Keywords:
Computational Economics; Agent-based models; Social Dilemmas; Collective Action; Public Goods.

INTERNATIONAL JOURNAL OF ECONOMICS AND ECONOMIC POLICY
DEPARTMENT OF ECONOMICS
UNIVERSITY OF PARMA
VIA KENNEDY 8
43100 PARMA
mraimondi@live.it

M. Raimondi¹

Computational rationality and voluntary provision of public goods: an agent-based simulation model.

(JEL Classification: C63 ; D64 ; D80 ; H41)

¹ I wish to thank Franco Donzelli and Alessandro Arrighetti for their unstinting help and supervision of my work. I would like to thank Giorgio Fagiolo for his advice and helpful observations. Last but not least, I would particularly like to thank Alessandro Vezzani, author of the source code of the software developed here, for his valuable collaboration in the work and creative support in carrying out the simulations. Finally I would also like to thank Salvatore Curatolo, Andrea Lasagni, Mario Menegatti, Antonio Affuso, Davide Tondani, and Vincenzo Dall’Aglio, for their advice and helpful observations. All errors are mine.
Abstract

The issue of the cooperation among private agents in realising collective goods has always raised problems concerning the basic nature of individual behaviour as well as the more traditional economic problems. The Computational Economics literature on public goods provision can be useful to study the possibility of cooperation under alternative sets of assumptions concerning the nature of individual rationality and the kind of interactions between individuals.

In this work I will use an agent-based simulation model to study the evolution of cooperation among private agents taking part in a collective project: a high number of agents, characterised by computational rationality, defined as the capacity to calculate and evaluate their own immediate payoffs perfectly and without errors, interact to producing a public good. The results show that when the agents’ behaviour is not influenced either by expectations of others’ behaviour or by social and relational characteristics, they opt to contribute to the public good to an almost socially optimal extent, even where there is no big difference between the rates of return on the private and the public investment.

Introduction

Economic theory offers many models that explain the free riding dominance in terms of strategic rationality, while Experimental Economics presents many experiments where individual rationality, in conjunction with other factors, influences the agents’ behaviour in such a way as to produce some cooperation, at least in the early phases of the game.

However, one question remains unexplained: whether (and how) agents without either strategic rationality or social and emotional factors affecting their choices can cooperate and successfully produce public goods.

The Computational Economics literature on public goods provision can be interpreted as an attempt to study the possibility of cooperation under alternative sets of assumptions concerning on one hand the nature of individual rationality and on the other the kind of interactions between individuals. By fixing the structure of the interactions and by making them as neutral as possible, one can study the evolution of cooperation under a variety of assumptions about the agents’ rationality, from “zero-intelligence” agents up to sophisticated agents. Alternatively, by fixing the type of individual rationality, there are interesting ways to study how the different interaction structures affect the emergence of cooperation.

This work is concerned with the second type of study. I use an agent-based simulation model to study the evolution of cooperation among private agents taking part in collective projects, where a high
number of agents interact with a view to producing a public good. Simulation models, in fact, are particularly useful to model the agents’ rationality within a world where virtual players interact under a given set of rules.

My virtual agents are characterised by computational rationality, defined as the capacity to calculate and evaluate their own immediate payoffs perfectly and without errors.

The absence of errors and the lack of any individual strategic behaviour or relational social characteristic, such as kindness, sense of belonging or fear of exclusion, give rise to unusual dynamics in agents’ behaviour which produce a significantly higher level of voluntary contribution to public goods provision than that prevailing in reference models in both Game Theory and Experimental Economics.

My analysis shows that if the agents’ behaviour is not influenced either by expectations of others’ behaviour or by social and relational characteristics, they opt to contribute to the public good to an almost socially optimal extent, even where there is no big difference between the rates of return on the private and the public investment.

This paper is organized as follows: in the next Section, I introduce some assumptions on rationality in public goods provision contexts and, in particular, I define the notion of computational rationality as employed in the simulations model. In Section 2 I briefly describe the reference model. In Section 3 I exhibit the simulation findings and discuss the results of the validation tests. Finally, Section 4 contains some concluding remarks and suggestion for future research.

1. Assumptions on rationality and outcomes of collective action

The use of computational models has been justified by three key aspects. First, computer simulations are used aside traditional theory as basic computational tools able to solve the computing difficulties arising in multi-equational models. Secondly, simulations are useful in solving more complicated models that cannot be solved analytically in an easy or complete way. In this case, computational systems allow one to clarify the solution structure by illustrating hidden dynamic properties and testing the influence of parameters and assumptions onto the results. Lastly, simulations have found useful applications in a certain class of problems for which to apply a mathematical equation system does not appear immediate or useful (Axtell, 2000).

Agent-based models are particularly useful to study interaction mechanisms among individuals leading to collective outcomes because they are suitable for studying classes of systems characterised by two

---

2 See Raimondi, M (2009, b).
typical properties: they comprise interacting units and possess emerging properties, that is, properties which develop precisely from the interaction between agents (Bruun, 2004).

Since the first applications of game-theoretic frameworks, Social Sciences built a robust theory of cooperation based on strategic rationality and dominant strategies. Subsequently, this theory has been supplemented by many theoretical studies based on other perspectives such as Evolutionary Game Theory, Experimental Economics and Computational Economics, in order to analyse what are the reasons behind the failure associated with the Game Theory predictions and the strategic rationality assumption.

In this Section I present the three kinds of rationality characterising the agents in Game Theory models, in the experiments on public goods provision and in my agent-based model, i.e. strategic rationality, rationality affected by social motives and computational rationality respectively.

1.1 Strategic rationality

In traditional Economics, the voluntary production of public goods is thought to be doomed to failure. Given the two principal characteristics of public goods, non rivalry and non excludability, the individual finds it rational not to take part in the production of a public good, but only to take advantage from it once it has been collectively produced.

A generalized free riding is the outcome that can be predicted under the assumption of strategic rationality: a selfish agent maximizes his payoff expecting the other agents to be rational as well. Even if the agent knows that the collective outcome would be the Pareto-optimal solution, his individual earnings from the provision of the public good are uncertain and the strategy of “no contribution” is the dominant strategy whatever strategy will be chosen by the other players.

The social dilemma is that the Pareto–efficient result for the group (collective action) is doomed to failure because of (1) selfish interest prevails and (2) there are multiple problems of uncertainty which make it difficult to evaluate expected individual and collective benefits as well as the differential between costs and advantages of cooperation. Uncertainty is particularly evident in the forecast of individual behaviour and the agents are not able to predict others’ behaviours.

As a result, the most common type of behaviour in consumption and production phases of public goods tends to be opportunistic, and even when payoffs are higher, cooperation is less widespread than individualistic behaviour. The dominance of this type of individual behaviour leads to the failure of collective action or inefficient outcome.

This outcome is strongly predicted by the hypothesis of strategic rationality. This kind of rationality is one of the most common assumptions made in Game Theory, along with common knowledge of
rationality. Strategic rationality implies that every player is motivated by maximizing his own payoff (or his utility), thus being able to perfectly calculate the probabilistic result of every action.

In the theoretical decision of static games (where the moves of players are simultaneous) the players choose strategies that maximize expected payoff given some beliefs about others' strategies not contradicted by anything they know. In other words, rational players always choose the dominant strategy, i.e. the strategy that, for every choice of strategy of the other players, produces the greater payoff.

The assumptions of common knowledge and rationality imply that:

1. the specification of the game and the players’ preferences among the outcomes, together with everything that can be logically deduced about the game, are common knowledge among the players;
2. the players are rational in the sense that they always seek to maximize their own expected utilities, and this is common knowledge among all players.

The form of rationality specified in the second statement is usually interpreted as referring to the axioms of rational choice under uncertainty while the term “common knowledge” was introduced into Game Theory by Lewis (1969) and later formalized by Aumann (1976).

To summarize, the standard knowledge and rationality assumptions of Game Theory include the assumption that players choose their moves or strategies rationally in the sense of expected utility theory and that the fact that they do this is common knowledge.

The backward induction argument, that leads to the expected outcome of generalised free riding in finitely repeated games, also works on the basis of these assumptions. Most theorists, however, find a conflict between backward induction reasoning and other kinds of reasoning. Much effort has gone into trying to solve the problem. Virtually all of these efforts exploit the extensive-form structure of the above games or the fact that they are played over time.

1.2 Rationality in experiments

Experimental economists as well as sociologists and political scientists state that cooperation is possible and in fact, we see collective behaviour in many aspects of real life.

Many experiments specifically test this theory in the finitely IPD (Andreoni and Miller, 1993; Cooper, DeJong and Forsythe, 1996) and the Centipede game (McKelvey and Palfrey, 1992): they show that

---

3 For further explanation of these assumptions see, e.g., Colman (1997), Colman and Bacharach (1997), Hollis and Sugden (1993), Sugden (1993, 1992).
4 See, e.g., Savage (1954) that formalizes the axioms of rational choice under uncertainty within a Bayesian framework of subjective probabilities.
5 The Nash Equilibrium concept is justified by reference of the concept of common knowledge. A fact p is common knowledge if all players know p, all players know that all players know p, all players know that all players know that all players know p, and so on ad infinitum. It is important to distinguish common knowledge from the situation of general knowledge in which all members of the group merely know the proposition to be true.
explanations in terms of incomplete information and reputation-building cannot explain all violations of backward induction. These results suggest that a small proportion of people behave cooperatively or altruistically irrespective of any assumptions that they may hold about their co-players’ rationality or any attempt to bolster their own reputations.

The typical results of Experimental Economics are high rates of contribution to the public goods at the beginning of the game and a gradual reduction in contributions towards the end of rounds until a null or quasi-null outcome is reached.

As in Game Theory predictions, however, the final result is free riding. The most important variables involved in this drastic decrease of voluntary cooperation are identified in strategic behaviour of the agents and in learning processes about the rules, the payoffs and the structure of games.

But the most significant findings are the highly cooperative behaviours that the agents carry out in the first phases of the game and which are uncommon in the strategic rationality perspective.

In experiments, the rationality characterising agents is of a different type. Here, the rationality of the agents is bounded, and they are also subject to errors in making assessments linked to their social environment. This bounded rationality is not thus exclusively selfish: it is somehow interested in social and collective motives such as altruism, sense of group belonging, fear of social exclusion, reciprocity and imitation.

Positing this type of rationality, experiments have shown atypical cooperative behaviour in the starting rounds of the game, with a total absence of generalised free riding, and a gradual decline of contributions in subsequent rounds: the experimental perspective is interesting particularly in that it stresses the importance of not economically measurable factors and allows one to discriminate among many concurrent explanations.

Briefly, the experimental findings to explain cooperation are typically organized into three categories: decision errors and confusion; strategic cooperation rising from selfish motivations; “other-regarding” preferences related to kindness, fairness, reciprocity, warm glow and social ties. The hypothesis of confusion, simply states that the agents cooperate because they do not understand the dominant strategy; strategic cooperation means that the agents cooperate because they understand that if all other agents will contribute, they can obtain higher payoffs; finally, voluntary cooperation is promoted by social motives to which experimental literature collectively refers as kindness or altruism.

All these concomitant factors have as consequence that the rationality of the subjects involved in experimental analysis is bounded and produces collective results that, like those produced by Game Theory, are not Pareto-optimal.

---

6 The Experimental Economics literature on voluntary production of public goods has dealt widely with this subject and has also included consideration of exogenous variables such as the number and gender of subjects, incentives for cooperation, the presence or absence of thresholds and repetition, the initial information structure, the level of altruism, loyalty or friendship in a group, as well as variables reflecting the legal framework such as whether subjects may communicate and how, etc. For a complete review of this literature, see Raimondi (2009, b).
1.3. Computational rationality

The question of how agents interact adapting their behaviour to achieve a collective goal is strongly influenced by the characteristics of the agents’ rationality. In this paper I attempt to measure the outcome and the dynamics of collective action assuming that the agents are characterised by computational rationality.

With computational rationality, agents neither make errors nor are affected by confusion: they are able to calculate and assess their payoffs with no margin of error, to compare the present payoffs with those obtained in the previous rounds of the game and to make choices of contribution and investment exclusively on the basis of this assessment.

Computational rationality excludes any type of strategies or social motives affecting choices and represents the basic and more innovative element of this model.

In comparison with the bounded rationality of experimental models, computational rationality has the clear aim of overcoming the limits imposed by social motives and relational influences, as well as the limits imposed by confusion, i.e. the difficulty of calculating payoffs and understanding strategies.

Indeed, the comparison with strategic rationality is more complex and critical. Computational rationality is not just a simplified kind of rationality arrived at by ignoring important conditions characterising social dilemmas: rather, it intends to correct or eliminate some of the weaknesses inherent in the conceptual framework where strategic rationality acts.

These weaknesses, which can be traced to an analysis by Luce and Raiffa (1957) of the “backward induction paradox” in the finitely repeated PD, appear to show that strategic rationality and common knowledge imply self-defeating behaviour in certain mutually interdependent decision situations involving sequential choices.

A strategically rational agent, in fact, knows the payoff structure perfectly, so that he also knows that only coordination in collective action leads to a Pareto-optimal outcome. However, his strategic behaviour leads to defection from the common objective, producing an economic outcome that is worse than it would be otherwise.

Empirical evidence suggests that human decision makers do not always follow the backward induction path even when they are capable of understanding the logic of the argument.

Note that this involves an element of incomplete information rather than incomplete rationality on the part of the players: the players are assumed to be strategically perfect rational, and it is the information rather than the rationality assumptions that are relaxed.

---

7 See, for example, Kreps, Milgrom, Roberts, and Wilson (1982). They showed that two rational PD players who each believe that there is a small probability that the other player is irrational will deviate from the prescriptions of backward induction in an attempt to influence the other player by building their reputations for cooperativeness.
Computational rationality removes the inherent weaknesses relaxing the assumption of complete information and allows the agent to assess only the economic efficiency of available alternatives. This prompts behaviour that leads to the best outcome according to the agent's own evaluation. We can define this sort of rationality as “pure rationality”, in the sense that it is purified from informative weaknesses affecting the agents in forming their expectations.

2. Theoretical and simulation models

In this Section I present the model and the structure of agent-based simulations.

2.1 The voluntary contribution mechanism (VCM)

The agent-based model outlined here has the same structure as a typical public good game in experimental models. Its aim is to assess the outcome in a collective process of a system where emotional and environmental factors typically influencing individual choices are “neutralised.” Procedural behaviour is based exclusively on the correct assessment of the agents’ own previous actions.

As stated above, Game Theory predicts an outcome of complete free riding, while the typical result of Experimental Economics is a high rate of contribution to the public good at the beginning of the experiment and a gradual reduction in contributions over time until a null or quasi-null outcome is reached.

The model used here is based on a model typically used in Experimental Economics (Andreoni, 1995) where agents have the free choice of allocating their resources to either public or private investments. I now present in detail the basic framework of the linear public goods game (LPGG) with voluntary contribution mechanism (VCM).

In a LPGG both the production function of the public good and the payoffs are linear. We have a group of subjects endowed with amounts of “tokens” that can either be kept (private investment) or contributed (public good). Subjects play the same game for a finite number of periods. In each period, every subject is endowed with an endowment of $w$. The subject must then divide this endowment between a contribution to a private account ($y_i$) that yields a constant return to the private investor only, and a contribution to a public account ($c_i$) where consumption benefits accrue to all group members.

---

8 The LPGG was introduced by Isaac, Walker and Thomas (1984) and has been widely used in subsequent experimental literature. For a complete analysis of the use of LPGG in Experimental Economics, see Zelmer (2003).
In formal terms, we have two goods, a private good ($Y$) and a public good ($X$) and a set of $N$ individuals:

$n = 1, \ldots, N$.

Each individual receives the endowment $w_i$ at the beginning of each round and the public good is realised by a production function given by:

\[ X = g(C) \]

where $C$ is the total amount of the contributions to the public good, that is

\[ C = \sum_{i=1}^{N} \epsilon_i \]

and $\epsilon_i$ is the individual contribution to the public good given by

\[ \epsilon_i = w_i - y_i \]

For each subject, the individual earnings from the public good are

\[ x_i = g(C) = (\alpha / N)C \]

where $\alpha$ is the rate of return on the public good and $\alpha / N$ is the marginal per capita return of contribution (MPCR). Each agent has the following payoff function:

\[ U_i(y_i, x_i) = py_i + x_i \]

where $p$ is the marginal return to a unit of private good. So, the marginal rate of substitution between the private and the public good is given by:

\[ M = (\partial U_i / \partial y_i) / (\partial U_i / \partial y_i) \]

---

9 MPCR depends on the number of subjects in the group; given $N$ for a group, MPCR is also a constant.
In a LPGG, both the parameters $a$ and $p$ are constant and chosen to create a social dilemma:

$$\frac{a}{N} < p < a$$

Typically, in experiments parameterized in this way, the agent maximizing the utility function, subject to a budget constraint $(w_i = y_i + c_i)$ and a non-negativity constraint $(c_i \geq 0)$, has a dominant strategy that is to contribute nothing to the public good. The Nash Equilibrium is thus to invest in the private good (full free riding). In contrast, for the group as a whole the Pareto efficient outcome is to invest all endowments in the public account.

In order to identify variables that increase and decrease cooperation levels, previous research studies have made various parameters and factors to vary in the respective models of voluntary production of public goods. Some of these factors have big effects on the rate of contribution while others have little explicative power.

Some variables in fact do not have a definite impact either because evidence is ambiguous or because other partially reveals factors cannot be separated out. For these and other reasons, the issue of collective action and coordination is particularly suitable for treatment with ACE, which can include and assess elements typically excluded from traditional models.

2.2 The agent-based model

This work focuses on the definition of computational rationality characterising agents.

I define that agents, at every instant, can perfectly calculate their own payoffs and positively or negatively assess the outcome of previously made choices. If their actions are not guided by assumptions on others’ behaviour, by expectations on the future outcome of collective action (i.e., typically, collective action failure), by social motives or by behaviours that are conform with the group behaviour, they will continuously adapt their actions solely on the basis of actual costs and benefits.

I find evidence that, even where the differential between public and private profit levels is low and the starting population is heterogeneous, the dynamics of the system will work towards a high level of contribution to the public good on the part of almost all agents.

The computational model can exclude the above cited determining factors of experimental work: confusion, altruism and strategic behaviour. Virtual agents are modelled as computationally rational entities: they “only” can calculate correctly the level of incentive and payoff right from the first step. By definition, they are not altruistic: they act only on the basis of their assessment of their own individual result from each decision. In the same way, the agents’ behaviour can be defined as “neutrally strategic” rather than opportunistic.
The model counts $N$ private generic agents who interact for a length of time that is not fixed \textit{a priori}. Agents are heterogeneous in their initial endowments of the private good, and this heterogeneous nature is private information for each agent.

At each step of the game, an agent decides how much of his endowment to invest in the private market, for positive individual profits, and how much to invest in the public good, which yields collective profit equally divided between the group members\footnote{For the sake of simplicity it is assumed that costs of searching and aggregation between individuals are too low to matter.}.

Agents are placed on a lattice which measures 40x40 for a total of 1600 virtual agents. The dynamics of the system is deterministic (myopic): at each instant in time, each agent only knows the state of the system in the preceding period and chooses his investment assuming the future to be identical with the present.

The model has the following parameters and functions:

\begin{align*}
p & = \text{profit rate on private good;} \\
\alpha & = \text{profit rate on public good;} \\
q & = \text{technological constraint on realisation of good;} \\
f_i & = \text{allocation algorithm that assigns to each agent in the lattice a class of contribution valid for the initial period ($t=0$);} \\
k_0^i & = \text{initial characteristic of each agent determined by function } f_i; \\
\end{align*}

The variables used in the algorithms and in the definition of the rules of interaction are as follows:

\begin{align*}
n & = 1, \ldots, N \text{ number of agents;} \\
g & = \text{number of agents aggregated;} \\
w_i^t & = \text{amount of endowment (10 “tokens”) in each period;} \\
y_i^t & = \text{amount of endowment invested in private good;} \\
w_i^t - y_i^t & = \text{amount of endowment invested in public good;} \\
\pi_i^{t,CA} & = \text{instant payoff from “collective action”;} \\
\pi_i^{t,IA} & = \text{instant payoff from “individual action”;} \\
\Pi_i^t & = \text{total instant payoff } = \pi_i^{t,CA} + \pi_i^{t,IA} \\
\end{align*}

\footnote{See also the Appendix for further information on simulation design.}
First, I attempt to test the effects of different kinds of spatial interactions among the agents (global or local interactions). Interaction dynamics in these simulations can have two structures: (1) a simplified global interaction structure, in which each of the \( N \) agents virtually interacts at the same time with all other agents and (2) a local interaction structure, in which each of the agents “locally” interacts with a subset of agents, on the basis of the following rules.

At the first step, \( t = 0 \), each agent “asks” his first and second neighbours how much they will invest in the public good. If the answer is an amount greater than zero which, in conjunction with the agent’s own amount, satisfies the minimum requirement \( q \), the two agents form a group. Otherwise they remain isolated agents.

In the simulated model, for simplicity’s sake, no technological constraint is imposed on the realisation of a public good so that the value of \( q \) is always zero: hence, to form a group, it is sufficient that the answer of the neighbour is greater than zero.

In the same step, aggregated agents invest in the public good the share of endowment determined by the initial distribution, and isolated agents invest their entire share in the private market. The first step ends with the calculation and payment of individual and group payoffs.

In each subsequent step (\( t > 0 \)), the same mechanism of questioning is at work: all agents, both aggregated and isolated, continue to question their first and second neighbours, incorporating these agents into existing “groups”, creating new groupings or continuing to pursue individual action.

But in all subsequent steps there also exists a second procedural rule governing the evaluation of the outcome of previous actions: the agent assesses the “opportunity cost” obtained from previous action.

\[
\pi_{i, CA} = \left[ \sum_{i=1}^{\infty} (w_i^l - y_i^l) (l + \alpha) \right] / n \\
\pi_{i, LA} = (y_i^l) (l + p)
\]

and

\[
P_t = \text{cumulative payoff} = \sum_{t=0}^{T} \Pi_t
\]

---

12 Local interactions are those which occur in the real world, where individuals interact with a proper subset of agents in the system rather than directly with the whole.

13 The first and second neighbours of each agent number eight in all. The selection mechanism for questioning is always random. Second neighbours are also considered: these are agents positioned at corners and are useful because, in every round, they guarantee that there is at least one agent free to questioning for each agent.

14 This mechanism also allows for the questioning among groups. If an agent from one group at random questions an agent belonging to another group, and the two agents’ contributions are greater than zero, the result is the aggregation of two whole groups.

15 The formation of groups has no effect on the final result but is rather the consequence of collective action which started up locally because of the intrinsic characteristics of the rules of interaction. It is seen in the association of agents who agree to produce one and the same public good. In abstract terms, the public good produced by one group at a particular point of the lattice is the same type as a good produced at another point.
Following this “opportunity cost” rule, the agents assess the opportunity cost of investing their tokens alternatively in the public good or in the private good by calculating this value:

\[ A_i^t = \pi_t^{CA} - [(w_i^t - y_i^t)(1 + p_t)] , \]

that is the difference between the (actual) payoff obtained from contributing \((w_i^t - y_i^t)\) to the public good and the payoff that they should obtained from contributing the same \((w_i^t - y_i^t)\) to the private good at the time \(t\). Note that \((A_0^t = 0)\).

An alternative rule of assessing is the “net benefit” rule: the agents only must assess if the previous action produced a result greater than zero, i.e. if the difference between the benefit and the cost of cooperation is greater than zero\(^{16}\).

In these simulations, I decided to use the “opportunity cost” rule in order to make the comparison between the public and the private investment for the agents more specific\(^{17}\).

In deciding what to do at time “\(t+1\)” the agent has to take into account relative, rather than absolute variations, so that his own \(A_i^t\) is compared with the \(A_{i-1}^t\) (the result obtained in the previous step)\(^{18}\).

Lastly, the model also specifies the variable \(\Delta(t)\) which represents the variation of the action each individual undertakes over the public good investment in the preceding period. \(\Delta(t)\) may take only the values +1 and -1.

On the basis of the “opportunity cost” rule of evaluation of the results, the dynamics of the model runs as follows:

\[ \Delta_i^t \in \{+1,-1\} \]

\[
\Delta_i^t = \begin{cases} 
A_{i-1}^t & \text{if } A_{i-1}^t > A_{i-2}^t \\
-A_{i-1}^t & \text{if } A_{i-1}^t < A_{i-2}^t \\
-1 & \text{if } A_{i-1}^t < A_{i-2}^t \text{ and } A_{i-1}^t, A_{i-2}^t < 0 
\end{cases}
\]

\[(w_i^t - y_i^t) = (w_i^{t-1} - y_i^{t-1}) + \Delta_i^t\]

\(^{16}\) To follow this rule, the value that the agent must calculate is given by \(B_i^t = \pi_t^{CA} - (w_i^t - y_i^t)\).

\(^{17}\) Series of simulations were also carried out using the rule based on net benefit: the results were qualitatively similar to those produced by the presented model.

\(^{18}\) It did not appear appropriate to measure \(A_i^t\) in absolute terms, that is, without comparing it with its value in previous steps and supposing that at each step the agent assesses whether \(A_i^t\) is positive or otherwise, as this could distort and amplify incentives for cooperation.
Note that a rule of “progressive disinvestment” is introduced: subjects gradually reduce their investments in the public good until they completely disappear, if they obtain two consecutive absolute negative results (loss on investment)\(^{19}\).

This rule is introduced to provide incentives for individual action, and appears to be consistent with what happens in the real world: subjects become more risk-averse when they meet persistent negative results in the short term, while they tend to stop “negative” investments when their perception of loss lasts over time, even where a long term view might favour holding on to an investment. The rule is intended to represent what typically happens on a stock exchange when, for example, investors prefer to sell stocks and shares that have made frequent losses over a short period.

From the point of view of endogenous mutation, i.e. of changes caused by the rules internal to the model, we register not only creation of groups but also group disruption: when a group finds itself with an average level of contributions below the constraint \(q\) (which means, in this case, when all the members of group are contributing zero), the group breaks up and the agents participating in the group become individual agents once again.

In the subsequent steps, the agents start to question their neighbours again and may form new groups. In each new round of questioning after group disruptions, the level of contribution of the agents concerned is no longer the random level assigned by function \(f\) at the start, but the last level of contribution to the public good made by the group before its breakup.

The next table summarizes the basic “working” of the model.

\(^{19}\) If they obtain two consecutive relatively negative results, but greater than zero in absolute terms, they do not disinvest but make further attempts to invert their action.
3. Results

The aim of this Paper is to design an agent-based model to study the evolution of cooperation among private agents taking part in collective projects. These agents are characterised by computational rationality. The question is the following: if the agents’ behaviour is not influenced by expectations both on the others’ behaviour and social factors, will they contribute to the public good, even where there is not a big difference between the rates of return on the public and the private good?

Three series of simulations have been carried out. The first series aims at verifying the main hypothesis, that is to establish whether agents having computational rationality, acting on the basis of rules that
take into account only updated individual payoffs at each step, can reach high levels of voluntary contribution.

I also test the sensitivity of the model to initial parameters (i.e. the algorithm allotting the individual characteristics and the rates of return on both the private and the public good investment).

The next series tests the model's robustness by means of an empirical validation test: it tests how the system responds to “noise”, that is the possibility of error and “confusion” (with regard to rules of behaviour) on the part of agents.

The last series introduces free riding into the model in order to test whether and to what extent opportunistic agents produce the failure of collective action and no production of public goods, as happens in experimental models and Game Theory.

3.1 Contribution levels, profit differentials and the heterogeneous nature of initial characteristics

The main result of the first series of simulations is that in the model there occurs an extremely high level of voluntary contribution to the public good. This is an interesting result in terms of both the final outcome and the dynamics through which it is reached.

The agents follow exclusively the rule according to which they assess the payoffs of the most recent rounds only and thereby “adapt” their next actions on the basis of the individual satisfaction. They go through a short period of attempts of action inversion; then, as a result of the higher financial payoffs that can be obtained from the public investment, they gradually and systematically synchronise, “directing themselves” towards the increase of contributions to the public good and, consequently, towards high average contribution levels.

Note however that the profit differential between public and private goods is not great and that the level of public profit assumed in the simulation model is lower than that used in experimental models: in fact, in experimental models the profit rate on public goods is usually twice that of private investment ($\alpha = 50\%$), and sometimes even higher. In my simulations the profit rate on public investment is only $30\%$ higher than that on private good ($\alpha = 30\%$).

I check whether different degrees of heterogeneity in the initial population may influence the outcome. Replications of the simulation showed that there exists a critical threshold of rate of return on the public good, $\alpha^*$ and when this threshold is exceeded, different initial distributions of characteristics do not affect the final result. In this series of simulations, $\alpha^* = 0.28$ and it is lower than the rate of return traditionally used both in Experimental Economics models and in my other simulations: this is useful in not excessively reducing the incentives to free riding and in assessing that high observed levels of cooperation are not twisted by a greater differential between the rates of return.
Therefore, with the purpose of giving virtual agents characteristics as similar as possible to those revealed \emph{a posteriori} in typical experiments on public goods, I restricted the support of the distribution function of the initial characteristics to the \([3 – 7]\) interval.

The main result of the model is robust and independent of initial parameters. Agents who are perfectly able to calculate incentives and payoffs without error or confusion, without altruistic or opportunistic motives, contribute to the supply of public goods at a level between 85 and 95 percent of their endowment.

Moreover this form of collective action does not decline over time, but rather it makes steady through a cyclical repetition: in each simulation, the system shows four-period cycles in which cooperation increases and decreases and, finally, after about 20 steps, becomes steady.

These cycles probably occur because there are lower and upper limits to what the agents can invest: no agent is allowed contribute with tokens worth less than zero or worth more than the maximum current endowment of ten units, so that cyclical limit states are created. This condition also tested in the validation tests in order to prove it does not affect the simulation results.

This result needs to be interpreted bearing in mind that the agents are involved in a process of asynchronic and double mutation. On the one hand, it is due to individual heterogeneity (i.e. the different characteristic distributed in the initial round that defines the agents’ capacity to contribute); on the other hand, it is also due to the temporary heterogeneity. Temporary heterogeneity is shown in the dynamics of communication and aggregation of agents so that they are in fact making decisions on different time planes.

At each moment there are some agents increasing their own contribution at the same time as other agents are reducing theirs: when the system reaches its stability, all agents are contributing around the 80% of their endowments and they do not subsequently modify this level.

Sensitivity tests were carried out by way of numerous simulations of the model, by changing individual characteristics and simulating the simplest model (with global interaction structure), that is the same model presented in the previous section, but with the agents not placed on a lattice, having no communication between each other, and having only to choose their levels of contribution to the public good contemporaneously and regardless of group membership.

In the first place, to study the impact of different distributions of initial agents’ characteristics on contribution levels, all possible uniform distribution functions were tested, that is to say, all possible combinations where the \(k_0\)’s are selected with a discrete probability in the interval \([0,10]\). I tested other intervals of distributions in other simulations, and I verified that the initial heterogeneity has a neutral effect on the outcomes.
The relation of the different distributions\(^{20}\) to the various levels of profit on the public investment \(\alpha\) was thus identified in order to test its potential influence on results.

Figure 1 shows that, above a certain threshold of profit on public investment, the model is not in fact sensitive to changes in the distribution parameter.

\emph{Fig. 1: Average contribution to the public good as a function of initial characteristics of the population and profit from public good.}

The graph shows systems of 1600 agents. In these simulations, after a short period of time (10-20 rounds), in which there occur “four” period cycles, the system synchronizes toward a high level of cooperation.

The average contribution to the public good is represented by colour and is obtained by calculating the average individual contribution once the system make itself steady.

The graph shows the critical threshold of profit on the public good, a function of the initial conditions, given by \((1 + \alpha^*) = 1.28\). Below this threshold the average contribution is practically zero (blue zone), while above it most agents invest in public goods with an average contribution of eight \((w_i^t - y_i^t) = 8\) (red zone).

Where the initial characteristics are uniform for all agents \((\Delta k_i^t = 0)\), the outcome is full cooperation\(^{21}\).

This is predictable given that the agents start the game at the same time with identical positive

\(^{20}\) For the purposes of verification, the variable defining variability in initial distribution was defined as follows:

\[\Delta k_i^t = k_i^{t_p} - k_i^{t_m}\] where \(k_i^{t_p}\) represents the maximum value of the distribution, 10, and \(k_i^{t_m}\) represents the possible difference with respect to the average value of 5 occurring from game to game.

\(^{21}\)
contributions and increase their levels of contribution each time until all their resources are invested in public goods.

All other initial configurations produce a curve in contribution defined by public profits. For levels between \((1 + \alpha^*) = 1.1\) and \((1 + \alpha^*) = 1.28\) it shows an increase in the average contribution as the variability in initial distribution \(\Delta k^j\) increases. Once the critical threshold \(\alpha^*\) is passed, there are almost full contribution levels to the public good.

It is interesting to look at the dynamics leading to this final state. Figure 2 shows the evolution over time of the average, minimum and maximum level of contribution.

Fig.2: Average contribution and minimum and maximum contributions to the public good (realisation of public goods).

The graph shows that during the transition period, the difference between minimum and maximum contributions declines rapidly until the point where all agents are “synchronised” and contemporaneously increase their contribution to the public good. When some of the agents reach the maximum possible contribution, \((w^j_i - y^j_i) = 10\), the synchronised system stabilises itself.

Note that in the first phase, where maximum contribution is falling, important changes in behaviour are caused by progressive disinvestment: when agents investing heavily in the public good repeatedly obtain negative results, they disinvest rapidly, thereby making the average contribution more uniform with less cooperative agents “obliged” to contribute.

\[^{21}\text{Only in the case of a uniform initial distribution with 0 for all } i, \text{ does the game simply end at the first step with complete absence of contribution to the public good.}\]
Lastly, note that in the final state there are different levels of contribution, between ten and six, which are all giving agents positive results.

Table 2 summarises my main findings.

### Tab. 2: Benchmark simulation in “neutral environment”

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support of the distribution function of $k_i$</td>
<td>3 - 7</td>
</tr>
<tr>
<td>$(1+\alpha)$ = profit from the public good</td>
<td>1.3</td>
</tr>
<tr>
<td>$(1+p)$ = profit from the private good</td>
<td>1</td>
</tr>
<tr>
<td>$q$ = technological constraint</td>
<td>0</td>
</tr>
<tr>
<td>% of the agents in groups</td>
<td>100%</td>
</tr>
<tr>
<td>Average contribution of group (as percentage of endowments)</td>
<td>85% - 95%</td>
</tr>
<tr>
<td>Time for reaching steady state</td>
<td>About 20 rounds</td>
</tr>
</tbody>
</table>

### 3.2 Levels of voluntary contribution and “errors” in the system

In the second series of simulations, I assessed the stability of results given that the $k_i$ are constrained to lie in the range [0, 10] due to the impossibility of negative investments and the limitation of individual resources.

I introduced random noise during some simulations: each agent $i$, with probability $\rho$, can make a random move $\Delta i$ which does not follow the standard procedural rules.

For low levels of noise, the results were confirmed and proved stable. But the noise pushed the critical threshold $\alpha^*$ somewhat higher (see Fig. 4).

Figure 3 shows the simulation with the same starting conditions as the first series but with random noise introduced with a probability of $\rho = 0.02^{22}$.

---

$^{22}$ Noise with probability of $\rho = 0.02$ means that 2% of the total population will make random errors in their actions.
Note that a low level of noise completely eliminates the peculiar behaviour occurring in the deterministic case of constant and uniform distribution (total cooperation).

With higher noise levels, with probabilities higher than $\rho^i = 0.05$, the results no longer depend on starting conditions and the critical threshold $\alpha^*$ rises noticeably to levels even higher than one.

In this simulation, for any initial distribution of characteristics, there tend to exist two distinct sets of solutions: no contribution and contribution by almost all agents. Figure 4 shows the incidence of different levels of noise on the results, that is the relationship between parameters $\rho^i$ and $\alpha^*$, and how cooperation increases as noise and the critical profit level rise.
The absence of cooperation is the blue area. Average levels of contribution to the public good between \((w_i^t - y_j^t) = 4\) and \((w_i^t - y_j^t) = 9\) are shown by blue–red shading. The diagram shows that as noise increases there is always a transition period when contributions tend to standardise, and subsequently a phase where the agents synchronise and systematically increase their contribution to the public good.

When the maximum level of contribution is reached, there is another four-period cycle during which the average contribution alternately rises and falls.

The following graphs show the complex dynamics resulting from an increase in noise and the wider variety in contribution behaviour. In particular, the graph shows the simulation where \(\alpha = 0.5\) and \(\rho = 0.01\). We can observe that there are agents who make very little contributions: 23 agents contribute only with one token and 8 agents are completely free riders.
Fig. 5: Average and maximum contribution to the public good as a function of time and with 2% probability of individual error.

Fig. 6: Average and maximum contribution to the public good as a function of time and with 7% probability of individual error.
3.3 Levels of voluntary contribution and free riding among agents

The third simulation aims at showing that opportunistic strategic behaviour is the true determinant of collective action failure, once factors of confusion and altruism are neutralised. I introduce a genotype of free riding which is exogenous and systematic.

A proportion of agents are characterised by “unconditioned” opportunistic behaviour, where “unconditioned” means *a priori* and exogenously set. These agents, whatever may happen during the simulation, defect from the collective goal and, anticipating its failure, try to obtain the highest possible individual benefit for themselves immediately.

A parameter $F_h$ is added to the model, at time $h$, to introduce a percentage of free riders exhibiting right from $h$ a gradually opportunistic strategic behaviour. $F_h$ is the percentage of agents in the system that, by time $h$ always decreases its contribution to the public good by one unit at every step. As in experimental economics models this leads to a gradual fall in contribution levels and eventually to the failure of collective action.

Figure 7 shows contribution to the public good as a function of the proportion of strategic agents $F_s$ and the yield of public investment.

Note that the behaviour dynamics of non-opportunistic agents are identical to those occurring in simulations without noise and with the critical rate of return on the public goods $\alpha^*$: these dynamics are also characterised by four-period cycles before synchronization.

By introducing $F$, the system exhibit a dynamics of cooperation only when $\alpha^*$ is higher than a critical threshold given by the identity $(1+\alpha^*)/(1-F) = 1,15$, that is

$$\alpha^* = 0,15-1,15 F$$
In this case it was possible to derive analytically the critical threshold of free riding agents which leads to system failure in relation to the profit rate on public investment.

The curve represents the theoretical phase diagram and can be analytically derived by considering that the system with $F_i$ strategic agents is equivalent to a system where $(1 - F_i)$ agents act according to the usual rules.

The critical threshold of the rate on public goods is lower than $\alpha^*$ and increases when the presence of non contributing strategic agents also increases.

Figure 8 shows what happens in the model when contribution to the public good is close to zero.
In this case, there is an initial transition period when the largest contributors gradually disinvest. But the low differential between profits from public and private investments prevents the system from synchronizing and the agents from increasing their contributions to the public good.

Table 3 summarises the main results of this series of simulations.

**Tab. 3: Simulation with opportunistic agents (hypothesis of free riding).**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support of the distribution function of $k'$.</td>
<td>$3 - 7$</td>
</tr>
<tr>
<td>$(1+\alpha) = \text{profit from the public good}$</td>
<td>1.3</td>
</tr>
<tr>
<td>$(1+p) = \text{profit from the private good}$</td>
<td>1</td>
</tr>
<tr>
<td>$q = \text{technological constraint}$</td>
<td>0</td>
</tr>
<tr>
<td>$F_h = % \text{ of free riding agents}$</td>
<td>20%</td>
</tr>
<tr>
<td>$% \text{ of agents in groups}$</td>
<td>0</td>
</tr>
<tr>
<td>Time for reaching the steady state</td>
<td>About 20 rounds</td>
</tr>
</tbody>
</table>
4. Conclusions

This Paper presents simulations obtained by means of an agent-based model where agents search solutions for collective action in order to provide a public good. My aim was to verify the hypothesis that computational rationality, in the absence of errors and conditioning, and taking exclusively into account the efficiency of own options in individual investment, can lead to the success of collective action.

The Paper illustrates the basic initial assumptions, the procedural rules regulating agent interaction and the preliminary results of a series of simulations.

My findings show a varied picture. There is a real possibility of collective action succeeding at significant levels where there is an absence of strategic behaviour and other factors influencing the results of experimental studies. The introduction of opportunistic behaviour on the part of some agents, however, leads to the same results as obtained in Experimental Economics: agents opt for selfish behaviour and collective action fails.

In both cases, a higher expected level of profit from collective action, or a higher differential between profits on public and private investments, can influence the pace of coordination and grouping dynamics. But under a certain threshold the profit differential does not appear to influence the level and spread of cooperation.
1. The simulation model

1.1 Individual Variables

A generic individual $i$ placed in the lattice is characterized by:
- pair of coordinates $(x,y)$ that determines their position;
- $i$ coordinate for individual identification;
- initial distribution in a contribution class;
- colour to identify the agents’ status;
- a Boolean coordinate “Isjoined”: true for agents aggregated in groups, otherwise false;
- $j$ coordinate identified by the aggregate that it is connected to ($j = 0$ if isolated);
- $d$ coordinate identified by the dimension of the aggregate that it is connected to ($d = 1$ if isolated);
- $p_i(t)$ coordinate equal to the average instant payoff at the time $t$;
- $P_i(t)$ coordinate equal to the cumulative payoff at the time $t$;
- action undertaken during the current step;
- part of the endowment invested in the private market;
- part of the endowment invested in the public good.

1.2 Group variables

A generic group is generated by:
- the union of two individuals;
- the consolidation of a group and an isolated individual;
- the consolidation of two groups.

A generic group is characterized by:
- $k$ coordinate identified by the aggregate;
- colour to identify their status;
- $p$, coordinate ($t$) equal to the instant payoff at the time $t$;
- value of the average allotment contribution.

1.3 Simulation parameters

The simulation is characterized by the following customizable user parameters:
- minimal technological constraints for the public good;
- private market return rate;
- public good return rate;
- parameter for saving data related to graphs.
Questionnaire post ACE study

1. Links with the literature

- Is your model based on some existing model in the simulation literature? □
- Is your model based on some existing model in the non-simulation literature? □
- Does the paper contain a survey on the theoretical background of the phenomenon that is investigated?
  - □ Long
  - □ Brief
  - □ None
- Does the paper contain a survey of the relevant simulation and non-simulation models?
  - □ Long
  - □ Brief
  - □ None

2. Structure of the model

- Have you clarified:
  - the goal of your model (empirical or theoretical) □
  - whether the implications are testable with real data □
  - the evolution of the population (static or dynamic)
    - if static: the total number of agents □
    - if dynamic: birth and death mechanisms □
  - the treatment of time (discrete or continuous) □
  - the treatment of fate (deterministic or stochastic) □
- Have you classified your model with respect to:
  - the topological space (no space, nD lattices, graphs…) □
  - the type of agent behaviour (optimising, satisfying..) □
  - the interaction structure (localized or non-localized) □
  - the coordination structure (centralized or decentralized) □
  - how expectations are formed (rational, adaptive or other) □

3. Analysis

• Have you clarified the objective of the analysis (full exploration or partial exploration)? □

• Have you clarified the focus of the analysis (equilibrium at micro level, equilibrium at macro-level, out-of-equilibrium)? □

• Has statistical testing of the properties found in the artificial data been performed? □

• Are the parameters of the model been estimated / calibrated on real data? □

• Has sensitivity analysis been performed? □

• Has validation been performed? □

4. Replicability

• Is the presentation detailed enough to allow the replication of the experiment/results? □

• Have you used a simulation platform to implement your model? □

• If any, have you clarified which simulation platform you have used? □

• Can the simulation be run online? □

• Graphical presentation of the model structure:
  □ UML diagrams (specify)
  □ Other diagrams (specify)
  □ None

• Code availability:
  □ Web-site
  □ Upon request
  □ None

• Is your exploration performed only on a subset of the parameters’ space? If yes, please state why.

• Which kind of statistical analysis have you performed on the artificial data?
  □ graphical
  □ descriptive statistics
  □ multivariate analysis (metamodelling)
  □ stationarity / ergodicity tests on artificial time series
  □ other (please specify)………………………………..
• If multivariate analysis / statistical tests have been performed, please list the methods you have used.

• Please put down all meaningful parameters that had to be initialized and indicate the method(s) you used for estimation or calibration (e.g. beta: calibrated/estimated from statistical data/empirical data collection). Please indicate a reference for each method.

• Please mark those features that you tested for sensitivity.
  □ Random seed variation
  □ Variation in the level of data aggregation
  □ Noise type and noise level variation
  □ Variation in the decision processes and capabilities of the agents
  □ Parameter variation
  □ Variation of sample size (esp. small sample properties)
  □ Temporal model variation (discrete to continuous time or from fixed to random updating of cells)
  □ Other: ........................................

• Please indicate the method(s) you applied for testing the model’s sensitivity on input variation (please give a reference for each method).

• Please state the type of validity that you claim for your model.

• Please indicate the method(s) you applied for testing the model’s validity (please give a reference for each method).

• Comments on this questionnaire
You have worked through this questionnaire that aims at increasing methodological rigour in agent-based social and economic simulation. Do you have any comments or recommendations for us that could improve this questionnaire?
Bibliography


