An Analysis of Context-Dependent Preferences in Voluntary Contribution Games with Agent-Based Modeling

Stephen Wendel & Joe Oppenheimer Both of Department of Government and Politics University of Maryland College Park, MD 20742 USA swendel@mail.umd.edu joppenheimer@gvpt.umd.edu

For a more up to date presentation of our thinking on this project, see Paradox Lost at http://www.bsos.umd.edu/gvpt/oppenheimer/research/paradoxlost.pdf

Abstract

In “Skating on Thin Ice,” Frohlich and Oppenheimer (2006) describe a phenomenon they observed in laboratory experiments on the production of public goods that is rarely discussed in the literature. They report individual contributions to the public good are often inconsistent over time, appearing to fluctuate between two distinct contribution levels. Although they conjecture that individuals have complex context-dependent preferences, they did not develop a full specification of the theory.

We develop an agent-based simulation of these conjectures, provide a possible specification of a theory of complex context-dependent preferences, and demonstrate how this theory can, in fact, generate the pattern of contributions observed by Frohlich and Oppenheimer. We then conduct sensitivity analyses, examining the behavior generated by fourteen scenarios. Two main theories are considered: that inconsistent contributions arise either from a deterministic avoidance of exploitation or from a probabilistic response to exploitation. The former theory clearly fails, the latter theory, under certain conditions, does produce the observed pattern of contributions. Two simple alternative theories are also considered, that of a highly-stylized “probabilistic guilt” and of goal-oriented but non-utility maximizing behavior (with stable preferences). Both alternatives, under certain conditions, are also able to generate the observed pattern. We develop an analysis of situations in which the predictions of these theories diverge and suggest that one could discriminate between them in laboratory settings. Finally, we consider a possibly fruitful relationship between simulation and experimentation to consider the implications of one’s models and conjectures.

1 Thanks are due to Norman Frohlich for his detailed suggestions as we redrafted.
# Table of Contents

Abstract ................................................................................................................................................................i
Table of Contents ..............................................................................................................................................ii
Introduction ........................................................................................................................................................1
   The Model ......................................................................................................................................................3
   The Environment .....................................................................................................................................3
   The Preferences ........................................................................................................................................3
Avoiding Exploitation ..................................................................................................................................4
Two Models of Context Dependent Behavior .................................................................................................5
   Avoiding Exploitation with a Continuous Other-Regarding Utility Function .......................................5
   Responding to Exploitation with a Probabilistic Other-Regarding Utility Function ..............................6
Methodology .......................................................................................................................................................8
   Design of the Simulation Model .................................................................................................................8
   Execution of the Model ...............................................................................................................................9
   Initial Scenarios Considered ......................................................................................................................10
Results .............................................................................................................................................................11
   Benchmark Models ................................................................................................................................11
      Summary ..............................................................................................................................................11
   Sensitivity Tests ......................................................................................................................................14
      Alternative Parameters with the Benchmark Models ...........................................................................14
      Mixing Elements of the Benchmark Models ......................................................................................16
   Alternative Models of Jagged Contributions ..........................................................................................17
Analysis ...........................................................................................................................................................19
   Review of Simulation Results ..................................................................................................................19
   What is Causing the Simulation Results? .................................................................................................19
   Insight into Real World Behavior ............................................................................................................20
Conclusion .........................................................................................................................................................21
References .........................................................................................................................................................22
Introduction

In “Skating on Thin Ice”, Frohlich and Oppenheimer (2006) (F&O) describe a phenomenon they observed in a voluntary contribution mechanism (VCM) experiment concerning public goods. In VCM experiments, any amount of contribution contradicts pure rational choice theory, an issue which has been extensively studied (Ledyard, 1995). F&O observe that not only are an individual’s contributions not zero, but the individual’s contributions to the public good fluctuate (often between two distinct contribution levels) and hence appear to be inconsistent over time. They report the finding as if it is widespread though rarely discussed in the literature: This pattern of contributions, which they refer to as “jagged contributions,” is illustrated in Figure 1. They conjectured context-dependent preferences might explain the observation, but do not develop a full specification of the theory to explain the observation.

The field of experimental economics has a long history of testing the predictions of economic theory in the laboratory (e.g., Smith, 1994). In particular, critics of rational choice theory in psychology and economics have often used experimental studies to demonstrate empirical problems with some of the key assumptions. Kahneman and Tversky’s (1979) work on prospect theory, Fehr and Schmidt’s (1999) work on inequality aversion, and Laibson’s (1997) work on discounting are a few examples.

In their piece, F&O develop a wide ranging discussion of how preferences might be formulated to take into account these and other empirical findings. Traditionally, the formulations of preferences by economists have included assumptions of stability, uniqueness, self-interest and continuity. All of these presumptions are taken to task in their essay. F&O argue that individuals must have complex “context-dependent preferences,” where the context the individual believes herself to be in shapes her preferences and thus the decisions she makes. They advocate an expanded understanding of individual utility and preferences, including self-interested monetary payoffs, altruistic behavior, and context-dependent modifiers.²

² This theory is distinct from Tversky and Simonson’s (1993) model of “context-dependence preferences,” which illustrates how a background context of options that are infeasible in the given environment may still shape individual choices. Tversky and Simonson question rational choice theory’s assumption of independence of irrelevant alternatives; F&O question the assumptions of a (solely) self-interested utility function and of stable preferences.
Self-interest has been suspect for considerable time, and experimentalists are currently fully engaged in tests, speculations, and theories regarding how to deal with the problems of non-self-interested behavior. Traditionalists such as Elizabeth Hoffman (Hoffman et al., 1996) have argued that such behavior is only manifest when third party or publicity effects are brought to the fore. Others, such as Fehr and Schmidt (1999), Frohlich and Oppenheimer (1984), and Cox et al. (2001) have argued that other-regarding behavior is an essential element of motivation. Explaining the act of giving in prisoner dilemma games without resort to the considerable indeterminacies of the folk theorem is quite straight forward with an element of other-regarding behavior, and we continue that tradition.

Two different sorts of moves have been made in the modeling of other-regarding behavior. The initial work was constructed by making preferences broader, to include the welfare of others, but with no other strong presumptions. The preference functions remained stable, twice differentiable, and expressible in traditional terms (Valavanis, 1958; Frohlich, 1974). But experiments led to observations of conditionalized responses and the theory was to follow with more complicated formulations of preferences (see Rabin, 1993; Cain, 1998; Fehr and Schmidt, 1998; and Frohlich et al., 2004). In this literature, an individual’s behavior was conditionalized on that of others. Context dependent behavior here was modeled as a form of adaptation: The individual maximizes in a fashion conditionalized to the environment they find themselves in. But often, the basic functions utilized in theories of other regarding behavior remained continuous and twice differentiable.

However, in some of these formulations, discontinuities (e.g. Fehr and Schmidt, 1999), are formulated via ‘dummy variables,’ making the function ‘context dependent.’ Behavior could then be expected to exhibit ‘instabilities’ as changes in context occurred. Such complications reflected the data in question that shows not only giving, but also considerable persistent instability at the individual level. How preferences relate to such behavior then was the focus of F&O. They argued on the basis of cognitive processes for less orderly structures of conditionalized instability.

In this essay we sketch a few alternative formulations of preferences that could support the observed micro-behavioral patterns. We build upon F&O by constructing an agent-based model of the author’s VCM experiments and use them to analyze the possibility of alternative theories of context-dependent preferences. Specifically, this study has three goals:

1. To specify fully a model of context-dependent preferences and test whether it can generate the phenomenon of jagged contributions observed in F&O’s VCM experiments;
2. To explore plausible alternative hypotheses for jagged contributions in order to gauge the degree to which the theory of context-dependent preferences has unique explanatory power in this circumstance; and
3. To discover conditions under which the predictions of context-dependent preferences and alternative hypotheses diverge from one another, providing scenarios under which one can discriminate between the theories and falsify invalid ones.

In sum, we attempt to induce jagged contributions in simulated agents participating in a computerized VCM environment mirroring behavior in F&O experiments. This allows us to flesh out the theoretical conjectures posed in F&O and use the agent based environment to consider both the implications and alternatives of the conjectures.
The Model

The Environment

The experimental data discussed in F&O stemmed from a typical symmetric VCM experiment. The size of the group was 5, and each individual had an endowment of 10. The individual could hold on to any proportion of her endowment or contribute it to the public good. The public good had a 40% rate of return for each individual. (See Table 1 for details.)

The Preferences

In keeping with the conjectures in F&O, the basic model of behavior can be said to be a combination of self-interest and other-regarding motives. Given that each unit contributed generates .4 units of benefits for each person, the other-regarding motives must be a function of those benefits. The benefits from \( i \)'s donation (call that \( x_i \)) for the 4 others in the group then can be specified as \( 4 \times 0.4 \times x_i \) or \( 1.6x_i \). The weights being assigned to other-regarding motives relative to the self-interest motives will depend upon the behavior of others in the environment that the individual is in. More specifically, in any round, \( t \), the behavior of others in the previous round, \( t-1 \), and conceivably in earlier rounds, will affect the individual’s behavior.

We assume that the weight the player places on other-regarding motives will be dependent upon two factors. First, each individual has a default preference for benefiting the group, \( \alpha_i \). Second, this preference is conditional on a response function regarding the behavior of other players. If the given player believes she is being exploited, she may withdraw support for others.4

To keep things straight, we write the motivations more formally. The normal individual, \( i \), with an endowment of 10, making a decision at time \( t \) within the context of the VCM game has 3 purely self-interested components, and an other-regarding component which is conditional on the player’s experiences.

The Self-interested Components:

1. She gets to keep everything that she doesn’t contribute to the public good (10 - \( x_{i,t} \)).
2. She gets a 40% return on anything she donates in any round (\( x_{i,t} \))
3. She gets a 40% ‘bonus’ from anything contributed by others (\( \Sigma x_{j,t} \)).

Other-regarding Motivations:

---

3 In multiple rounds, we will refer to \( x_{i,t} \) as \( i \)'s donation in the \( t \)th round.
4 A third element, the player’s desire to signal to others in the group regarding her disposition to contribute in future rounds (Lohmann 1994, 2000), could obviously also be considered and factored in, but we do not do so in this initial paper.
1. She may have a motive to help the group. We denote the weight on other-regarding motives as $\alpha_i$. This will factor in to her payoff via some function $f_i$.

2. Her willingness to help the group will be conditional on her previous contributions relative to the previous contributions of other players (the typical other shall be referred to as $j$), making $f_{i,t}$ a function of $x_{i,t-1}, x_{j,t}$.

These aspects of the motivation for any individual, $i$, can then be written as:

\[
U_{i,t} (x_{i,t} \mid x_{i,t-1}, x_{j,t-1}) = (10 - x_{i,t}) + 4(\sum x_{j,t} + x_{i,t}) + f_{i,t}(x_{i,t} \mid \alpha_i, x_{i,t-1}, x_{j,t-1})
\]

**Equation 1:** Basic utility function

This may be thought of as the basic equation or objective function being maximized by the individual, $i$. In the formulation in Equation 1 the result could be a smooth, twice differentiable function or it could be discontinuous, depending on the form of $f_i$. We will consider both those formulations that generate unique and stable utility functions as well as those which generate discontinuities reflecting contextual changes. But our next step is to illuminate what might affect $f_{i,t}$, the roles of both $\alpha_i$ and $x_{j,t}$, and to consider the operationalization of these relations imposed by the particular VCM being modeled.

### Avoiding Exploitation

The sense of “being exploited” is a complex construct and we will present a stylized version. We assume that her other-regarding behavior, at whatever level it may be by default, is contingent on her avoiding exploitation. If she finds herself giving, and being exploited, she may decrease her other-regarding motivation by some factor, we call $r$. Then one way that her effort to avoid exploitation could be formulated would be that $\alpha$ is modified by $r$ multiplicatively as in $\alpha(1 - r)$. To do this we elaborate on the aspect of the arguments we referred to as $f_i$ in Equation 1.

The structure of the institutions affects how this sense of being exploited is manifest. For example, in F&O the full vector of giving by others in the group ($j$) is neither known nor knowable. So $i$, who is concerned about relative effort can only consider her behavior relative to the mean. In this case, $i$ might be motivated by whether she expects to give more than the mean in the group, or:

\[
r_{i,t} = \begin{cases} 0, & x_{i,t} \geq \bar{x}_{j,t-1} \\ \frac{x_{i,t} - \bar{x}_{j,t-1}}{x_{i,\text{max}}} & x_{i,t} < \bar{x}_{j,t-1} \end{cases}
\]

**Equation 2:** Not being taken advantage of (the vector of others’ giving is unknown)

Given that each player in the VCM was given a budget of 10, $x_{i,\text{max}}$, the maximum possible contribution in this VCM game, is 10. Hence $r$ has a minimum value of 0, and a maximum of 1. When $x_{i,t} > \bar{x}_{j,t-1}$, the player expects to be exploited and $r$ is positive, thus decreasing $\alpha$ and the

---

5 On the other hand, were members of the group able to know each other’s giving, they might use a ranking $\rho$ of how much they gave in the last round. Then $r$ could take a different form. Being ranked 1 is being ranked as the biggest donor. Obviously the median, in this game, would be 3. One measure of how much of a sucker one has been is how one ranks as a donor relative to the median. If $i$ finds herself as above the median, in giving, she will pull back ($r$ would then be negative). If she finds herself as below the median, her $r$ would be 0.
player’s other-regarding motives.\(^6\) When \(x_{ij} \leq \bar{x}_{j,t-1}\), i.e., \(r\) is 0, \(\alpha\) is left unmodified. So obviously, were \(i\) to set \(x\) to the average others gave in the last round, \(r\) would be 0. But this need not be maximizing behavior for \(i\).

### Two Models of Context Dependent Behavior

Consider, now, two alternatives for putting these elements together into a single preference structure for the individual. As stated above, we believe that these concerns modify the subject’s other-regarding behavior. We do so by developing a set of alternative models, each designed to conform to our notion that a given player’s behavior varies as the behavior of others cause the decision context to shift.

#### Avoiding Exploitation with a Continuous Other-Regarding Utility Function

To begin with, consider the simplest case: where there is a continuous, twice differentiable utility function. The individual has the two classes of motivations listed earlier: selfish and other-regarding. But the other-regarding motive is degraded when the individual feels that she will be exploited. In this model, we assume that the individual player conditions \(\alpha\) directly on the deviation between her contribution and the mean of those of others: \((1-r_{ji})\), and she can have a range of concern over \(r\), which we will label \(\theta_i\). The player’s other-regarding motive then has the form:

\[
\text{Equation 3: Other-regarding motive}
\]

\[
f_{ij} = \alpha_i \cdot (1 - r_{ji})^{\theta_i} \cdot x_{ij}
\]

Our basic Equation 1 can now be rewritten as:

\[
U_{ij} (x_{ij} \mid x_{j,t-1}, x_{i,t-1}) = (10 - x_{ij}) + 0.4(\sum x_{j,t} + x_{i,t}) + \alpha_i \cdot (1 - r_{ji})^{\theta_i} \cdot x_{ij}
\]

\[
\text{Equation 4: Detailed objective function}
\]

Where \(r\) only affects \(\alpha\) when \(r\) is positive.

The basic decision of the individual agent will be dictated by Equation 4. Behavior in the first round is not specified, however. Contributions in the first round will reflect exogenously developed expectations, which are assumed to lead to random contribution levels across individuals. After that, the individual will decide how much to give by maximizing Equation 4. Changes in the player’s utility maximizing contribution level are driven by \(\bar{x}_{j,t-1}\). (In the terminology developed presented above, \(\bar{x}_{j,t-1}\) is the time-dependent “context” of the agent’s decision.) Figure 2, below, provides the utility-maximizing contribution level for a given \(\bar{x}_{j,t-1}\), with \(\alpha_i = 1\) and \(\theta_i = 2\).

---

\(^6\) Here we tie the expectation to \(\bar{x}_{j,t-1}\) but it could be modeled in a more nuanced manner involving learning. But the added complexity of learning is not necessary for our analysis at this point, and we revisit the issue in the sensitivity analyses.
Responding to Exploitation with a Probabilistic Other-Regarding Utility Function

An alternative model of context dependent preferences can be developed with only small modifications to our presumptions. Reconsider the general utility function, Equation 1. If players are not constrained to avoid exploitation over a continuous utility function, a plausible alternative is that they respond to other players’ actions in a probabilistic fashion.

While numerous specifications could be considered, for simplicity and direct comparison with the first model, we will consider how an individual could probabilistically respond to the degree of “exploitation”. We will model this response with a random variable, $\lambda_i$, which plays an analogous role to $r$ from Equation 2. $\lambda_i$ is zero when the agent gave no more than the average in the previous round and thus has no reason to feel exploited. In this case, she will maintain her default level of other-regarding motivation, incorporating her default concern for others alongside her monetary self-interest. If the agent’s contribution in the prior round, $x_{i,t-1}$, was greater than the average contribution level of the other players, $\bar{x}_{j,t-1}$, there will be a $(x_{i,t-1} - \bar{x}_{j,t-1})/x_{imax}$ chance that the agent will respond negatively to this exploitation. Again, the maximum possible contribution is 10, and the probability of a negative response lies in the range [0,1]. For simplicity, we assume that the response is directly proportional to the exploitation. $\lambda_i$ can be specified in full as follows:
• if $x_{i,t-1} \leq x_{j,t-1}$, i.e., the agent was not exploited, $\lambda_{i,t} = 0$ with probability 1;
• if $x_{i,t-1} > x_{j,t-1}$,
  - $\lambda_{i,t} = 0$ with probability $1 - \frac{x_{i,t-1} - x_{j,t-1}}{10}$, and
  - $\lambda_{i,t} = \frac{x_{i,t-1} - x_{j,t-1}}{10}$ with probability $\frac{x_{i,t-1} - x_{j,t-1}}{10}$.

As a first pass, we can write the player’s other regarding motive as:

$$f_{i,t} = \alpha_i \cdot (1 - \lambda_{i,t})^\theta \cdot x_{i,t}$$

**Equation 5: Other-regarding motive with a probabilistic response, Version 1**

Where our generic Equation 1 would be rewritten for this model as:

$$U_{i,t}(x_{i,t} | x_{i,t-1}, x_{j,t-1}) = (10 - x_{i,t}) + 4(\sum x_{j,t} + x_{i,t}) + \alpha_i \cdot (1 - \lambda_{i,t})^\theta \cdot x_{i,t}$$

**Equation 6: Probabilistic decrease in other regarding motives, Version 1**

In comparing this model with the previous model, it should be clear that we have made two departures. First, we have given exploitation a probabilistic effect on the agent’s utility function, by replacing $r$ with the random variable $\lambda$. Second, we have changed the agent’s context-dependence from avoiding exploitation in the current round to responding to exploitation from the previous round. The impact of each of these departures will be made evident as we present the results of the simulations. Since we have replaced $r(x_{i,t})$ with $\lambda(x_{i,t-1})$ and since $\lambda(x_{i,t-1})$ is constant for any given round, the new utility function is linear in $x_{i,t}$, making the utility-maximizing contribution trivial and not compelling. When we add the additional assumption of decreasing marginal returns to other-regarding behavior and scale alpha to remain comparable with the previous model, we get a more compelling utility function:

$$U_{i,t}(x_{i,t} | x_{i,t-1}, x_{j,t-1}) = (10 - x_{i,t}) + 4(\sum x_{j,t} + x_{i,t}) + \frac{\alpha_i}{10^{\gamma_{i,t}}} \cdot (1 - \lambda_{i,t})^\theta \cdot x_{i,t}$$

**Equation 7: Probabilistic decrease in other regarding motives, Version 2**

Where $\gamma_i \in (0,1)$. Changes in the player’s utility maximizing contribution level are driven by $\lambda_{i,t}$, which forms the time-dependent “context” of the agent’s decision. Figure 3, below, provides the utility-maximizing contribution level for a given $\lambda_{i,t}$, with $\alpha_i = 1$ and $\theta_i = 2$, $\gamma_i = 0.5$, and the derived probability of the context evoking a pure self-interested response.

---

7 An analysis of the first model, that of avoiding exploitation, will show that $r(x_{i,t})$ played a similar role in the utility function. Since we have replaced $r(x_{i,t})$ with $\lambda(x_{i,t-1})$, we need to explicitly state an assumption of decreasing marginal utility.

8 As $X$ increases, Gamma will distort the impact of alpha, however. To retain the same alpha when $X$ is at its maximum value, we apply a scale factor.
One obvious question arises as to how the agent should judge exploitation and how to respond to it. In this simple implementation, each round the agent has a probabilistic decrease in $\alpha_i$ based solely on the degree to which she was exploited in the previous round. We thus assume that the agent “starts fresh” each round, and is willing to forgive previous transgressions after she has responded in kind. This strategy is analogous to a probabilistic version of Tit For Tat in the Iterated Prisoner’s Dilemma game, where a single round of punishment is meted out for a single round of transgression.

**Methodology**

**Design of the Simulation Model**

To develop an agent-based simulation to mimic the above theoretical models we simulate a group of participants, i.e., agents, give them each an endowment at the beginning of each round, and the allow them to contribute some or all of it to the production of a public good. Conceptually, our simulation consists of three parts: the VCM itself, the experimental subjects (here the agents) participating in the game, and the set of tools to visualize and record data from the game and to explore the sensitivity of the simulation model to its parameters.9

---

9 The model was built upon the RePast Simphony agent-based modeling platform (Tatara et al. 2006), with subsequent elements programmed in Java. RePast provides helpful visualization and data logging tools.
The design of the VCM is straight-forward, and directly follows F&O. At the end of each round the software totals the contributions given and provides each agent with its payoff. At the end of the game, the software outputs a complete record of the contributions and payoffs for each agent for later analysis by the researcher. During the game, agents are unable to communicate with each other, make their decisions simultaneously, and are unaware of the total duration of the game. The researcher sets parameters to specify the endowment given to each player, the number of players, the rate of return of the public good, and the number of rounds of play. In all simulations presented here, we used parameters drawn from F&O: $10 for the endowment, 5 players, a 40% rate of return, and 15 rounds of play.

The design of the game’s agents is considerably more complex, and provides the flexibility needed to simulate context-dependent preferences and a range of alternative theories that might generate jagged contributions. Here we will only focus on the components used in the two theoretical models given above, and introduce other components used in sensitivity analyses later on. At the start of each round, each agent receives information about the average contribution of the other agents in the prior round, if any. The agents then maximize their utility over the range of potential contributions via a simple numerical approximation: one hundred evenly spaced contribution increments are evaluated in their utility function, and the contribution that generates the highest utility level is chosen.

The simulation model requires two sets of input files. The first file provides the rules of the VCM game, i.e., the endowment given to players, the rate of return of the public good, and the number of rounds of play. The second file provides the full specification of each agent in the game including the type of utility function, and the variables necessary for that function ($\alpha$, $\theta$, and $\gamma$). The input files are archived along with the output of each simulation.

**Execution of the Model**

The agent-based simulation was executed in two ways. First, we employed the Graphical User Interface provided by RePast Simphony to load the VCM game. A number of graphs were developed on which to watch the progress of the simulation, including the agents’ individual contributions over time, the agents’ payoffs over time, and the average total contributions for the group over time. In addition, all of this information is logged to file for later analysis. Once the software is loaded, the user can adjust parameters and re-run it instantaneously.

Since the benchmark model is stochastic, individual runs are inconclusive on their own. It would be tedious, however, to execute the software manually a sufficient number of times in the Graphical User Interface to generate meaningful results. For each theoretical model, we executed the simulation in batch mode across one hundred unique random number seeds. The results were logged to a file and analyzed in the statistical package R. This design is inspired by Catherine Dibble’s concept of a “Computational Laboratory”, which employs an agent-based model at the center of a larger process of sensitivity testing, optimization, and statistical analysis (2006).10

---

10 As with the computational laboratory described in Dibble (2006), this simulation model has the capability to perform parameter sweeps over a set of variables (not simply random number seeds), and perform optimization over an expansive parameter space. These features were not used in the final results presented here, but may be of value for further research on this model.
Initial Scenarios Considered

Thus far, we have discussed two potential models of context dependent preferences, and the set of parameters which are required to complete the utility function for each player in the game. We used the following scenarios in the initial analysis:
<table>
<thead>
<tr>
<th>Scenario</th>
<th>Types of Players, by Utility Function</th>
<th>Distribution of Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5 Continuous Utility Agents</td>
<td>Homogeneous: $\alpha=1; \theta=2$</td>
</tr>
<tr>
<td>B</td>
<td>5 Continuous Utility Agents</td>
<td>Heterogeneous: $\alpha=(1, .8, .6, .4, .2); \theta=2$</td>
</tr>
<tr>
<td>C</td>
<td>4 Continuous Utility Agents, 1 Purely Self-Interested Agent</td>
<td>Heterogeneous: $4(\alpha=1; \theta=2); 1(\alpha=0; \theta=NA)$</td>
</tr>
<tr>
<td>D</td>
<td>5 Probabilistic Response Agents</td>
<td>Homogeneous: $\alpha=1; \theta=2; \gamma=0.5$</td>
</tr>
<tr>
<td>E</td>
<td>5 Probabilistic Response Agents</td>
<td>Heterogeneous: $\alpha=(1, .8, .6, .4, .2); \theta=2; \gamma=0.5$</td>
</tr>
<tr>
<td>F</td>
<td>4 Probabilistic Response Agents, 1 Purely Self-Interested Agent</td>
<td>Heterogeneous: $4(\alpha=1; \theta=2; \gamma=0.5) 1(\alpha=0; \theta=NA; \gamma=NA)$</td>
</tr>
</tbody>
</table>

**Results**

**Benchmark Models**

**Summary**

The Probabilistic Response model successfully replicated the phenomenon of jagged contributions under scenarios with heterogeneous preferences (E,F) while all other scenarios failed to replicate the phenomenon. Both the success of Scenarios E and F, and the failure of the other versions provide insights into a possible mechanism driving jagged contributions in the laboratory.

**Scenarios A-C: Continuous Utility Agents with Homogeneous and Heterogeneous Preferences**

In each of these scenarios, and for each stochastic simulation, the Continuous Utility Agents settled on a contribution level within the first four rounds, and remained at that level throughout the game. Scenario A, with homogenous preferences, showed a strong sensitivity to the initial (random) contributions made by the agents in the first round. Figures 4 and 5 provide sample simulation results with high and low initial mean contributions. The contributions of each of the 5 agents over the 15 round VCM game are shown, though the players’ contributions quickly converge and become indistinguishable. The lines are labeled “CU” to signify that the agents have a Continuous Utility function. In this scenario the agents are homogonous and thus their numbering is merely nominal.
In Scenario B, with heterogeneous preferences, each player adapted to the other player’s contribution, and found a unique level of utility maximizing contributions. There was relatively little sensitivity to initial contribution levels, as shown in Figures 6 and 7 below. Again, the contributions of each agent of the 5 agents over the 15 round VCM game are shown as separate lines. The agents are numbered in decreasing order of $\alpha$ ($\alpha_{CU\_Agent\_1} = 1.0$, $\alpha_{CU\_Agent\_2} = 0.8, \ldots$)

![Figure 6: Scenario B: Heterogeneous Continuous Utility Agents, High Initial Contributions](image)

![Figure 7: Scenario B, Heterogeneous Continuous Utility Agents, Low Initial Contributions](image)

In Scenario C, with four homogenous potentially-cooperative players and one purely self-interested player, the potentially cooperative players found their common utility maximizing level of contributions also by round 5. This joint contribution level was generally lower than without the purely self-interested player (Scenario A), which would be expected since the continuous utility agents avoid exceeding the average contribution level of the others. Figures 8 and 9 provide examples. The Self Interested Agent is labeled “SI_Agent”, the other potentially-cooperative Continuous Utility agents are nominally labeled “CU_Agent” 1-4.

![Figure 8: Scenario C, Four CU Agents with One Pure Self-Interested Agent](image)

![Figure 9: Another Example of Scenario C, Four CU Agents with One Pure Self-Interested Agent](image)

Scenarios D-F: Probabilistic Utility Agents with Homogeneous and Heterogeneous Preferences

In Scenario D, probabilistic utility agents confronted with clones of themselves find a stable medium level of cooperation no matter their initial starting contributions. Since they all maximize utility at the same level of contribution, their probabilistic withdrawal of cooperation is never
triggered, and no jaggedness in contributions occurs. Sample results, with different initial contributions, are shown in Figures 10 and 11. These Probabilistic Response agents are nominally labeled “PR_Agent” 1-5.

Once heterogeneous preferences are introduced in Scenario E, and some agents are potentially contributing more than others, jagged contributions appear. Not surprisingly, jaggedness occurs among the agents who by default would contribute more than the others. That is, jaggedness is not consistent across all players – players with greater other-regarding motives are more likely to be exploited, and thus exhibit jagged contributions. Figures 12 and 13 provide examples. The agents are numbed in decreasing order of $\alpha$ ($\alpha_{PR_Agent_1}=1.0$, $\alpha_{PR_Agent_5}=0.8$, ...)

Scenario F shows that the probabilistic agents are clearly being taken advantage of by one purely self-interested agent, the jaggedness increases. They tend towards a 70% level of contributions as seen in Scenario D ($\alpha=1$, as in Scenario D). However, there are frequent jumps to lower contribution levels, as shown in Figures 14. This jaggedness is not always consistent – in a small portion of the simulations, random chance led to fewer jumps over the 15 round game; an example is given in Figure 15.
Sensitivity Tests

Alternative Parameters with the Benchmark Models

Given that the benchmark theoretical models use specific sets of parameters for the given context-dependent utility functions, the question naturally arises as to the sensitivity of the models’ outcomes to these parameters. We developed ten sensitivity tests to determine the parameters that generate jagged contributions and thereby clarify the driving mechanism behind the model. These tests are divided into two groups: first, those starting from benchmark scenario A, which did not generate the phenomenon, to seek parameters that enable jagged preferences, and second, those starting from benchmark scenario E, which did generate the phenomenon, to seek parameters that restrict the generation of jaggedness.

The sensitivity tests which sought to induce jaggedness from Scenario A (with Continuous Utility Agents) were:

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Population of Agents, by Utility Function</th>
<th>Distribution of Parameters $\alpha$, $\theta$</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>5 Continuous Utility Agents</td>
<td>Homogeneous: $\alpha$ each of (0.1, 0.5, 1, 2, 4, 8); $\theta$=2</td>
<td>Does not induce jaggedness. $\alpha$ only affects the contribution level to which the agents stabilize.</td>
</tr>
<tr>
<td>H</td>
<td>5 Continuous Utility Agents</td>
<td>Homogeneous: $\alpha$= 1; $\theta$ each of (0.1, 0.5, 1, 2, 4, 8)</td>
<td>Does not induce jaggedness. A lower (but positive) $\theta$ makes the agent less sensitive to being taken advantage of, and, on average, it raises the final stable level cooperation (from random initial starting points). A higher $\theta$ does not destroy cooperation, but it does make the final level of cooperation more sensitive to the random starting point.</td>
</tr>
<tr>
<td>I</td>
<td>1 Continuous Utility Agent, 4 Purely Self-Interested Agent</td>
<td>Heterogeneous: $4x(\alpha=1; \theta=2)$ $1x(\alpha=0; \theta=NA)$</td>
<td>Does not induce jaggedness. The presence of more self-interested players lowers the stable level of contributions given by the continuous utility agent.</td>
</tr>
</tbody>
</table>
Change the population to consist of 4 random and one Continuous Utility agent. For Continuous Utility Agent: \( (\alpha=1; \theta=2) \)

Induces apparently random contributions, **not jagged ones**. The presence of other players contributing at random makes \( x_{i,t} \) vary randomly, and causes the agent’s behavior to appear random (though in fact it responds in a deterministic way to a random signal). In a single simulation contributions may appear “jagged”, but this is not consistent across simulations.

The sensitivity tests to remove jaggedness from Scenario E (with Probabilistic Utility Agents) are described in the next table:

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Baseline Scenario and Modified Parameter</th>
<th>Population of Agents, by Utility Function</th>
<th>Distribution of Parameters ( \alpha, \theta, ) and ( \gamma )</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>Test extremely small deviations in ( \alpha )</td>
<td>5 Probabilistic Response Agents</td>
<td>Heterogeneous: ( \alpha=(.91,.92,.93,.94,.95); \theta=2; \gamma=0.5 )</td>
<td>Removes jaggedness. Sufficient heterogeneity must exist for jagged contributions.</td>
</tr>
<tr>
<td>L</td>
<td>Decrease all values of ( \alpha )</td>
<td>5 Probabilistic Response Agents</td>
<td>Heterogeneous: e.g., ( (\alpha=(0,0.2,0.4,0.6,0.8); \theta=2; \gamma=0.5) )</td>
<td>Can remove jaggedness. Lowering all values of ( \alpha ) to the range ([0, 0.8)) causes more severe jaggedness for the remaining high-contribution players. Further lowering of ( \alpha ) causes all contributions to break down.</td>
</tr>
<tr>
<td>M</td>
<td>Change two agents to purely self-interested</td>
<td>3 Probabilistic Response Agents, 2 Purely Self-Interested</td>
<td>Heterogeneous: ( 3\times(\alpha=(.2,.4,.6,.8,1); \theta=2; \gamma=0.5) ) ( 2\times(\alpha=0; \theta=NA) )</td>
<td>Does not remove jaggedness.</td>
</tr>
<tr>
<td>N</td>
<td>Test values of ( \theta )</td>
<td>5 Probabilistic Response Agents</td>
<td>Heterogeneous: ( (\alpha=(.2,.4,.6,8,1); \theta each of (.1,.5,1.2,4,.8); \gamma=0.5) )</td>
<td>Does not remove jaggedness. Theta affects the depth of each response. Low theta leads to small jumps in contributions, higher theta leads to higher jumps, until agents jump from their default contribution down to 0.</td>
</tr>
<tr>
<td>O</td>
<td>Test values of ( \gamma )</td>
<td>5 Probabilistic Response Agents</td>
<td>Heterogeneous: ( (\alpha=(.2,.4,.6,8,1); \theta=2); \gamma each of (-1,.1,.3,.5,.7,.9,1,2) )</td>
<td>Can remove jaggedness. Decreasing ( \gamma ) lowers overall contributions ( (\gamma=0.3, 0.5) ) until all contribution breaks down ( (\gamma=-1, 0.1) ). Increasing ( \gamma ) ( (\gamma=0.7, 0.9, 1, 2) ) causes contributions to oscillate between 100% and 0% every round.</td>
</tr>
<tr>
<td>P</td>
<td>Force all agents to start at same contribution level</td>
<td>5 Probabilistic Response Agents</td>
<td>Heterogeneous: ( (\alpha=(.2,.4,.6,8,1); \theta=2; \gamma=0.5) )</td>
<td>Does not remove jaggedness. After the initial contribution, less altruistic players lower their contributions, triggering the others to probabilistically lower theirs as in scenario E.</td>
</tr>
</tbody>
</table>

Wendel & Oppenheimer, Simulation
Mixing Elements of the Benchmark Models

Of the two benchmark theoretical models studied thus far, the model of avoiding exploitation via utility maximization over a continuous utility function clearly fails to generate the phenomenon of jagged contributions, while the model of probabilistic responses to exploitation does generate the phenomenon. Since these models differ by only three assumptions, we examined the sensitivity of the results to each assumption. To recap, the three differences are:

- responding to prior exploitation versus avoiding current exploitation
- a probabilistic versus a deterministic impact of exploitation
- decreasing marginal utility versus linear utility in other-regarding motives.

We started the analysis with the unsuccessful Scenario B, that of a set of heterogeneous players who avoid exploitation via utility maximization over a continuous utility function, because it matched the parameters, if not theoretical assumptions, of the successful Scenario E. We then applied each assumption to determine the driving force behind the jagged contributions:

<table>
<thead>
<tr>
<th>Tests to induce jaggedness, changing the structure of the Continuous Utility Agents (i.e., from Scenario B)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scenario</strong></td>
</tr>
<tr>
<td><strong>Q</strong></td>
</tr>
<tr>
<td><strong>R</strong></td>
</tr>
<tr>
<td><strong>S</strong></td>
</tr>
<tr>
<td><strong>T</strong></td>
</tr>
</tbody>
</table>

There are four tests instead of six because two of the possible combinations are not meaningful: the agent cannot maximize over a probabilistic response to expected exploitation (with or without decreasing marginal utility).
Scenarios Q and S, which both produced oscillations between two contribution levels, spurred us to briefly test other deterministic responses to prior exploitation. While a full examination of potential models of learning behavior is not focus of this paper, we found that non-oscillating, jagged contributions did not occur under a number of learning models. If agents equally weighted the past N rounds of the game, the period of oscillation increased and then dampened, but jaggedness did not occur. The addition of a discount factor in the agents’ calculation of prior rounds’ exploitation also did not cause jaggedness. Finally, making the agents heterogeneous in their discount factors and time horizons, also failed to generate clearly jagged contributions. Again, these results are only preliminary and more detailed models such as Bayesian updating could be investigated in future work.

Alternative Models of Jagged Contributions

We developed two alternative theoretical models which could potentially generate jagged contributions among the players of the VCM game. The first is a variation on the probabilistic utility function given above, where a decrease in self-interest is triggered when personal contributions fall below the average – a stylized form of guilt. The second is a departure from utility maximization, where individuals instead have multiple goals and change their mind over time as to which goal to pursue.

**Simple Alternative Model 1: “Context-Dependent Guilt”**

In the probabilistic response model discussed above, the agent has a default preference for other-regarding behavior that is degraded probabilistically withdraw if other agents are exploiting her. We can generate a simple alternative model by modifying the default preference and response. For example, assume that the agent places the most weight on her self-interest. But now, let us add guilt: if the agent contributed significantly less than the others in the previous round, it will probabilistically (and temporarily) decrease its tendency for self-interested action and thus increase its contributions. The agent’s utility function could then be expressed something like:

\[ U_{i,t}(x_{i,t} | x_{i,t-1}, x_{j,t-1}) = (10 - 0.6 \beta \lambda_{ij}^\theta x_{i,t}) + 0.4 \sum x_{j,t} + \alpha_i x_{i,t} \]

Equation 8: Probabilistic Decrease in Self-Interested Motives, Version 2

Where \( \lambda_{ij}^\theta \) is defined as above, but for contributions that are below the average instead of above it. We simulated a population of 5 such probabilistically guilty agents with heterogeneous
preferences; sample results are included in the Figure below. This function also generates jagged contributions.

Simple Alternative Model 2: “Changing Your Mind”

If we retreat from theoretically-grounded models of behavior and look simply for models that could generate jagged contributions, we find many. For example, consider an agent that has a stable set of preferences for self-interest and other-regarding behavior. Instead of maximizing a standard utility function, the agent probabilistically selects among self-interest and other-regarding behavior with frequencies given by these preferences. Acting alone, without any other players, this type of agent can generate jagged preferences. A sample is given below.

Naturally, we do not believe that this is a realistic model of behavior. We find it useful, however, to use such extreme examples to illustrate the limits of the technique we have employed: an issue that we reconsider in the next section.

---

12 In designing these mock agents we were inspired by the work of Gode and Sunder (1993).
Analysis

We now consider the significance of these simulations in three ways. First, we narrowly analyze the results – to judge the successes and failures of the various theoretical models. Second, we consider what aspects of the logic programmed in the simulations led to the particular results with which we are concerned – why jagged contributions occurred (or did not) in the simulations. Third, we consider what insights about real world behavior, if any, one may glean from these simulations.

Review of Simulation Results

The model of probabilistic response among heterogeneous agents clearly replicates the phenomenon of jagged contributions observed in F&O’s VCM experiments. These jagged contributions are robust to changes of the initial contribution levels of the players and the degree to which players withdraw their cooperation in the face of exploitation ($\theta$). Sensitivity analyses showed that changes in the characteristics of the players can hinder the ability of the probabilistic model to generate jagged contributions. This finding should not surprise us, however. If the “default” level of altruism ($\alpha$) among players is too similar, they quickly converge to their common level of contributions, and naturally have no reason to feel exploited (i.e. $r =0$). If $\alpha$ is too low, there is insufficient other-regarding behavior to withdraw, and hence, again, no jaggedness. Similarly, we would generally expect decreasing marginal utility to other-regarding behavior, and when we relax this assumption to have a linear utility function ($\gamma = 1$) or increasing marginal utility ($\gamma > 1$) degenerate results occur. On the other hand, the continuous utility model clearly does not generate the phenomenon, and that result is robust to changes in the core parameters ($\alpha$) and ($\theta$) in individual agents, as well as the distribution of these characteristics across the agents.

What is Causing the Simulation Results?

Taking a wide view of the main simulations, sensitivity analyses, and alternative specifications, a set of sufficient conditions for jagged contributions emerges. First in the set is heterogeneity: agents should be heterogeneous in their other-regarding motives, or else they all cooperate and have no reason to withdraw that cooperation. Second, is a context defined by the prior contributions of others. If the agents merely seek to avoid future exploitation, they will find the stable, optimal level of contribution to balance any exploitation against their other-regarding motive. When agents respond to prior exploitation, making a temporary shift in their preference for other-regarding behavior, their behavior varies over time. Third, to avoid a clear pattern of oscillation, agents should respond probabilistically to exploitation. Fourth, to avoid the utility maximizing level of contribution being constant or a step-function, the agent’s utility function must not be linear. Decreasing marginal utility in other-regarding behavior is one means ensuring non-linearity. Our sensitivity analyses have illustrated that when any of these elements are relaxed, the remaining elements are insufficient to generate jagged contributions. When all four are present, as in Scenarios E and F, jagged contributions were generated in our simulations.

It may turn out that despite our sensitivity analyses, the results were dependent upon the particular set or range of parameters chosen, or other unexamined assumptions in the model. While the goal of this paper lay elsewhere, in testing the power of a model of context dependent

---

13 A more complex version with noise in the observation process has also been tested and can also trigger a sense of exploitation and jagged contribution levels.
preferences, we should note that it can be extended to substantiate (or refute) our analysis of the logical relation between these behavioral mechanisms and the patterns of behavior observed. This raises the larger question of the relationship between simulation and real-world behavior.

**Insight into Real World Behavior**

Of course, the simulations don’t permit us to conclude that the individuals captured in the F&O data contributed in a jagged fashion because they had other-regarding preferences that were probabilistically decreased due to individuals feeling exploited by others. Alternative specifications and models also generated the phenomenon, and they form viable conjectures that can imply jagged contributions. The simple theory that people feel probabilistically guilty, or even that they have multiple goals and probabilistically select between them over time, generates the phenomenon of jagged contributions. Without solid data, and further logical development of implications, neither the “probabilistic guilt” nor the benchmark theory of a probabilistic withdrawal of other-regarding behavior may be rejected as plausible explanations of jagged contributions. Such a variety of plausible premises is to be expected. Although this study considered two very basic alternative explanations, other plausible theories could be developed that also generate the phenomenon.

As with any other model, our simulations allowed us to probabilistically test the logical ability of the model to generate the empirical data. In the agent-based modeling world, this is referred to as a “generative test;” such a generative test suggests strongly that we ought to exclude the model of continuous utility agents avoiding future exploitation. However, a generative test is insufficient to establish the validity of a model. We must delve deeper to find the significance of these simulations for real world behavior and vet the model. In order to eliminate alternative hypotheses, one can examine the conditions under which the theories make divergent predictions.

For example, consider the effect of placing either a “probabilistically guilty” or a “probabilistically selfish” player in two sets of environments. First, place the player among a heterogeneous set of other such players. Both the “probabilistically guilty” and “probabilistically selfish” can generate jagged contributions, depending on chance and the average level of contributions as determined above. Second, place the player in an environment where others contribute less than the given player. One would expect the “probabilistically guilty” player’s contributions to be constant, while the “probabilistically selfish” player to display jagged contributions. The other simple model of jagged contributions, we discussed, that of a player “changing her mind,” would be likewise easy to distinguish. Such divergent predictions establish experimental tests by which a researcher could empirically falsify one or more of the theories.

These simulations have not addressed the underlying causes of what we have referred to as “other-regarding behavior,” but they could be readily expanded to do so. Here, we have assumed a certain level of other-regarding preference for each agent, and then analyzed how this preference would interact with a changing context to create jagged preferences. Obvious explanations for other-regarding behavior come to mind from the literature: an inherent altruism, a concern for reputation, habit, conformism, or a preference for the best joint outcome. There has been significant research to distinguish between these motives in VCM games. However, there has been little on how these motives would generate jagged contributions. If people have an inherent sense of altruism, then it should not matter what other players contribute, and jaggedness would be unlikely. If people work towards the best joint outcome, then we would expect them to make some

---

14 One could make an inference that the benchmark context-dependent preference model is has a higher likelihood of being an accurate explanation of the observed phenomenon, but for a small data set of real human-subject experiments, this inference would have little leverage.
signal of their intentions, and the difficulty in interpreting signals in such an environment could also lead to jagged contributions. If people act out of a concern for reputation, we could not a priori eliminate the possibility of jagged contributions, but we would expect that removing all accountability should decrease the jaggedness (and contributions). Indeed, each theory provides a rich set of observable implications that could be examined both in agent-based models (for logical consistency), and in experiments for plausibility.

After this analysis, we can now make a small comment on the contributions of agent-based modeling to a larger experimental research program. Our simulation model allowed us to probabilistically test the logical ability of the model to support the empirical data. The agent-based modeling framework also allowed us to design intuitively attractive alternative specifications of the theory and check their implications. The framework then facilitated the rapid implementation and refinement of these theories, as we enforced and relaxed assumptions at will, making it possible to isolate the sufficient conditions for generating jagged preferences in the probabilistic response model. It allowed us to investigate the sensitivity of the various models to their parameters; while a full sensitivity analysis in a Computational Laboratory (Dibble 2006) would provide a more thorough and rigorous picture, these basic tests have helped guide our thinking. Finally, the agent-based model has identified the circumstances under which the theories provide divergent predictions, thus improving our leverage with small tests. Such modeling saves valuable time and money: providing a tool to sketch ideas and deduce their implications in a cost-effective and rapid manner. And when the real world is amenable to experimental tests, the simulations can help us identify the best strategies for repeated returns to laboratory testing. An agent-based model such as ours provides an invaluable tool for streamlining and providing leverage to a larger experimental researcher program.

**Conclusion**

This paper confirms the conjecture that under certain assumptions context-dependent preferences can generate the previously unexplained phenomenon of jagged contributions as observed in F&O’s experimental work on voluntary contribution games. It also helped identify alternative conjectures (e.g. “probabilistic guilt”), that also generate the phenomenon. A logical next step would be to go back to the details of the original experimental data to see if the conditions under which jaggedness occurred mirror the specified antecedents in our trials. Afterwards, a return to the laboratory would be called for in order to distinguish between and falsify theories based on the divergent behavior predicted by the agent-based model.

One of the anomalies brought out in many VCM experiments is that minimal communication leads to a break out of cooperation. Our results cast some light on this finding. If many or most are motivated to get to a cooperative outcome via some sort of communitarian values, then, as was conjectured in F&O, without communication the individuals try to signal their willingness to donate. In a 5 person VCM without communication, these signals would be very noisy, and difficult for others to decode.

The result could be a lot of misinterpretation and erratic, saw-toothed behavior as the response to lower valuations of cooperation by others in the group. We know that when subjects can communicate intentions with even rudimentary clarity, they almost always cooperate. The results from our probabilistic response model generate presumptive evidence for such a hypothesis.

If still further work is to be pursued on this topic, it would be first to examine in detail the experimental data that shows jagged responses, to see which antecedents occurred. Next one should consider setting up experimental tests to falsify one or more of the alternative theories based on
their divergent predictions of individual behavior under new scenarios, as described above. Of course, the method is also the message: This same agent-based framework could easily be extended to analyze additional theories of choice-making and additional institutions beyond those in the simple VCM games.

References


