 84	4, 10,	0.3, 0.75	yes	no	1	2
6	IW 88	4	0.3	yes	no	6	0.66

[32]The initial endowment, w, varied across these experiments ranging from 10 to 60. We ignore this here for the reasons given in section 2.3.1. All sessions had 10 rounds.

Line 4 and 5 come from Isaac, Walker and Thomas (1984) which preceded Isaac and Walker (1988) in terms of treatments. They used the same values of M and N for 4 treatments. They also used both inexperienced and experienced subjects so they had a 2x2x2 design. They had only one session for each treatment. We are reluctant to make too much of observations that are a single data point but we include them for the record. For the treatments involving experienced subjects, line 4, we get an NMSE = 0.88 when we compare the experimental data to the IELORP* simulations. For the treatments with inexperienced subjects, line 5, the NMSE is 2.0. The main distinguishing feature of line 5 is that it involves inexperienced subjects. One might therefore be tempted to argue that inexperience is a condition important in experiments that is not included in the IELORP* model. However, as we will see in the next paragraph, that is not supported in the IW(1988) experiments with inexperienced subjects. We believe the NMSE of 2 is an artifact of the small sample size of 1. It is not surprising that a single session might deviate from that average generated by IELORP*, even if the average of a larger number of experiment sessions would not. We believe that more experiments with inexperienced subjects, on these parameters for (N, M), are needed to reject IELORP*.

Line 6 comes from Isaac and Walker (1988). As they were providing experience for their subject pool for the IW(1988) experiments, Isaac and Walker conducted six sessions with $M = 0.3$, $N = 4$ and inexperienced subjects. When we compute the NMSE for the difference between the average contributions from these data and those from IELORP*, we get 0.66.

The reasonably low value of NMSE for line 6 suggests to us that the one observation from IW84 mentioned in the previous paragraph is indeed an outlier. More sessions could generate average contribution rates closer to IELORP*. Therefore, inexperience is not something that we need to control for in our IELORP* model.

7 Conclusions

Learning alone does not explain the VCM data. Other-regarding preferences alone do not explain VCM data. But combining them in an appropriate way does the job. We have merged our learning model, IEL, with a modification of standard other-regarding preferences, ORP. We then provided a distribution function over preference types. The resulting model, IELORP* is parsimonious and generates behavior similar to that in many VCM experiments. Further, the performance of the model does not change much over a wide range of values for its parameters. We have provided a robust explanation for the four stylized facts from VCM experiments.

The major finding is that our model generates behavior that is similar to that generated by many VCM experiments. It is easy to verify that IELORP* generates behavior that satisfies the following stylized facts.

(1) Average contributions begin at around 50% of the total endowment and then decline with repetition, but not necessarily to zero.

(2) There is considerable variation in individual contributions in each repetition.

(3) Increases in M, the marginal value of the contribution to the public good, lead to an increase in the average rate of contribution particularly in later repetitions. This is especially true for small groups.

(4) Increases in N lead to an increase in the the average rate of contribution particularly in later repetitions. This is especially true for small values of the marginal value of contribution to the public good.

The experiments we consider include two run by Isaac-Walker at different locations and two run by Andreoni. Each of these experiments had 6-8 observations.[33] Using a normalized mean squared error measure (NMSE), we find that the average contributions generated by IELORP differ from those in these experiments by only 3.4% to 6.6%. We find this to be compelling evidence in favor of our model. This is especially true when one notes that the three parameter values we use for IEL in the VCM are essentially the same values we successfully used for IEL in vastly different scenarios: a call-market with private goods in Arifovic and Ledyard (2007) and the Groves-Ledyard mechanism with public goods in Arifovic and Ledyard (2008).

A secondary finding is that our model is robust. It relies on a very small number of parameters. IEL has only 3 parameters: the size of the remembered set and the rate and range of the experimentation. ORP has only 3 parameters: those that determine the probability distribution on types. Further, the precise values of those parameters do not matter very much. We show that the NMSE between IELORP and experimental data changes very little as we change the 6 parameters over a fairly wide range. This suggests that if we want to transfer IELORP to other VCM experiments or to other experiments like call markets in a private goods world, there will be little need to re-calibrate the parameters.

There are some open issues that remain to be dealt with in future research. One involves the linear nature of the ORP model. It may be that the individual altruism parameter should depend on N. As we have modeled it, each agent cares only about the average profit of all agents and not the total. It is, however, likely that agents care about the total payoff but with diminishing marginal utility as N increases.

A second issue involves the myopia of the IEL model. As it stands, the retained strategy set does not consider sequences of moves, it only considers the next move. This means, in particular, that teaching - moving now to encourage future behavior on the part of others - is not a considered option by IEL agents. Spikes in contributions designed to encourage others to contribute would be one example of such behavior. This did not seem that important in our results for VCM. But it might be really important in more serious coordination games.

A third issue involves the transferability of the IELORP* model across widely varying mechanisms and environments. IELORP assumes subjects come to the lab with their other-regarding preferences built in. That is, it is assumed that nothing in the experiment triggers the altruistic or fair-minded behavior. If this is true then the IELORP* model should work[34] in environments in which economists typically don't consider other-regarding behavior. For example, IELORP* should work in both the call market and Groves-Ledyard environments we had analyzed earlier with IEL without ORP. If individuals really do have other-regarding preferences, then the fact that we don't seem to need those preferences to successfully model

[33]There are two other Isaac-Walker experiments with only 1 observation each for which the differences between IELORP and data were 9% and 20%. We believe a bigger sample size is needed before we take these differences seriously.

[34]By work we mean it should generate behavior similar to that in the experiments without significant calibration or changes in the model parameters.

behavior in, say, continuous markets should mean that there is no observable difference in the behavior of those with and those without a preference for altruism.[35] This needs to be researched more carefully.

Finally, there are some features of experimental designs that IEL is currently insensitive to.[36] (i) There is no provision for communication which we know can significantly increase contributions. (ii) There is no role for the matching protocols. In IELORP*, agents process past observations the same way whether they come from individuals with whom they will be re-matched or not. (iii) There is no provision for experience. It did appear that experience mattered in some of the Isaac-Walker experiments. (iv) In IELORP* we also treated all subjects the same with respect to IEL. That is, we made no provision for subject pool differences in computational or strategic skills that could affect the parameters J and μ. One reason this was not necessary was the insensitivity of the performance of IELORP in VCMs to the values of those parameters. But in more complex experiments this could be very important.

However, even without this future research to provide refinements, IELORP* has proven to be a parsimonious, robust model of behavior that requires little calibration to generate observations similar to those of experimental subjects.

[35]See Kucuksenel (2008) where he shows that, for example, in bi-lateral trading with asymmetric information, the existence of altruism leads to increased trading and efficiency.

[36]See Zelmer (2003) for a list of those features that seem to be important from the experimental evidence.

8 References

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

Andreoni, J. and J. Miller, 1991, Can Evolutionary Dynamics Explain Free Riding in Experiments?, Economic Letters, 36, 9-15.

Arifovic, J. and J. Ledyard, 2004. Scaling Up Learning Models in Public Good Games, *Journal of Public Economic Theory* ,6,2004, 205-238.

Arifovic, J. and J. Ledyard, 2007. Call Market Book Information and Efficiency, *Journal of Economic Dynamics and Control*, 31, 1971-2000.

Arifovic, J. and J. Ledyard, 2008. A Behavioral Model for Mechanism Design: Individual Evolutionary Learning, manuscript, Caltech.

Bolton, G. and A. Ockenfels, 2000. ERC: A Theory of Equity, Reciprocity, and Competition, American Economic Review, 90(1) 166-193

Camerer, C., and T. Ho, 1999. Experience-Weighted Attraction Learning in Normal Form Games, *Econometrica* 67: 82-873.

Charness, G. and M. Rabin, 2002. Understanding Social Preferences with Simple Tests, The Quarterly Journal of Economics, 117, 3, 817-869.

Cooper, D. and C.K. Stockman, 2002. Fairness and Learning: An Experimental Examination, Games and Economic Behavior, 41, 26-45.

Cox, J. C., D. Friedman, and S. Gjerstad, 2007. A Tractable Model of Reciprocity and Fairness, Games and Economic Behavior, 59, 1, 17-45.

Cox, J. C., D. Friedman, and V. Sadiraj 2008. Revealed Altruism, Econometrica, 76, 1, 31-69.

Cox, J. and V. Sadiraj, 2007. On Modeling Voluntary Contributions to Public Goods, Public Finance Review, 35, 2, 311-332.

Croson, R., 2007. Differentiating Altruism and Reciprocity, in the *Handbook of Experimental Economics Results*, C. Plott and V. Smith, eds., forthcoming.

Fehr, E. and K. Schmidt, 1999. Fairness and Retaliation: The Economics of Reciprocity, *Quarterly Journal of Economics*, 114, 817-868.

Holt, C. and S. Laury, 2002. Theoretical Explanations of Treatment Effects in Voluntary Contributions Experiments, forthcoming in the *Handbook of Experimental Economics Results*, C. Plott and V. Smith, eds., forthcoming.

Isaac, R. M., J. M. Walker, and S. Thomas, 1984. Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations, Public Choice, 43, 113-149.

Isaac, R.M.,K. McCue, and C. Plott, 1985. Public Goods Provision in an Experimental Environment, *Journal of Public Economics* 26, 51-74.

Isaac, R.,M. and J.M. Walker, 1988. Group Size Effects in Public Goods Provision: The Voluntary Contribution Mechanism, *Quarterly Journal of Economics*, 103, 179-199.

Isaac, R. M., J. M. Walker, and A. Williams, 1994. Group Size and the Voluntary Provision of Public Goods: Experimental Evidence Utilizing Large Groups, Journal of Public Economics, 54, 1-36.

Janssen, M. A., and T. K. Ahn, 2006. Learning, Signaling, and Social Preferences in Public Good Games. Ecology and Society 11(2): 21. URL: http://www.ecologyandsociety.org/vol11/iss2/art21

Kim, O., and M. Walker, 1984. The Free Rider Problem: Experimental Evidence, *Public Choice*, 43, 3-24.

Kucuksenel, S. 2008. Behavioral Mechanism Design, manuscript, Caltech.

Ledyard, J., 1971. A Convergent Pareto-Satisfactory Non-Tatonnement Adjustment Process for a Class of Unselfish Exchange Environments.. *Econometrica*, 39, No. 3, 467-499.

Ledyard, J., 1995. Public Goods: A Survey of Experimental Research, in The Handbook of Experimental Economics, eds J. Kagel and A. Roth, Princeton University Press

Marwell, G. and R. Ames, 1979. Experiments on the Provision of Public Goods 1: Resources, Interest, Group Size, and the Free-rider Problem. *American Journal of Sociology*, 84: 1335-60.

Roth, A. and I. Erev, 1995, Learning in Extensive Form Game: Experimental Data and Simple Dynamic Model in the Intermediate Term, *Games and Economic Behavior*, 8, 164-212.

Wendel, S., and J. Oppenheimer, 2007. An Analysis of Context-Dependent Preferences in Voluntary Contribution Games with Agent-Based Modeling, manuscript, University of Maryland.

Zelmer, J., 2003. Linear Public Goods Experiments: A Meta-Analysis , Experimental Economics, 6:299-310

9 Appendix: Asymmetric Endowments

If endowments are asymmetric, then the equilibrium is qualitatively the same as that with symmetric endowments. That is,

$$c^i = \left\{ \begin{array}{l} 0 \\ MAX\{0, w^i - \bar{w}^{-i} + \mu^i\} \\ w^i \end{array} \right\} \text{if} \left\{ \begin{array}{l} 0 \geq \left[(M - \frac{1}{N})\beta^i + M - 1\right] \\ \gamma^i(\frac{N-1}{N}) \geq \left[(M - \frac{1}{N})\beta^i + M - 1\right] \geq 0 \\ \gamma^i(\frac{N-1}{N}) \leq \left[(M - \frac{1}{N})\beta^i + M - 1\right] \end{array} \right\}.$$

With asymmetric endowments, for fair types in equilibrium, $c^i = max\{0, w^i - \bar{w} + \mu\}$ where μ is the average contribution across all i. Let N_a be the number of altruists, and N_f^+ be the number of fair-minded types for whom $c^i > 0$. Then in equilibrium, $\sum c^i = N_a\bar{w}_a + \sum_{i \in N_f^+}(w^i - \bar{w} + \mu)$. Therefore $N\mu = N_a\bar{w}_a + \sum_{i \in N_f^+}(w^i - \bar{w} + \mu) = N_a\bar{w} + N_f^+(\bar{w}_f^+ - \bar{w}) + N_f^+\mu$.

It follows that $\mu = \frac{N_a\bar{w}_a + N_f^+(\bar{w}_f^+ - \bar{w})}{(N - N_f^+)}$, where N_a is the number of altruists, $\bar{w}_a$ is the average endowment of altruists, N_f^+ is the number of fair minded agents contributing positive amounts in equilibrium, $\bar{w}_f^+$ is their average endowment, and $\bar{w}$ is the average endowment over all agents.

10 Figures

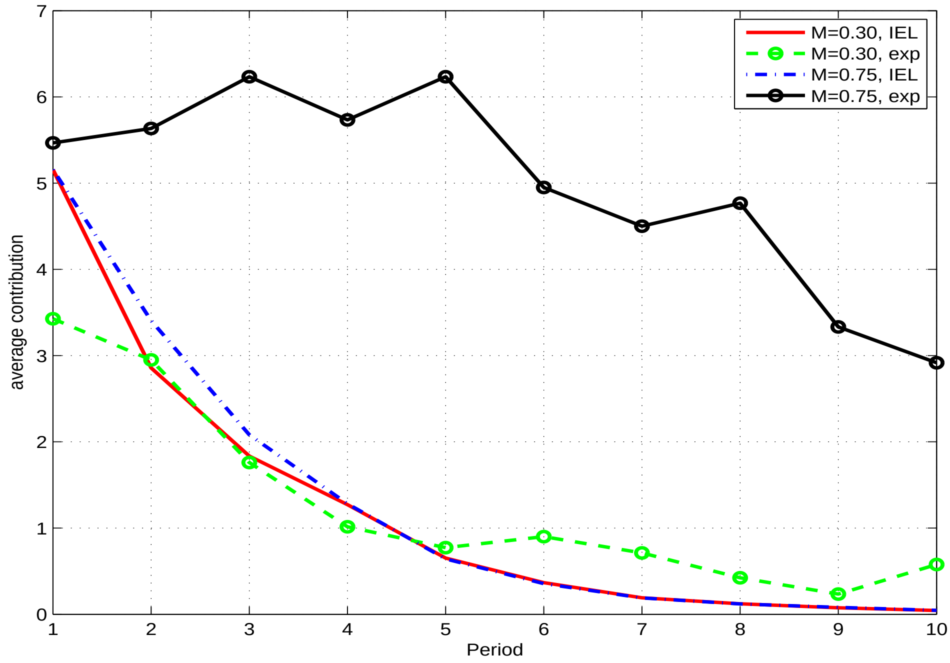

Figure 1: Simple VCM IEL
Exp. Data Source: IW(1988)

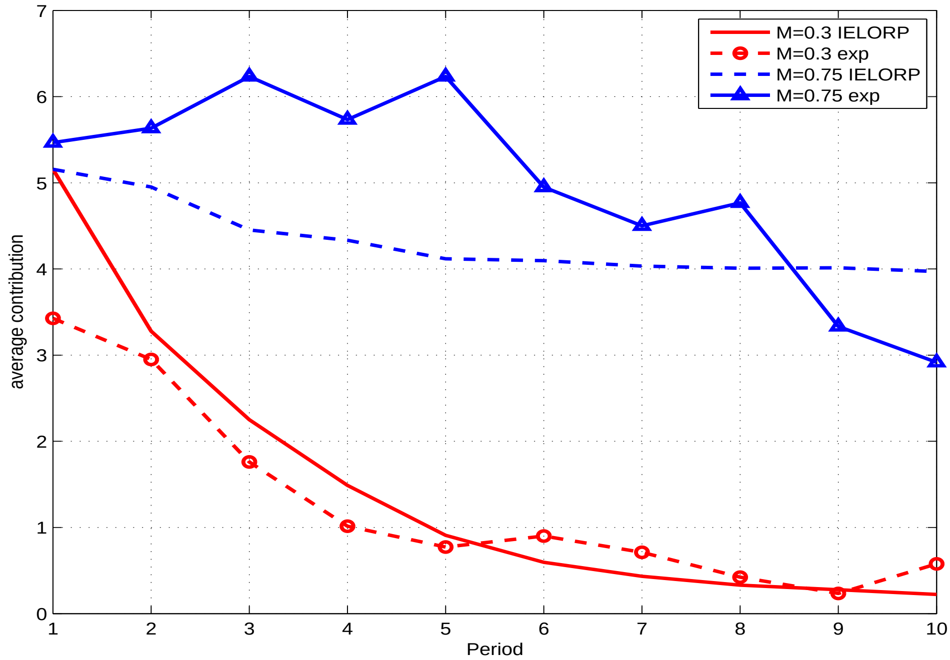

Figure 2: P=0.48, B=22, G=8, N=4.
Exp. Data Source: IW(1988)

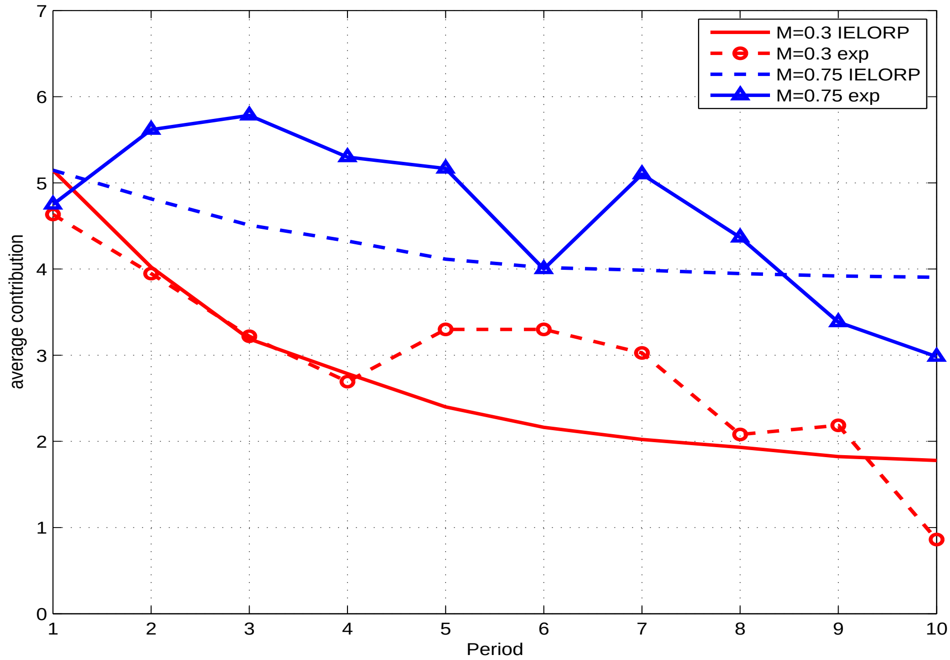

Figure 3: P=0.48, B=22, G=8, N=10.
Exp. Data Source: IW(1988)

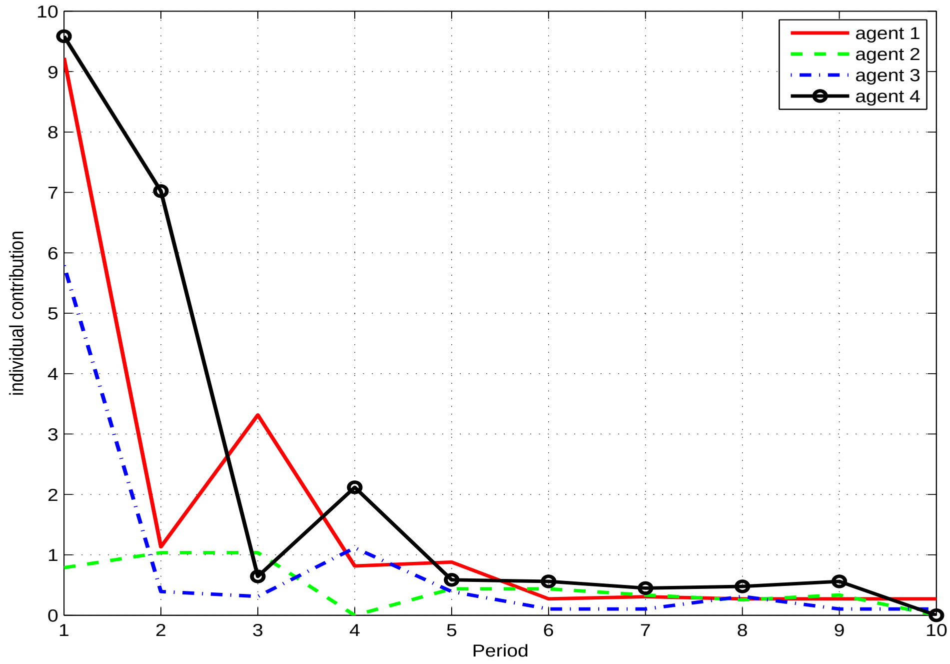

Figure 4: Individual contributions from IELORP*, N=4, M=0.3

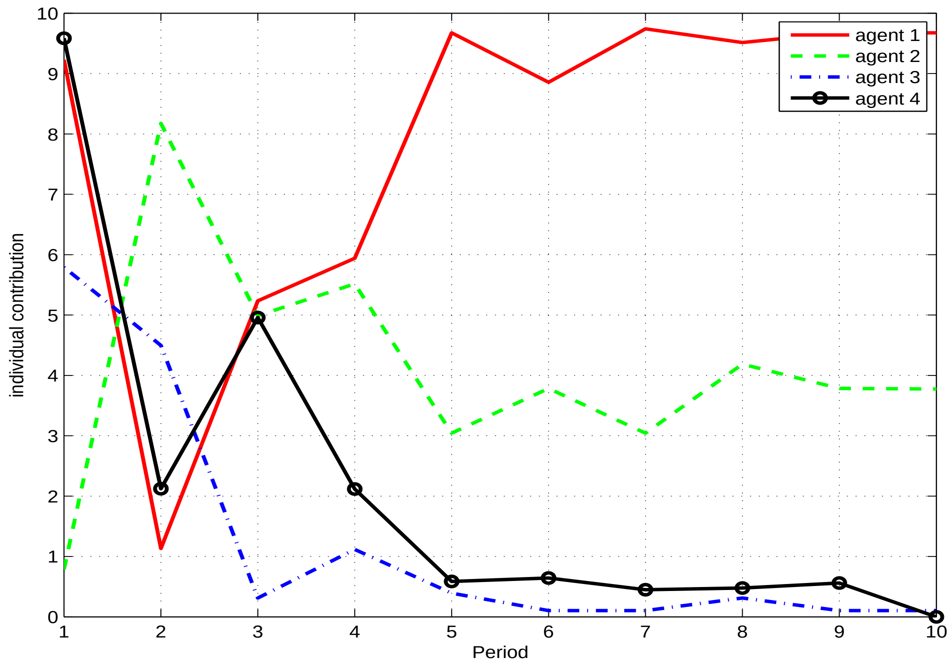

Figure 5: Individual contributions from IELORP*, N=4, M=0.75

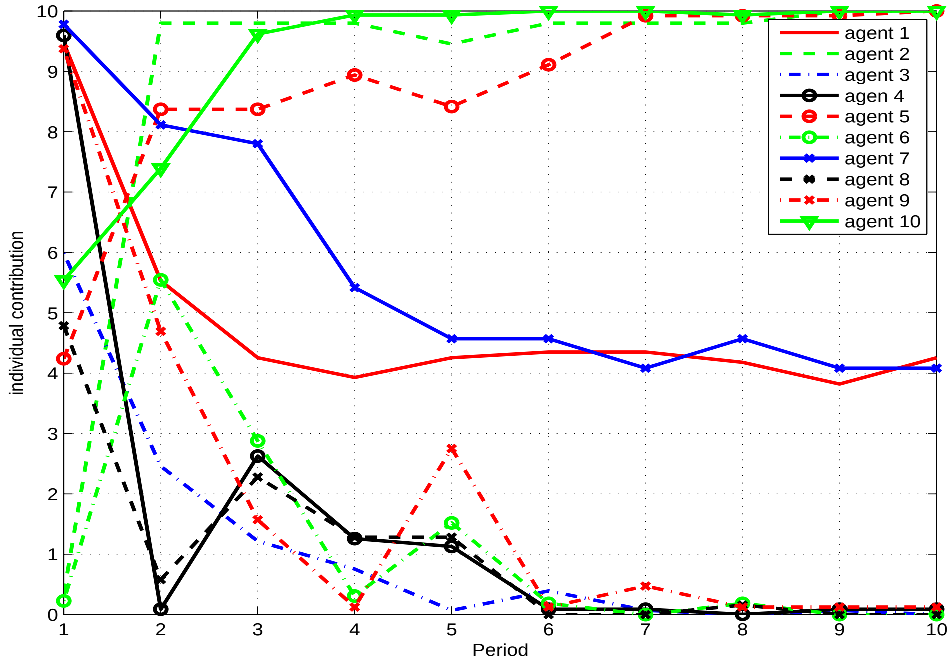

Figure 6: Individual contributions from IELORP*, N=10, M=0.75

Figure 7: NMSE landscape for P=0.48

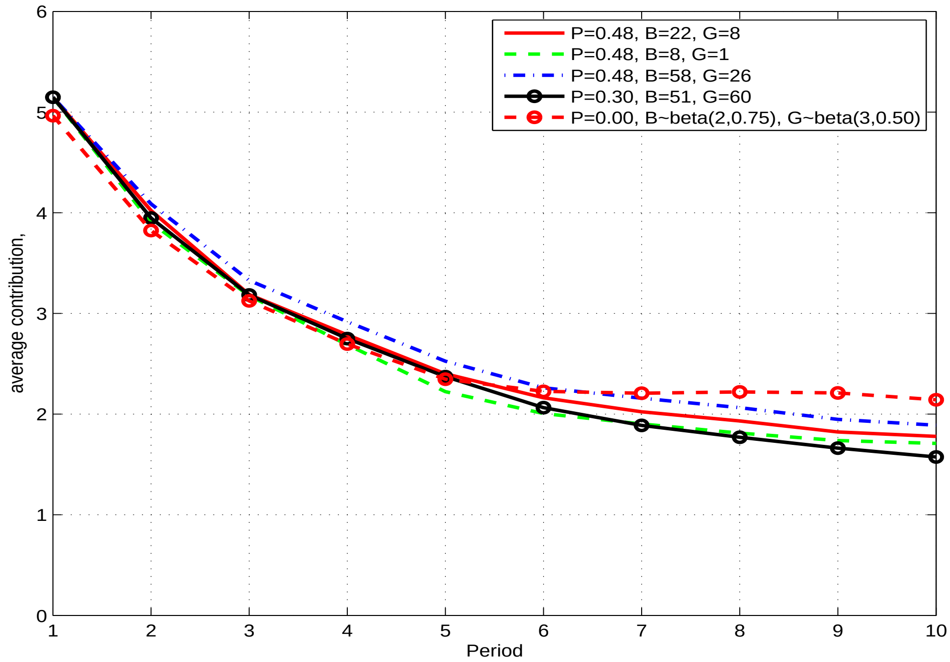

Figure 8: Average contributions for different parameter sets Using IELORP

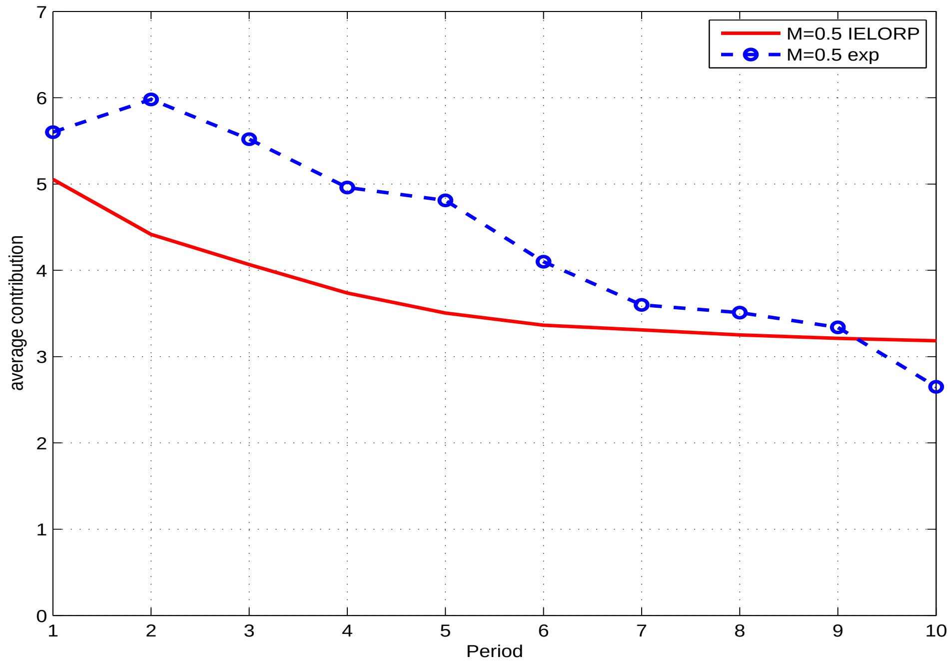

Figure 9: Average contributions, N=5 (total of 20 subjects) M=0.5 Exp. Data Source: Andreoni, 1995, regular data

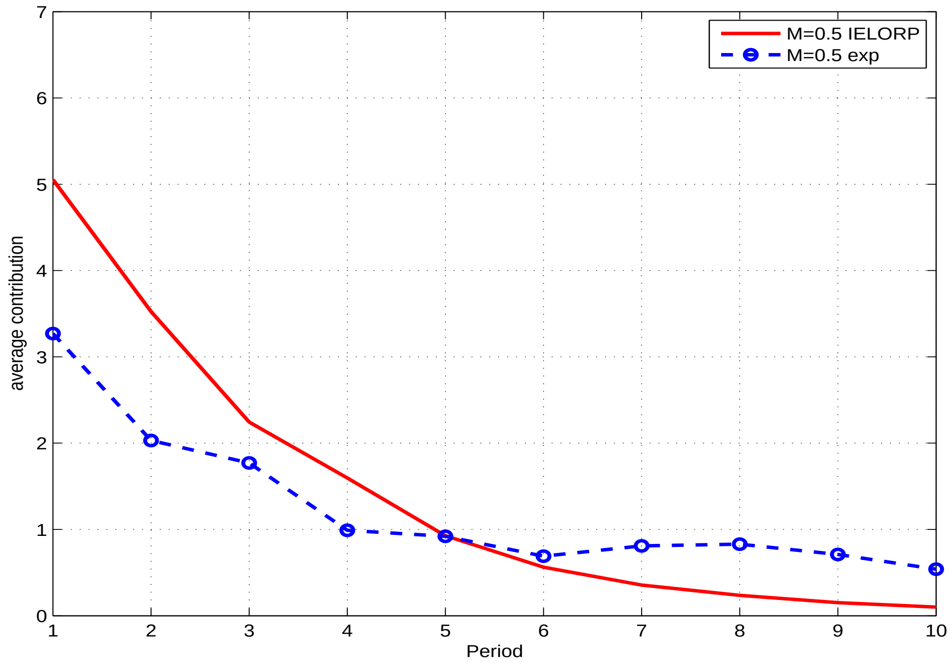

Figure 10: Average contributions, N=5 (total of 20 subjects) M=0.5
Exp. Data Source: Andreoni, 1995, rank data