Dynamic Voluntary Contribution to a Public Good: Learning to be a Free Rider

Christiane Clemens[*] *und* *Thomas Riechmann*

Diskussionspapier Nr. 240

März 2001

ISSN 0949 – 9962
JEL Klassifikation: H41 – C73 – D83 – C6
Keywords: bounded rationality, evolutionary games, experiments, learning, public goods

[*]Universität Hannover — Institut für Volkswirtschaftslehre, Abteilung Wachstum und Verteilung, Königsworther Platz 1, 30 167 Hannover, e–mail: *clemens@vwl.uni-hannover.de* und *riechmann@vwl.uni-hannover.de* .

Abstract

Dieser Beitrag behandelt die Frage, ob eingeschränkt rationale Wirtschaftssubjekte optimales Verhalten im Rahmen der freiwilligen Bereitstellung eines öffentlichen Gutes lernen. Der Lernprozeß wird durch einen evolutionären Algorithmus abgebildet. Im Verlauf des Lernprozesses konvergiert die bereitgestellte Menge des öffentlichen Gutes gegen das Nash–Gleichgewicht, obwohl vollständiges Trittbrettfahrerverhalten nicht beobachtet werden kann. Damit spiegelt unser Ergebnis Resultate aus Experimenten zu öffentlichen Gütern wider. Sowohl die Gruppengröße als auch die individuelle Experimentierfreudigkeit sind zentrale Einflußfaktoren des Lernprozesses.

Abstract

This paper explores the question whether boundedly rational agents learn to behave optimally when asked to voluntarily contribute to a public good. The decision process of individuals is described by an Evolutionary Algorithm. We analyze the learning process of purely and impurely altruistic agents and find that in both cases the contribution level converges towards the Nash equilibrium although, with pure altruism, *exact* free rider–behavior is never observed. The latter result corresponds to findings from experiments on voluntary contribution to a public good. Crucial determinants of the learning process are the population size and the propensity to experiment.

1 Introduction

This paper explores the question whether boundedly rational individuals, who possess only a minimum of information about the structure of the economy, learn to behave optimally when asked to contribute to a public good. The decision process of agents is described by an Evolutionary Algorithm (EA), which belongs to the general class of adaptive learning algorithms.[1]

The question whether agents voluntarily participate in the provision of a public good has at length been studied theoretically as well as in experiments (Isaac *et al.*, 1985; Bergstrom *et al.*, 1986; Andreoni, 1988*b*).[2] A widely accepted prediction on individual behavior is the Nash conjecture, i. e. each person maximizes utility taking other people's behavior as given. The results from standard theory are clear. There are no incentives to reveal individual preferences truthfully, the agents try to free ride on the contributions of others, and the equilibrium provision level falls short of the Pareto–efficient one.

But how do individuals actually know that free riding is the best response? If we consider the results from experimental economics, we find that (a) there is no significant evidence of free riding in one–shot games, and (b) there is convergence towards free riding in repeated games although there is no *exact* free riding (see Marwell and Ames, 1981; Isaac *et al.*, 1984; Miller and Andreoni, 1991). So, there is good reason to conclude that standard theory neglects relevant aspects of individual behavior.

Sugden (1985) discusses non–Nash behavior, but finds that taking account of individual beliefs on other's contributions even aggravates the problem of free riding. Andreoni (1988*b*) studies strategic interaction and learning in public goods experiments. He finds neither of both strongly supported by the data, and ascribes the results to non–standard behavior such as altruism, social norms or bounded rationality.

Recently, especially incomplete information about individual actions, bounded rationality, and the necessity to learn optimal strategies has come into focus of economic theory.[3] Bliss and Nalebuff (1984), Fershtman and Nitzan (1991), and Gradstein (1992) argue that in a dynamic context an agent has the opportunity to learn the response of other players. Here, learning is modeled as *Bayesian learning*, where an agent has statistical information about preferences or donations of others which are periodically

[1] Evolutionary algorithms (EAs) are a family of simulation methods resembling the basic working principles of biological evolution. The probably best known type of EAs are Genetic Algorithms (Clemens and Riechmann, 1996; Riechmann, 2001*b*).

[2] For a survey see Ledyard (1995) and Cornes and Sandler (1996).

[3] For an extensive discussion see Evans and Honkapohja (1999).

updated. The results are ambiguous. Fershtman and Nitzan (1991) discuss a dynamic model where the agents have incentives to free ride on current and future contributions, whereas Marx and Matthews (2000) find Bayesian equilibria with and without completion of the public project.

In our view, the Nash and the non–Nash conjecture, and the models of Bayesian learning share an important shortcoming. Neither do the individuals possess perfect information nor are they attributed with such advanced information processing capacities, those theories require them to have. We drop the assumption that *"I know that everyone else knows, and everyone knows that each knows, and we all know that we know … "*, i. e. that every player in isolation has full knowledge of the relevant data and can costlessly figure out all equilibria. Instead, our analysis relies on the assumption of *boundedly rational* agents. Basically our model consists of a repeated single–shot game, where subjects report their individual valuation of the public good. It is a dynamic approach, because a single–shot game is not sufficient to allow subjects to learn the incentives, or following Andreoni (1988*b*, p. 292) *"Repetition appears to be necessary for subjects to approach free riding."* The agents are neither informed about the group size nor about preferences and endowments of their co–players. We think that in general this layout corresponds to the standard design of experiments.

The process of learning and strategy adaptation is modeled by an evolutionary algorithm. In each round, a typical agent of the population reports his willingness to pay, which is his strategic variable. The provision level of the public good is determined, and the individual receives his utility. The player then revises his strategy in order to adapt to the changing environment. This is done by means of *learning by imitation* and *learning by experiments* (Riechmann, 1999). We do not consider multi–period decisions, so there is no discounting of future payoffs or contribution delays.[4] Moreover, we do not discuss the design of incentive structures and mechanisms that encourage individuals to reveal their true preferences. Inevitably, the agents of our basic model end up with some kind of free rider–strategy. But, it is important to stress that they have learned these strategies despite of starting from a point of minimum information.

We contrast two preference specifications which allows for a separate treatment of the effects stemming from impure altruism: The first case is the standard framework where the agent does not receive additional utility from the *'act of giving'*. The second case captures the notion of receiving a *'warm glow'* from the donation to the public good (Andreoni, 1988*a*, 1990).

We draw from the model of Miller and Andreoni (1991), who describe

[4]It would be straightforward to extend the model in this direction but this is beyond the scope of this paper. The same argument applies to the discussion of heterogeneity or congestion phenomena.

free riding as an outcome of an evolutionary game. The replicator dynamics of the adaptive learning process strengthen those strategies over time that perform well. Their findings support results from experiments on voluntary contribution to a public good. In their simulations the provision of the public good converges towards exact free riding. Furthermore, they demonstrate that convergence is delayed with an increase in group size, which also replicates results from experiments.

Unfortunately, in their model, no strategy can be learned which differs from the ones already contained in the initial population. The purely imitative learning of replicator dynamics cannot develop any *new* strategies or recover those which have been wiped out. If free riding is not part of the initial set of strategies, it cannot be learned, and if the game starts with a homogeneous population, no convergence towards free riding will be observed. Miller and Andreoni (1991) cannot give a plausible explanation for an important result from free riding experiments, namely, that *exact* free riding is hardly observed. The evolutionary algorithm we use, provides results which are closer to the experimental findings.

Our analysis will proceed as follows. Section 2 provides a short review of the underlying assumptions and corresponding results of the static game of voluntary contribution to a public good with incomplete information. Section 3 derives general results for the intertemporal performance of strategies. Section 4 describes the basics of the model of EA–learning, analyzes the learning dynamics, and establishes a link between evolutionary game theory and the theory of evolutionary algorithms. We examine two major learning principles, i. e. learning by imitation and learning by experiments. In section 5, we discuss the simulations. Section 6 concludes and gives an outline for future research.

2 A Benchmark: Nash Equilibrium with Incomplete Information

Consider the standard model of voluntary contribution to a public good (cf. Varian, 1994; Cornes and Sandler, 1996, Ch. 6). Individual preferences and endowments are common knowledge, so there is no uncertainty regarding the strategies of other agents. Each individual's utility maximizing choice will depend on everyone else's. A Nash equilibrium is an allocation at which each individual's chosen contribution is a *best response* to the other's. Typically, the Nash equilibrium is not Pareto–efficient. The equilibrium level of the public good falls short of the socially optimal one, which — in analogy to Cornes and Sandler (1996) — is referred to as *systemic free riding*. In the particular case of quasi–linear preferences, the extent of free riding increases with the group size.

A natural extension to this basic model is to relax the requirements on individual knowledge and to allow for private information. It is easy to demonstrate that this adds to the problem of free riding such that *informational free riding* can be observed. We consider a simple model with $n \geqslant 2$ consumers, one private good x and one pure public good G.[5] Each agent is endowed with exogenous wealth w_i. The individual consumer participates in a Lindahl *tâtonnement* process and divides wealth between private consumption x_i, and a contribution to the public good $\theta_i G$. The individual cost share θ_i corresponds to the personalized Lindahl price. For simplicity, we normalize the price of the private good, the public good, and marginal costs for both goods to unity. The utility function of consumer i is $U(x_i, G)$ and takes on the quasi–linear form

$$U(x_i, G) = x_i + \beta_i \ln G, \tag{2.1}$$

which allows for the exclusion of income effects. The general form of $U(x_i, G)$ is known to all agents, while the individual valuation of the public good, $\beta_i > 0$, is private information. In what follows, we refer to β_i as the *true* value of the preference parameter to distinguish it from the actually reported value b_i. In a first step, we compute the individual cost shares θ_i and the public good level G for an arbitrary report b_i according to the preferences implied by (2.1). Then, in a second step, we determine the best–response functions b_i^* which constitute a Nash equilibrium.

Substitution of the budget constraint $w_i = x_i + \theta_i G$ for x_i enables us to write the consumer's maximization problem for an arbitrary report of b_i as

$$\max_G \quad w_i - \theta_i G + b_i \ln G, \tag{2.2}$$

with the corresponding first–order condition $b_i / G = \theta_i$. The contributions of different individuals are regarded to be perfect substitutes. For this reason it is straightforward to use an additive aggregator function for the total provision of the public good, that is

$$G = \sum_{i=1}^{n} b_i. \tag{2.3}$$

The cost share of agent i can be derived as follows

$$\theta_i = f(b_i, b_{-i}) = \frac{b_i}{b_i + \sum_{j \neq i} b_j}, \tag{2.4}$$

where the subscript $-i$ denotes the corresponding variables of agents other than i. Given b_{-i}, agent i knows how the own reported value b_i impacts

[5] The model is identical to the one discussed by Cornes and Sandler (1996, Ch. 7.4), except for the underlying preferences.

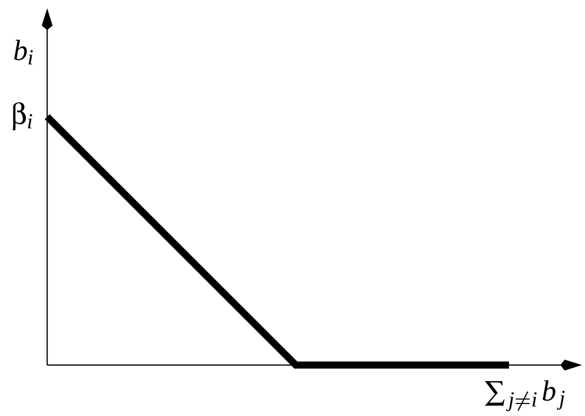

Figure1: *Best–response Function for Agent i*

on the Lindahl prices θ_i, θ_{-i} and the aggregate provision level of the public good, as given by (2.3). The consumer now chooses the best report b_i^* to maximize utility according to the true value β_i, while taking the reports of others as given

$$\max_{b_i} \quad w_i - \theta_i G + \beta_i \ln G. \tag{2.5}$$

Optimization leads to the following best–response function for agent i

$$b_i^* = g(\beta_i, b_{-i}) = \beta_i - \sum_{j \neq i} b_j \,. \tag{2.6}$$

The result is straightforward: Equation (2.6) indicates the well–known result that it is utility–maximizing for consumer i to significantly underreport the true value β_i. Figure 1 depicts the best–response function for agent i, where complete free riding occurs in case of $\beta_i = \sum_{j \neq i} b_j$. An announcement that preference data are used for the purpose of estimating Lindahl prices creates incentives to decrease voluntary contributions so as to understate true demands. In short, the agent has an incentive to misrepresent preferences, in order to get a lower personalized Lindahl price θ_i. Since this result applies to each of the n agents, the level of public good provision will be too low, if it is determined in accordance with the reported valuations. So, besides the systemic type of free riding, we additionally observe informational free riding.

The above described equilibrium is easily extended to a finitely repeated noncooperative game. The Nash prediction continues to hold, and each round is an exact replication of the one period game: Free riding (defect) is the best strategy. Although we may drop the assumption of full information with respect to the (possibly heterogeneous) form of $U(x_i, G)$ as well as perfect knowledge of β_{-i}, and w_{-i}, the agent of this model must at least be fully informed on the reports b_{-i} and the aggregation rule for G. Otherwise, he would not be able to compute the aggregate provision level of the public good and his cost share θ_i. So, despite the fact that we allow for incomplete information, we still demand a high amount of individual information processing capacities, especially when we deal with large communities or allow for repeated games.

The best–response function for the impure altruism set–up can be derived analogously. In this case we use a Cobb–Douglas type utility function

$$U(x_i, g_i, G) = x_i + \beta_i \ln\left(g_i^{\alpha} G^{1-\alpha}\right), \tag{2.7}$$

where $g_i = \theta_i G$ and $\alpha, 1-\alpha$ denote the weights, the agent puts on the private and the public part of his contribution. Equation (2.7) reduces to a private good model in case of $\alpha = 1$. The Nash–strategy of this framework is more difficult to learn than the simple linear strategy discussed above: On the one hand, we have a positive direct effect of g_i on utility, while on the other hand a utility maximizing choice of G implies a low individual contribution. Each of the agents has to learn to balance these two effects.

The repeated one–shot game will be the starting point for our learning model of the next chapter, but with a major difference. To capture the notion of *bounded rationality*, the agents of the evolutionary model are provided with as few information as possible regarding the structure of the game. They possess no knowledge with respect to group size, preferences, the aggregator function for the public good etc. We will demonstrate, that the agents learn the noncooperative strategy even with hardly any presupposition on individual knowledge.

3 Dynamic Performance of Strategies

We consider the model as introduced above with a population of n agents. For simplicity, we assume that all agents are identical with respect to their preferences and endowments, that is $\beta_i = \beta$ and $w_i = w$, $\forall\, i = 1, \ldots, n$. Moreover, we posit preferences, endowments, and the size of the population to be unchanging over time.[6] In each period individual i receives an endowment $w(t)$ and chooses a report $b_i(t)$ from the set of feasible contributions $\mathcal{B} = \{b \,|\, 0 \leqslant b \leqslant w\}$.[7]

[6]Learning in general does not occur in isolation but always includes learning from others. In assuming identical agents, we avoid the discussion on the quality of adopted strategies, which have been copied from other agents, who are totally different.

[7]For an easier understanding, imagine we assumed discrete contribution levels. Then $\mathcal{B}$ is the set of $h = 1, \ldots, k$ contribution levels.

Learning now involves that — as time goes by — the agent replaces poorly performing strategies by those performing well. But what makes one strategy perform better than another, and how does this information spread throughout the community? To answer this question, we will now focus on the dynamic evolution of an arbitrary strategy of time t, say the contribution level $b^m(t)$. Consider a group m of agents who decide to report the identical contribution $b^m(t) \in \mathcal{B}$. The time–t population share of type–m agents will be denoted with $q^m(t)$. The utility of a group–m agent signaling a contribution $b^m(t)$ in the standard case (2.1) is given by

$$U^m(t) = w(t) - b^m(t) + \beta \ln G(t). \tag{3.1}$$

with $G(t)$ as given by (2.3). Equation (3.1) displays the well–known result that the agent's utility is not only affected by his individual action $b^m(t)$. It is additionally determined by the strategic decisions of all members of the population which are reflected in the aggregate contribution level of the public good. Since $U^m(t)$ is a state–dependent variable, it changes over time as the agents revise their strategies.[8]

We are now interested in the performance of an arbitrary type of strategy, say $b^m(t)$ of period t. The quality of the performance of a type–m strategy can only be estimated in relation to the performance of the entire set of strategies which are actually played in the population. Hence, the quality of a strategy $b^m(t)$ is measured by individual payoff (3.1) relative to the average utility at time t, the latter given by

$$\overline{U}(t) := \sum_{m|b^m \in \mathcal{B}} q^m(t)\, U^m(t)\,. \tag{3.2}$$

The dynamic evolution of the population share $q^m(t)$ of agents playing $b^m(t)$ can then be described as

$$q^m(t+1) = q^m(t)\, \frac{U^m(t)}{\overline{U}(t)}\,. \tag{3.3}$$

The population share of of type–m agents in the next period, $q^m(t+1)$, is determined by two factors: First, by the population share of type–m agents in the current population, $q^m(t)$, and second, by the ratio of the utility of a type–m agent of the current population to the average utility of agents in period t, $\overline{U}(t)$. By (3.3), the growth rate of the population share of type–m agents can be derived as follows

$$\frac{q^m(t+1) - q^m(t)}{q^m(t)} = \frac{U^m(t) - \overline{U}(t)}{\overline{U}(t)}\,. \tag{3.4}$$

[8]Miller and Andreoni (1991) used a quite similar setup to run simulations based on replicator dynamics.

This is a typical representation of replicator dynamics in evolutionary game theory (Vega-Redondo, 1996, p. 44). Consider a strategy $b^m(t)$ that performs well in period t. The better this strategy, the more the payoff $U^m(t)$ exceeds the average payoff, and the higher the growth rate of the population share $q^m(t)$ of agents playing strategy m will be. The share of the population which plays any given strategy changes in proportion to its relative utility. In plain words, the success of strategy $b^m(t)$ is observed by agents who played a poorly performing strategy, say $b^k(t)$. Those agents will probably adopt the superior strategy and report a next–period willingness to pay of $b^m(t+1)$, which exactly reflects the dynamics of how better strategies spread throughout the population.[9]

Now, which strategies will perform well and which will not? In general, an agent perceives his impact on G as negligible. The direct utility loss of announcing *and* paying a comparably high $b_i(t)$ will exceed the indirect utility loss of the induced decrease in the provision level G. Since the subject of our model takes no account of other agents' strategic conjectures, he fares well with a low contribution. This argument, together with (3.4) predicts the dynamics to converge towards free riding behavior. So, our predictions on the performance of individual strategies closely resemble the Propositions 1. − 3. of the evolutionary game with replicator dynamics discussed by Miller and Andreoni (1991).

Yet, the most severe problem with replicator dynamics is that the concept relies on the assumption of an infinitely large population or, put differently, a 'continuum strategy' (Fudenberg and Levine, 1998). If populations are of finite size, equation (3.3) only displays the expected dynamics of the game. An attempt to derive the number of agents actually playing a specific strategy $b^m(t), b^k(t)$ from the population shares $q^m(t), q^k(t)$, would in any case of a finitely–sized population most probably result in something like *17.3* people signaling a demand of $b^m(t)$, *12.4* people announcing $b^k(t)$, *et cetera*. We will demonstrate below that the population size is a key determinant of the exact time path of population dynamics. So it is not convincing to work with a concept that presupposes infinitely large populations. Instead, our model relies on simulations of evolutionary algorithms (EAs), uniting the best of two worlds: On the one hand, EAs rely on finitely large populations. On the other hand, the dynamic characteristics generated by the type of EA used in this paper are quite similar to the of replicator dynamics given by (3.3).

[9] For details on strategy adoption, see section 4.

4 Learning Dynamics

Learning Concepts At this point we want to give a short non–technical sketch of the two learning concepts *learning by imitation* and *learning by experiments* (Riechmann, 1999), before going into technical details of the learning dynamics. The first step in imitation learning is that a type–m agent employs the information on individual strategies and corresponding utilities available to him in order to evaluate the success of his own strategy $b^m(t)$. At the end of each round of the game, the individual compares his resulting utility $U^m(t)$ to the utility $U^k(t)$ of other players, who played strategy $b^k(t)$. In a second step, he decides for period $t+1$ on whether or not to adopt a different strategy which has just proved to be promising. At this point it is important to stress that no new strategies are generated within this learning concept.[10] Those *new* strategies come into the population via experiments. The term *experiment* stands for a small–scale change in the strategy. The subject adds (subtracts) a small amount to (from) the chosen next–period contribution $b^m(t+1)$ before entering the new round of the game. This increases the variety in the pool (distribution) of available strategies and is the only way out of strategic lock–ins.

The evolutionary algorithm we use to model the learning process does not deviate much from the general design of genetic algorithms (Holland, 1975, 1992; Goldberg, 1989). In fact, the major difference is its real–valued coding. The algorithm consists of two genetic operators: selection / reproduction and mutation, where the first one reflects *learning by imitation* and the second one reflects *learning by experiments*.[11]

In the following, we separately discuss the consequences of the two learning concepts for the dynamics of strategy choice and the resulting provision level of the public good. Moreover, we focus on the similarities between evolutionary game theory and the theory of evolutionary algorithms (Riechmann, 2001*a*). We show that the dynamics generated by replicator dynamics are formally equivalent to the expected selection / reproduction dynamics of an EA.

Learning by Imitation The basic operator of the EA employed here is the operator of selection and reproduction stemming from the canonical genetic algorithm (CGA) introduced by Holland (1975, 1992) and Goldberg (1989).

[10] Moreover, this concept may even lead to absurd learning dynamics, if we assume agents who are heterogeneous with respect to their true preferences β_i. A strategy which in this context is successful for one person might as well perform poorly for another and hence is not worth to be copied.

[11] There is only one value coded in each round of the algorithm, namely the value of the strategy $b_i(t)$. Consequently there is no necessity for crossover.

At each time t, the population of the evolutionary algorithm consists of the strategies $b_i(t) \in \mathcal{B}$ played by the n agents of the economy. Note that in the terminology of genetic algorithms, here the population is not the set of agents but the set of strategies.

Members from a time–t population are transferred into the next–period one by means of selection and reproduction. The procedure is conducted a follows: First a strategy is drawn randomly (with replacement) from the old population and copied into the new one. Each strategy has a certain probability of being selected and reproduced, which depends on the quality of performance in the previous round of the game. This introduces state–dependency into our model. We use the *relative fitness* as a quality index of a strategy $b^m(t)$, which is defined as the ratio of the agent's fitness $U^m(t)$ to the *aggregate fitness* of the whole population $U(t) = \sum_{m|b^m \in B} N^m(t) U^m(t)$. $N^m(t)$ denotes the number of agents currently using strategy $b^m(t)$. We assume that the selection and reproduction probability equals the agent's *relative fitness*, which is a standard assumption in theory of genetic algorithms.

$N(t+1)$ is the total number of strategies in population $t+1$. Accordingly, the number of agents using strategy $b^m(t+1)$ in period $t+1$, $N^m(t+1)$, is given as

$$N^m(t+1) = \frac{U^m(t)}{U(t)} N^m(t) N(t+1). \tag{4.1}$$

In the spirit of replicator dynamics of evolutionary game theory, (4.1) can be translated into a notation which explicitly shows the development of the fraction of agents who play the respective strategy. Obviously, the share $q^m(t)$ of agents playing strategy $b^m(t)$ in period t and the average utility $\overline{U}(t)$ are defined as

$$q^m(t) := \frac{N^m(t)}{N(t)} \quad \text{and} \quad \overline{U}(t) := \frac{U(t)}{N(t)}. \tag{4.2}$$

Using (4.2), it is now easy to show that, for a selection / reproduction scheme according to relative fitness, the expected growth rate of a population share q^m can be written as

$$\mathrm{E}\left[\frac{q^m(t+1) - q^m(t)}{q^m(t)}\right] = \frac{U^m(t) - \overline{U}(t)}{\overline{U}(t)}. \tag{4.3}$$

Notice, that (4.3) is structurally equivalent to (3.4), with the only difference, that there is no randomness in ordinary replicator dynamics. This establishes a link between evolutionary algorithms and evolutionary game theory. By the *law of large numbers*, the selection / reproduction dynamics of EAs converge towards the replicator dynamics of evolutionary game theory with an increase in population size. Thus, for finite populations, the process of selection / reproduction in evolutionary dynamics can be regarded as a good first approximation of replicator dynamics.

The agent based simulation of the problem presented here will contain the operator of CGA selection / reproduction as described above. We already mentioned in the beginning that the behavioral interpretation of this operator is the one of learning by imitation: Agents who recognize that their strategy was not as successful as the strategies of others or, alternatively was below average, drop their strategy and imitate another.

From the formal analysis we know that the more successful a strategy was in the last period, the more likely it is to be imitated in the current one. This interpretation also holds for replicator dynamics, where it is usually called *evolutionary learning*. But this result does not predict, that the speed of learning increases while superior strategies spread throughout the society, because the expected growth rate of the fraction $q^m(t)$ of people playing $b^m(t)$ itself depends on the evolution of the entire population.

Yet, learning by pure imitation suffers from a severe drawback (Riechmann, 1999): Strategies not contained in the very first population can never be learned, and strategies wiped out throughout the learning process can never be recovered. This implies that learning by pure imitation is a process which is highly path–dependent. In contrast to replicator dynamics, here, learning by imitation includes the case that even superior strategies may die out due to the random element in the selection / reproduction operator. From this follows immediately that the learning process might eventually *lock in* at uniform contribution levels $b > 0$ which are far from the optimal free riding–level of the Nash–strategy $b = 0$. This phenomenon is usually called *genetic drift*.

Learning by Experiments An individual learning process which purely relies on imitation of other agents' strategies is indeed not a very convincing learning concept, since it completely ignores the creative part of learning. For this reason we introduce the mutation operator as a means of *learning by experiments*.

The agent chooses a preliminary contribution level for the public good. We denote this preliminary level with $\tilde{b}^m(t)$. For an easier understanding, imagine that selection / reproduction has already taken place and the preliminary level $\tilde{b}^m(t)$ is the starting point for the development of a new strategy $b^m(t)$. The value $\tilde{b}^m(t)$ is then subject to a refinement. A small amount is added to the individual contribution for the public good, or subtracted respectively. The resulting strategy $b^m(t)$ is the one, the agent pretends to be his true demand for the public good in the next iteration of the game.

This experimental process is modeled in the simulations as follows: The final strategy $b^m(t)$ is derived by adding a term $\varepsilon_i(t)$ to the agent's preliminary time–t strategy, $\tilde{b}^m(t)$, where $\varepsilon_i(t)$ is a random number drawn from a Gaussian distribution with zero mean and a finite variance σ^2:

$$b^m(t) = \tilde{b}^m(t) + \varepsilon_i(t), \qquad \text{with} \quad \varepsilon_i(t) \sim \mathcal{N}(0, \sigma^2)\,. \tag{4.4}$$

By using a Gaussian distribution for the *size of the experiment*, $\varepsilon_i(t)$, we sustain two plausible behavioral facts: Experiments are equally likely in any direction from the original strategy, and small experiments are more likely than big ones. We interpret the mutation variance σ^2 as the *propensity to experiment* (Riechmann, 2001*b*). It is a measure for the extent of experiments an agent is willing to undertake.[12] For convenience, the σ^2 is assumed to be identical for all agents and constant over time in our simulations.

With the introduction of mutation we simultaneously established a way out of the local optimum or *lock in*–dilemma discussed in the previous section. With a positive propensity to experiment, there is always a positive probability that local optima will be left in finite time, and that already vanished superior strategies may be regained.

5 Simulations

So far, what do we know about the learning of optimal strategies in the case of voluntary contributions to a public good? The theoretical analysis of the preceding sections has shown that (a) superior strategies have better chances to survive and spread throughout the society, (b) non–astonishingly, a superior strategy is to pay less in the standard public good model, (c) purely imitative learning does not necessarily lead to the Nash–strategy $b = 0$, and (d) the learning concepts act together in such a way, that an individual potentially may learn to free ride, even if it was totally unacquainted with free riding behavior in the outset.

Those theoretical observations now have to pass the empirical test. We performed numerical simulations to underline our findings and were especially interested in the following questions: First, how does the size of the population impact on the learning process? Second, how does the propensity to experiment affect the convergence properties of the learning process? Third, how does impure altruism affect the learning process? Forth, to what extent does convergence towards free riding behavior occur?

The simulations are based on the evolutionary algorithm as described in the previous section. In order to derive results regarding the sensitivity of the learning process with respect to the population size, we performed

[12]Due to the real–valued coding of the EA, mutation differs substantially from the *flipping the bit*–procedure of the binary–coded CGA.

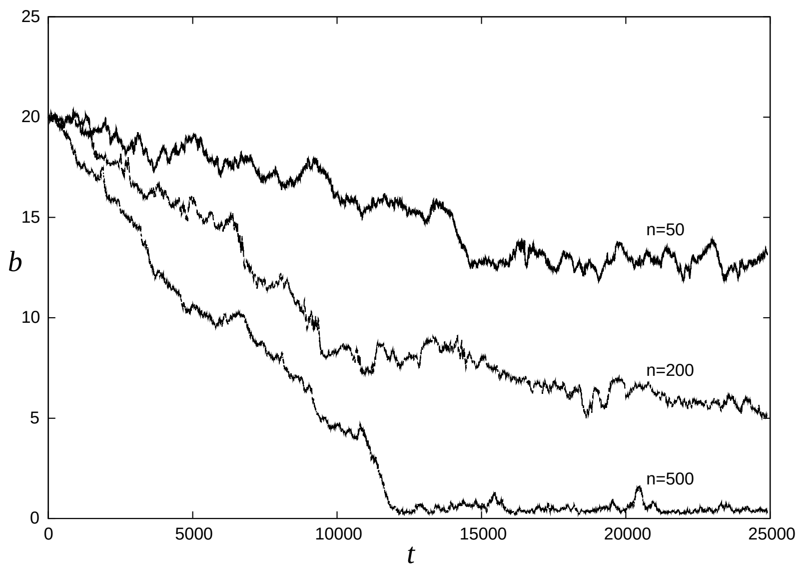

Figure2: *Convergence and Population Size: Fixed initial Population*

simulations with $n = 50$, $n = 200$ and $n = 500$ agents respectively. Remember that the population size n does not explicitly appear in the agent's problem (3.1). It is implicitly part of of the aggregate provision level of the public good $G(t)$, which is not subject to individual optimization. Of course, the population size does not affect the true valuation β or initial endowments w. Those parameters were assumed to be constant and identical across agents. The parameter settings are $w = 100$, and $\beta = 20$, which corresponds to a true report of $b = 20$. The propensity to experiment is set constant with $\sigma^2 = 0.03$.

Experimental results suggest that there is no significant free riding in the first period of a repeated single–shot game (Marwell and Ames, 1981). For this reason we assume our agents initially to be *naive* or some kind of good–natured. They reveal their preferences truthfully in the initial period, so we give them a reason for learning to be selfish. With identical agents, our initial population is uniform with $b_i(1) = 20$,$\forall i = 1, \ldots, n$.[13]

Group Size Figure 2 displays the society's average contribution to the public good for three simulation runs with $n = 50$, $n = 200$ and $n = 500$ agents. Independent of the population size, the contributions converge towards the free riding level $b = 0$ in the ultra long–run. Nevertheless, it becomes obvious from Figure 2 that convergence requires a considerable amount of time. In fact, the main reason for this is the initially homogeneous popu-

[13]At this point recall our critique on replicator dynamics, where a initially homogeneous population would not learn anything.

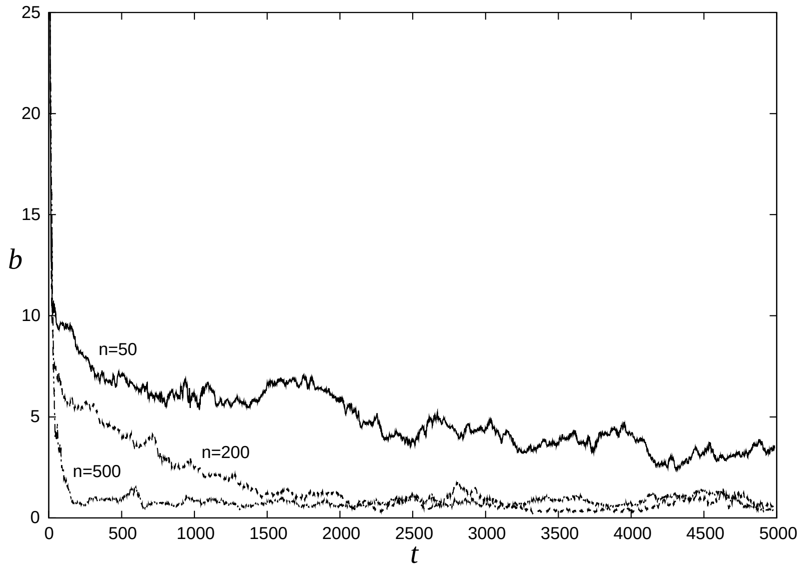

Figure3: *Convergence and Population Size: Random initial Population*

lation of strategies $b_i(1) = 20$ which does not give way to imitative learning. Learning progress can only be achieved by experiments. Once there is some heterogeneity in the population, imitation is effective and improves the learning process.

Figure 3 displays the simulation runs for an identically specified model, except for the initial population. In this case the initial population of strategies is randomized. The initial contributions $b_i(1)$ are i. i. d. and restricted to the interval $b_i(1) \in (0, w)$. The effects of imitative learning being immediately effective are obvious. Convergence occurs within a fraction of time required by the uniform–population algorithm. Our results are a direct application of the *Fisher principle* (Metcalfe, 1994; Birchenhall *et al.*, 1997) which says that learning occurs faster the more heterogeneous a population is.

Propensity to Experiment Another result displayed in Figure 2 as well as in Figure 3 is that larger populations converge considerably faster than smaller ones. This effect can primarily be attributed to the propensity to experiment. The chance (i. e. probability) of finding an agent, who carried out a comparably large and successful experiment is greater in a large population than in a small one. Thus, we expect the effects of experimentation to be more significant in larger populations. This outcome contradicts the results from laboratory experiments on voluntary contributions to a public good (see Miller and Andreoni, 1991, and references therein), where the speed of convergence was found to be inversely related to group size.

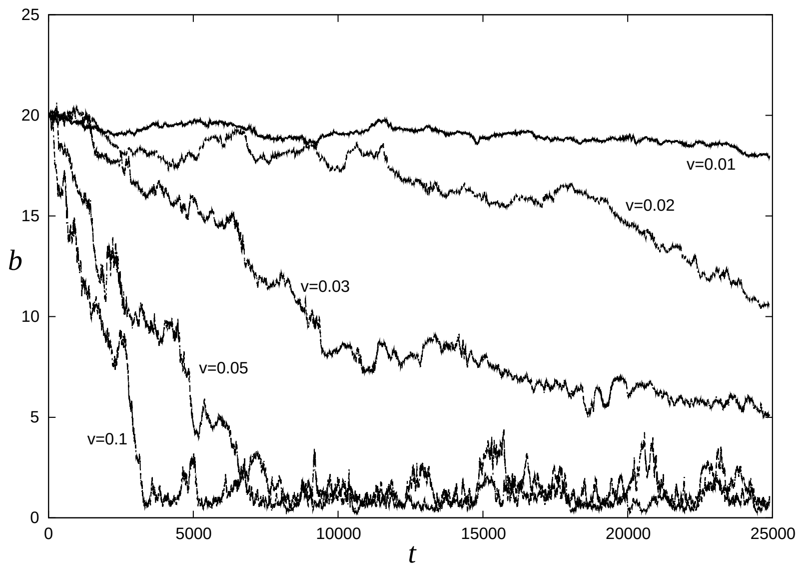

Figure4: *Convergence and Propensity to Experiment*

How can these contradictory results be explained? We believe that additional factors have to be taken into account. The group size ranged from 4 to 100 agents in the laboratory experiments cited above. In such small groups, an agent receives a comparably stronger feedback to his contribution than in a larger ones. The effect of his own actions on the aggregate provision level of the public good G is no longer negligible. For this reason, we think that the agent might be inclined to respond stronger to payoff signals by undertaking more or bigger experiments in order to find out the optimal strategy. We suggest that in small groups the propensity to experiment, σ^2, should be higher than in large ones. Figure 4 displays the society's mean contribution level for an initially homogeneous population of size of $n = 50$, with β and w as specified above. We let the mutation variance vary according to the following values $\sigma^2 = \{0.01, 0.02, 0.03, 0.05, 0.1\}$. Figure 4 shows the expected result, that a high propensity to experiment accelerates the convergence process.

Impure Altruism The qualitative result we expect from the assumption of exclusive utility received from individual donations is: A Nash equilibrium is characterized by a situation where each agent significantly offers a positive contribution to the public good. The amount crucially depends on the preference weight α. We simulated the learning process with a group of size $n = 200$, a comparably high propensity to experiment $\sigma = 0.05$, and let the preference parameter vary according to the values $\alpha = \{0.25, 0.50, 0.75, 1\}$ which corresponds to the Nash equilibria $b = \{5, 10, 15, 20\}$. Truthful rev-

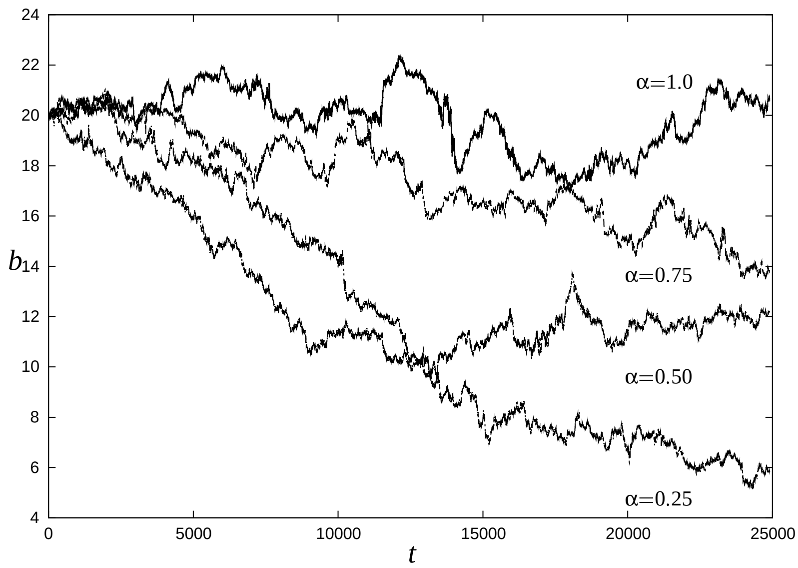

Figure5: *Impure Altruism*

elation of β is the dominant strategy, if the contribution is a private good. The simulation results in Figure 5 show, that obviously free riding strategies are learned faster. For small values of α, the agents does not put too much weight on the *'warm glow of giving'*. Hence, the principles of free riding are learned quite fast. The utility gains from free riding (small b) and the utility gains from the contribution (large b) start to offset each other, as α increases. This slows down the learning process and leads to deviations from the Nash–strategy which may last a considerable amount of time.

Extent of Convergence We finally turn to the last question, namely, to what extent convergence of the learning process occurs. From the theoretical solution we know, that especially in large groups, the optimal individual level of contribution is $b_i = 0$. From laboratory experiments we know that *exact* free riding is hardly ever observed. Our simulations replicate the results from experimental economics and thus differ from the evolutionary learning dynamics discussed by Miller and Andreoni (1991). Our results can mainly be ascribed to the fact that the agents of our model never stop to experiment.[14] So, what we derive is pretty close to the Nash–solution: The subjects learn the principles of free riding but never stop to look for a better strategy.

[14] Astonishingly experimentation does not even stop if the mutation rate is endogenized. First results on Meta–mutation indicate that the propensity to experiment does not converge to zero as the learning process approaches the free riding solution.

6 Conclusions

In this paper, we explored the question whether or not boundedly rational agents learn optimal strategies over time, when requested to voluntarily contribute to a public good. Our analysis relied on the standard Nash–Cournot approach of public economic theory that predicts free riding behavior, which — in the special case discussed in this paper — even increases with the group size.

The agents of our model played a repeated one–shot game and were endowed with only a minimum of information regarding the structure of the economy. The learning process was modeled by means of an evolutionary algorithm and analytically decomposed into two learning mechanisms: *learning by imitation* and *learning by experiments*. We demonstrated that the first concept in expectation equals the replicator dynamics discussed in evolutionary game theory.

Simulations support our major theoretical result, that boundedly rational individuals actually learn the principles of free riding. Better strategies are adopted over time and the provision level of the public good converges towards zero. The convergence speed is affected by the size of the population and the propensity to experiment, both positively correlated to learning progress. The learning process never comes to a rest in a situation of *exact* free riding, which reproduces results from laboratory experiments. This finding is due to the fact, that the agents of our dynamic model never stop to experiment in order to find a better strategy. In our view, this is an intuitively plausible and realistic description of individual behavior. Learning is slowed down, if the initial population is homogeneous. In this case, imitative learning is ineffective and the learning progress is mainly achieved by means of experiments.

We found that impure altruism slows down the learning process. The result can be ascribed to the counterworking direct and indirect way the individual contribution affects utility.

We discussed a rather elementary dynamic learning model which by intention was close to the standard game–theoretic approach. A straightforward extension to this model would be to allow for more heterogeneity regarding preferences or endowments. Another noteworthy extension is to discuss the effects of discrete public goods on the learning process. In our model, a positive amount of the public good was always provided as long as an agent decided to offer a positive contribution. The strategic situation changes substantially if at least a number k of n agents must announce a positive contribution to the public good, which otherwise would not be provided (Gradstein, 1992; Dixit and Olson, 2000).

References

Andreoni, James (1988*a*), Privately Provided Public Goods in a Large Economy: The Limits of Altruism, *Journal of Public Economics*, **35**, 57–73.

— (1988*b*), Why Free Ride? Strategies and Learning in Public Goods Experiments, *Journal of Public Economics*, **37**, 291–304.

— (1990), Impure Altruism and Donations to Public Goods: A Theory of Warm–Glow Giving, *Economic Journal*, **100**, 464–477.

Bergstrom, Theodore, Blume, Lawrence, and Varian, Hal (1986), On the Private Provision of Public Goods, *Journal of Public Economics*, **29**, 25–49.

Birchenhall, Chris, Kastrinos, Nikos, and Metcalfe, J. Stan (1997), Genetic Algorithms in Evolutionary Modelling, *Journal of Evolutionary Economics*, **7** (4), 375–393.

Bliss, Christopher and Nalebuff, Barry (1984), Dragon–Slaying and Ballroom Dancing: The Private Supply of a Public Good, *Journal of Public Economics*, **25**, 1–12.

Clemens, Christiane and Riechmann, Thomas (1996), Evolutionäre Optimierungsverfahren und ihr Einsatz in der ökonomischen Forschung, Diskussionspapier 195, Universität Hannover, Fachbereich Wirtschaftswissenschaften, Hannover.

Cornes, Richard and Sandler, Todd (1996), *The Theory of Externalities, Public Goods and Club Goods*, Cambridge University Press, Cambridge, 2nd edn.

Dixit, Avinash and Olson, Mancur (2000), Does Voluntary Participation Undermine the Coase Theorem?, *Journal of Public Economics*, **76**, 309–335.

Evans, George. W. and Honkapohja, Seppo (1999), Learning Dynamics, in: J. B. Taylor and Michael Woodford (eds.), *Handbook of Macroeconomics*, North–Holland, Amsterdam, pp. 449–542.

Fershtman, Chaim and Nitzan, Shmuel (1991), Dynamic Voluntary Provision of Public Goods, *European Economic Review*, **35**, 1057–1067.

Fudenberg, Drew and Levine, David K. (1998), *The Theory of Learning in Games*, MIT Press, Cambridge/MA.

Goldberg, David E. (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison–Wesley, Reading.

Gradstein, Mark (1992), Time Dynamics and Incomplete Information in the Private Provision of Public Goods, *Journal of Political Economy*, **100** (5), 581–597.

Holland, John H. (1975), *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor.

— (1992), *Adaptation in Natural and Artificial Systems*, MIT Press, Cambridge/MA.

Isaac, R. Mark, Walker, James M., and Thomas, Susan H. (1984), Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations, *Public Choice*, **43**, 113–149.

Isaac, R. Mark, McCue, Kenneth F., and Plott, Charles R. (1985), Public Goods Provision in an Experimental Environment, *Journal of Public Economics*, **26**, 51–74.

Ledyard, John (1995), Public Goods: A Survey of Experimental Research, in: John H. Kagel and Alvin E. Roth (eds.), *The Handbook of Experimental Economics*, Princeton University Press, Princeton, pp. 111–194.

Marwell, Gerald and Ames, Ruth E. (1981), Economists Free Ride, Does Anyone Else?, *Journal of Public Economics*, **15**, 295–310.

Marx, Leslie M. and Matthews, Steven A. (2000), Dynamic Voluntary Contribution to a Public Project, *Review of Economic Studies*, **67**, 327–358.

Metcalfe, J. Stan (1994), Competition, Fisher's Principle and Increasing Returns in the Selection Process, *Journal of Evolutionary Economics*, **4**, 327–346.

Miller, John H. and Andreoni, James (1991), Can Evolutionary Dynamics Explain Free Ridings in Experiments?, *Economics Letters*, **36**, 9–15.

Riechmann, Thomas (1999), Learning and Behavioral Stability — An Economic Interpretation of Genetic Algorithms, *Journal of Evolutionary Economics*, **9**, 225–242.

— (2001*a*), Genetic Algorithm Learning and Evolutionary Games, *Journal of Economic Dynamics and Control*, **25**, 1019–1037.

— (2001*b*), *Learning in Economics. Analysis and Application of Genetic Algorithms*, Physica–Verlag, Heidelberg.

Sugden, Robert (1985), Consistent Conjectures and Voluntary Contributions to Public Goods: Why the Conventional Theory does not Work, *Journal of Public Economics*, **27**, 117–124.

Varian, Hal R. (1994), Sequential Contributions to Public Goods, *Journal of Public Economics*, **53**, 165–186.

Vega-Redondo, Fernando (1996), *Evolution, Games, Economic Behavior*, Oxford University Press, Oxford.