JENA ECONOMIC RESEARCH PAPERS

2011 – 032

On the Evolution of Preferences

by

Astrid Gamba

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

Jena Economic Research Papers 2011 - 032

On the Evolution of Preferences*

Astrid Gamba†

Jun 2011

Abstract

A common feature of the literature on the evolution of preferences is that evolution favors nonmaterialistic preferences only if preference types are observable at least to some degree. We argue that this result is due to the assumption that in each state of the evolutionary dynamics some Bayesian Nash equilibrium is played. We show that under *unobservability* of preference types, conditional on selecting some *self-confirming equilibrium* as a rule for mapping preference into behavior, non-selfish preferences may be evolutionarily successful.

JEL classification: A13, C72, D64, D83.

Keywords: evolution of preferences, altruism, learning, self-confirming equilibrium.

*I would like to thank Pierpaolo Battigalli, Topi Miettinen, Matthias Messner and Werner Güth for useful suggestions. The usual disclaimer applies.

†Max Planck Institute of Economics, Jena, Germany. Contact: gamba@econ.mpg.de. Telephone number: ++49 3641 686634; Fax number: ++49 3641 686623

1 Introduction

Most of the literature which studies the evolutionary foundations of other-regarding preferences adopts the *indirect evolutionary approach* (Güth and Yaari, 1992; Güth, 1995; Güth and Kliemt, 1998). While the standard approach in evolutionary game theory consists in investigating whether a *strategy* is robust to evolutionary selection, the indirect evolutionary studies are about whether certain *preferences* are evolutionarily successful. Suppose there is a heterogeneous population composed of individuals with different preference types (e.g., altruistic and selfish). From this pool, agents endowed with their specific preferences are recurrently and randomly drawn to play a *basic game with material payoffs* (e.g., monetary payoffs). In each round, players behave rationally, maximizing their expected utility associated to their own preferences (which does not necessarily match the underlying expected material payoff). The evolutive success of a certain preference type is evaluated on the basis of the material payoffs (*objective fitness*) induced by the profile of strategies adopted. Agents whose preferences lead to higher material payoffs (higher fitness) tend to reproduce faster than those with lower material payoffs (lower fitness).

It should be noted that the use of this methodology poses a potentially difficult problem. Indeed, while in the traditional approach individuals are identified with strategies (each agent is programmed to play a specific strategy), in the indirect evolutionary approach individuals are identified with preference types. Consequently, we need to choose a *rule which maps profiles of preferences into behavior* to evaluate the evolutive fitness of a certain preference type, given the preference types of others. We can think of the evolution of preferences as a two-speed dynamic process. There is a *fast adaptation process* whereby, given the distribution of preference types in the population and the information structure, players adjust their behavior until they reach some plausible stationary state (equilibrium play). These states, in turn, constitute the rounds of a *slower evolutionary process* along which the population composition adjusts according to a fitness criterion. Thus, the static solution concept which captures players' behavior in any relevant state of the evolutionary dynamics (when the distribution of preferences is given) becomes crucial.

Existing studies in the literature on the evolution of preferences adopts Nash equilibrium (or variants of it) as a rule to describe behavior in any state of the evolutionary dynamics. A common feature of these studies is

the result that the survival of nonmaterialistic preferences strictly depends on the observability of the preference types of others. On one side, some authors conclude that evolution favors non-selfish preferences *if* players observe their opponents' preference types at least to some degree. For example, Güth and Yaari (1992) argue that observability is the driving force for the evolution of interdependent preferences (altruistic, reciprocal,...) by means of a commitment effect. Bester and Güth (1998) show that if the context exhibits strategic complementarities and players can observe their opponent's preference type, natural selection favors altruism.[1] Conversely, several other authors conclude that *if* opponents' preferences are *not* observable, evolutionary forces favor preferences which coincide with the material payoff (Ok and Vega-Redondo, 2001). Similarly, Ely and Ylankaya (2001) show that if preferences are unobservable, outcomes which are supported by stable preferences distributions must be Nash equilibrium outcomes. These negative results are emphasized by Samuelson (2001) and Robson and Samuelson (2011). In particular, Samuelson (2001), in his introduction to the symposium on the evolution of preferences, asserts: "*these papers* [i.e., Ok and Vega-Redondo, 2001; Ely and Ylankaya, 2001] *highlight the dependence of indirect evolutionary models on observable preferences, posing a challenge to the indirect evolutionary approach that can be met only by allowing the question of preference observability to be endogenously determined within the model*„ (p. 228).

Dekel, Ely, and Ylankaya (2007) study the evolution of preferences by means of a very general indirect evolutionary model and confirm the necessity result of the previous works. As suggested by Samuelson (2001), they also investigate its robustness. They assume that in each state of the long-term dynamics, agents play a Bayesian Nash equilibrium (BNE), given their preferences and the information about others' preferences. They show that, if players know the distribution of preference types in the population but do not observe their opponent's preferences, any non-Nash equilibrium outcome of the underlying game with fitness payoffs can be destabilized by a mutant with materialistic preferences. They conclude that this result holds true even if we allow for a small degree of observability.

Samuelson and Robson (2011), when commenting on the actual state of the art of the literature on the evolution of preferences, conclude that

[1] For other results on the evolutionary success of non-selfish preferences under observability of the preference types of others, see also Herold and Kuzmics (2008, 2009).

"*the indirect evolutionary approach with unobservable preferences gives us an alternative description of the evolutionary process, one that is perhaps less reminiscent of biological determinism, but leads to no new results*„ (p. 234). They propose to change the evolutionary scenario so as to encompass the evolution of signals of own preferences (and of reactions to these signals), as suggested by Samuelson (2001). One thing is certain: if we do not want to discard the indirect evolutionary approach, we need to overcome the strict dependence of the evolutionary success of nonmaterialistic preferences on observability.

In the next section, we present a new evolutionary scenario applied to the Centipede Game. We show that by adopting the solution concept of *self-confirming equilibrium* to capture the short-run play, there is room for altruistic preferences to evolve *even if* preferences are unobservable.[2] Before going into the details of our model, it is worth discussing two preliminary observations.

First, we want to stress that it is hard to justify the use of Nash equilibrium to capture the short-run play. In general, we can identify two possible justifications for a solution concept. The *eductive interpretation* (Binmore, 1987) assumes that the basic game is played only once and agents logically derive and implement the solution. To justify a BNE play eductively, we would have to assume rationality and common certainty of rationality, given common knowledge of the (Bayesian) game.[3] Yet, rather than implying in general that a BNE is played, these assumptions naturally deliver the larger set of (interim) rationalizable profiles of strategies for the given Bayesian game.[4] But even if they deliver only the unique BNE, they remain strong

[2]We pick altruistic preferences (joint material payoff maximizing) for various reasons. First, these altruistic preferences are a very simple form of other-regarding preferences and constitute a clear alternative to selfish preferences. Second, their evolution has already been studied (Bester and Güth, 1998). Third, in the context of the Centipede Game described in the next section, a population of all altruists is evolutionary stable and induces an (aggregate) outcome different from the Nash equilibrium outcome of the basic game. Since our main purpose is to show that we can find a stable distribution of preferences which induce (stable) non-Nash behavior, to consider a general model with every possible specification of preferences, as in Ely and Ylankaya (2001) and Dekel, Ely, and Ylankaya (2007), goes beyond the scope of this paper.

[3]By "knowledge" we mean a true, justified belief based on observation and deduction while, by "certainty" we mean a probability-one belief.

[4]On interim rationalizability, see Ely and Pesky (2006), Dekel, Fudenberg, and Morris (2007) and Battigalli et al. (2008).

epistemic assumptions. If we alternatively assume that the equilibrium play in each state of the evolutionary dynamics is the result of an adaptation process (*adaptive interpretation*),[5] we would need to be more explicit about the learning dynamics to show that the play surely converges to some BNE. In particular, we would have to specify what players can observe regarding the outcomes of previous interactions in order to explain how they could end up in equilibrium, holding a *common and correct belief about the play* (whatever their preferences). For example, in static games a player can learn the correct probabilities of each of his opponents' actions by observing only the actions played in every round by his opponents.[6] On the other hand, if the underlying game is dynamic, assuming that players observe ex post only the actions played by their opponents may not be enough to support the convergence to a BNE. We would need to make the implausible assumption that contingent plans (i.e., behaviour at counterfactual nodes) are fully observed ex post.[7] With these considerations in mind, we may want to adopt a solution concept *weaker* than BNE.

Second, we want to stress that in the above cited studies, observability of co-players' preferences emerges as a necessary condition to sustain the evolutive success of non-materialistic preferences *only because* it is assumed that in every state of the evolutionary dynamics a BNE is played. Indeed, if preferences are not observable, the correctness of conjectures about opponents' behavior implies that selfish players (expected material payoff maximizers) will always obtain the highest possible material payoff. Consequently, they will inevitably perform either the same or better in terms of fitness than the nonmaterial utility maximizers.

We suggest that self-confirming equilibrium (SCE) is an appropriate solution concept to describe the play in the relevant states of the evolutionary

[5] This is an implicit assumption in Güth and Yaari (1992), Bester and Güth (1998) and Dekel, Ely, and Ylankaya (2007).

[6] Notice that under the assumption of private values, learning the probability of actions would lead players to hold a common and correct belief about the play. By contrast, if we had interdependent values (but this is not the typical case in the cited literature), learning the probability of actions would not be enough. We would need to assume, for example, that players have access to some public joint statistics on actions *and* preference types of their opponents.

[7] Or you need to assume that long-lived and very patient players experiment *enough* to learn their opponents' strategies. On optimal experimentation, see, e.g., Fudenberg and Levine (1993b). However, there are no studies which offer general results on the convergence of learning processes to Nash equilibrium.

dynamics. Indeed, since self-confirming equilibrium has a natural learning foundation, it is suitable to represent the stationary states of the adaptive processes.[8] We adapt to extensive-form games the version of SCE proposed by Dekel et al. (2004), similar in spirit to the notion of *conjectural equilibrium* proposed by Battigalli (1987) and Battigalli and Guaitoli (1997). Essentially, the SCE describes situations in which players choose best replies to their conjectures about opponents' play (*rationality condition*) and the information on the actual play revealed *ex post*, after that the choices are made, does not induce them to revise those conjectures, independent of whether they are correct or not (*conjectures' confirmation property*). The feature of a SCE which matters here is that it allows situations where players hold *heterogenous beliefs about the strategies* as long as these beliefs are not contradicted by the evidence. From an adaptive perspective, we can justify such an equilibrium situation by arguing that, by behaving differently, individuals with different preferences may accumulate different experiences through the learning process (if they rely only on their own observations). In the next section, we show that by weakening, in this sense, the assumption of correctness of conjectures and allowing for individual learning, we gain further insights into the evolution of nonmaterialistic preferences.

2 The indirect evolutionary model

2.1 The scenario

Consider a population composed of individuals with *heterogeneous preference types* whose members interact with each other in pairs. Each player i can be either an altruistic or a selfish type, that is, for every i, $\Theta_i = \{\alpha, \sigma\}$, where α means altruistic and σ selfish. An altruistic type aims at maximizing the joint (material) payoff, whereas a selfish type aims at maximizing his own material payoff. We define the material consequences of players' actions with the functions $\mathbf{m}_i : Z \to \mathbb{R}$, $i \in N$, where Z is the set of terminal histories. While the utility of a selfish player i coincides with his material payoff, $U_{i,\sigma}(z) = \mathbf{m}_i(z)$, the preferences of an altruistic player i are represented by utility functions of the form: $U_{i,\alpha}(z) = \mathbf{m}_1(z) + \mathbf{m}_2(z)$.

We denote q the measure of altruistic types in the population. We name

[8] On learning foundations of SCE in extensive-form games, see Fudenberg and Levine (1993b) and Fudenberg and Kreps (1995).

q the state of the evolutionary dynamics in which the measure of altruistic types is q. We do not make any explicit assumption on the players' knowledge of q: players might not have any clue of the distribution of preference types in the population. Actually, they might not even be aware of the existence of preference types different from their own.

We assume that agents endowed with their specific preference types and beliefs are *recurrently* and *randomly* drawn to play a two-player dynamic game with material payoffs (monetary payoffs). Suppose that each agent may end up with probability $\frac{1}{2}$ in the role of player 1 and $\frac{1}{2}$ in the role of player 2. Hence, the probability that an altruistic type is drawn to play in the role of player i is exactly q, the measure of the set of altruists in the population. We assume that players do not receive *ex ante* any signal about their co-player's preference type.[9] Moreover, suppose that, after playing, agents observe *only the actions* actually performed by their co-player. In our scenario, this is equivalent to assuming that players observe ex post the terminal history.[10] We assume that in each round, given this information structure and a preference distribution q, agents play a SCE of the underlying game.

Generally speaking, a SCE describes a situation where (i) players best respond to their (heterogeneous) conjectures about their opponent's behavior by maximizing their *perceived* expected utility, and (ii) the information revealed ex post confirms their conjectures. To our knowledge, there is no study that sytematically analyzes different definitions of this notion of equilibrium according to the assumed scenario. Consequently, there is no off-the-shelf definition which is appropriate for our context (i.e., a dynamic game with incomplete information played recurrently with random matching). Hence, it is worth defining explicitly our version of SCE. Let S_i denote the set of strategies for player/role i. A profile of strategies $\beta_i = (s_{\theta_i})_{\theta_i \in \Theta_i} \in S_i^{\Theta_i}$ is a

[9] We make this assumption because this is the condition under which the negative result of the previous papers obtains, given the assumption of a BNE play.

[10] Note that in our scenario, we do not assume that there is an *ex ante* stage where players are ignorant about their preferences and form contingent plans of actions which depend on the realization of their types (due to a chance move). Indeed, in our context, it is more natural to assume that characteristics (i.e., types) are fixed for every agent. What is random is the matching of individuals endowed with their own type. This matching is, in turn, determined by the objective probability q, which measures the altruistic types in the population. Hence, by observing the terminal history, players obtain (possibly partial) information only about their opponent's moves and not about their opponent's preference type.

behavioral rule for player i, that is, a specification of a pure strategy for every preference type θ_i. In each round, for any given profile of behavioral rules $\beta = (\beta_1, \beta_2)$, the measure of altruistic types q induces a probability distribution over terminal nodes and thus a probability distribution over messages. Call Z the set of terminal histories. We denote $\pi(z|s_i; q, \beta_j)$ the *objective* probability that an agent playing in role i observes terminal history $z \in Z$ when he plays s_i, the measure of altruistic types is q and the behavioral rule for the other player/role is β_j. We denote $p(z|s_i; \mu_i)$ the *subjective* probability of observing z as assessed by an agent playing in role i with beliefs $\mu_i \in \Delta(S_j)$. Define $\zeta : S_1 \times S_2 \to Z$ and call $\zeta(s)$ the terminal history induced by the strategy profile s.

Definition 1 *A profile of behavioral rules* $\beta = ((s_{\theta_i})_{\theta_i \in \Theta_i})_{i=1,2}$ *is a* self-confirming equilibrium *if for each preference type* $\theta_i \in \Theta_i$ *of each player* $i \in \{1, 2\}$ *we can find a belief* $\mu_{i,\theta_i} \in \Delta(S_j)$ *such that*

i) $s_{\theta_i} \in \arg\max_{s_i \in S_i} \sum_{s_j \in S_j} \mu_{i,\theta_i}(s_j) U_{i,\theta_i}(\zeta(s_i, s_j))$ *and*

ii) $\forall z \in Z$, $p(z|s_{\theta_i}; \mu_{i,\theta_i}) = \pi(z|s_{\theta_i}; q, \beta_j)$.

The first condition states that each preference type chooses a pure strategy which maximizes his expected utility, given his beliefs (*rationality condition*).[11] The second condition states that for each preference type the statistical distribution of observations over terminal histories generated by the population composition and the moves of players must coincide with his subjective probability distribution. That is, the empirical frequencies of actions as observed by any preference type must not induce that preference type to revise his beliefs (*confirmation condition*).

Each SCE play determines the *objective fitness* of each preference type involved in the game by means of the material payoff obtained on average by agents endowed with that preference type. We evaluate the evolution of altruistic preferences by considering a *replicator dynamics*. Denote $\overline{m}_\theta(q)$ the average payoff of preference types θ in a typical state of the evolutionary dynamics where the fraction of altruists in the population is q. Call $\overline{m}(q)$ the

[11] In a dynamic game, one may strengthen the rationality condition by requiring that agents' maximizing also be conditional on unexpected information sets. But this does not change the set of SCE outcome distributions. This modification may be relevant for the analysis of *rationalizable SCE's*, but it does not affect our example.

current average payoff in the population. The growth rate for the population fraction q with altruistic preferences equals the difference between the current average payoff of individuals with these preferences ($\overline{m}_\alpha(q)$) and the current average payoff in the population ($\overline{m}$), that is: $\frac{dq}{dt} = [\overline{m}_\alpha(q) - \overline{m}(q)]\, q$.

With this scenario in mind, we next discuss the example of the (three-stages) Centipede Game. We consider an evolutionary dynamics where in each state, given a preference distribution q, players play a SCE of the Centipede Game, given the ex post information structure specified above and their beliefs about their opponent's play in that state. We will show that starting within a large subset of the simplex of all possible beliefs such an evolutionary dynamics favors altruistic preferences.

2.2 The evolutionary Centipede Game

Consider this three-period version of the Centipede Game:[12]

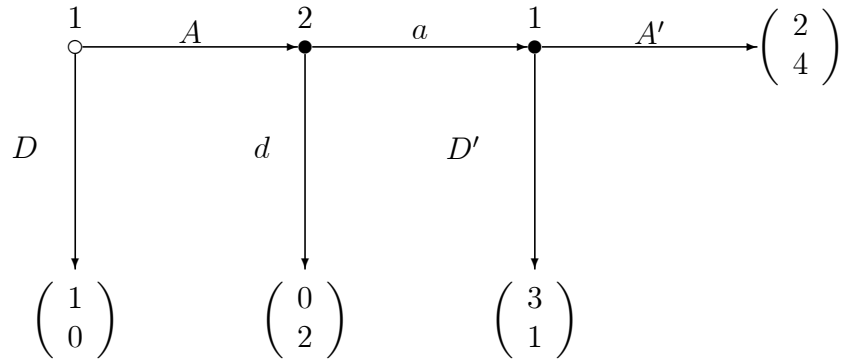

Fig. 1. Centipede Game with monetary payoffs

Suppose that selfish agents drawn to play in the role of player 2 attach a small probability to action A' being played conditional on (A, a), that is, $\mu_{2,\sigma}(AA'|A) < 1/3$ holds. Suppose that selfish agents drawn to play in the role of player 1 attach a small probability to action a being played by

[12] Notice that the numbers attached to the terminal nodes represent players' material payoffs (money) and do not necessarily coincide with their utilities.

player 2, that is, $\mu_{1,\sigma}(a) < 1/3$ holds. Hence, the unique restriction we impose to the initial system of beliefs is that all selfish individuals believe that strategies AA' and a are unlikely to be played.[13] According to these beliefs, selfish agents who have to play in player 2's role will choose d, while selfish agents who have to play in player 1's position will choose D. We do not have to impose any restriction to the system of beliefs of altruistic types. Indeed, whatever their beliefs, given that they want to maximize the expected joint payoffs, they will play "across" whenever it is their turn to make a move (i.e., AA' when they take player 1's role and a-whatever the strategy of player 1-when they take player 2's role). In every stage of the evolutionary dynamics, this profile of behavioral strategies constitutes a SCE of the basic game, given the system of beliefs specified above and the assumption that players can observe *ex post* only the actions played by their co-players. In fact, from a *static perspective* each preference type of each player is maximizing his (perceived) expected utility, and given the ex post information structure, none of them can revise their (possibly wrong) beliefs about the equilibrium strategies of their opponent. In particular, in any state q the correct probability of strategy a is exactly q. However, selfish types playing in the role of player 1 believe that the probability of a is very small, that is, $\mu_{1,\sigma}(a) < \frac{1}{3}$, which is wrong in those states where $q \geq \frac{1}{3}$. By playing D, they prevent themselves from realizing that a is actually more likely than expected. Similarly, the true probability that strategy AA' is chosen by player 1 coincides with the probability that the type drawn to play in player 1's role is altruistic, that is, q. After A, A' should be expected with probability one. However, selfish types in player 2's role believe that A' is unlikely and that by playing d, they cannot revise their wrong beliefs. From an *adaptive perspective*, by sticking to playing "down" whenever they can, selfish players do not have the opportunity to *learn* the correct probabilities of their co-player's strategies. By contrast, altruistic types may end up in equilibrium having a complete picture of the behavior of the opponent, but the correctness of their beliefs is not an issue here. Given their preferences,

[13] Obviously, without such a restriction we would face a multiplicity of SCE's, given the ex post information structure. We motivate such beliefs by reasoning that agents do not have access to any public statistics before playing. We could imagine that their initial beliefs are based on introspection: since they do not have any clue of the average behavior, they make the simple assumption that the opponent is likely to behave as they would behave it they were in his shoes. Moreover, note that these beliefs are non-doctrinaire, meaning that they attach positive probability to every possible terminal history.

they choose to play "across" whenever they have the possibility to do so, whatever their beliefs. Regardless of whether they learn the exact frequencies of each action or not, they would behave altruistically anyway.

Notice that this SCE is also *rationalizable*, meaning that it is consistent with rationality (i.e., consistent with conditionally expected utility maximization), confirmation of conjectures and initial common beliefs in rationality and confirmation of conjectures.[14] It is immediately clear that when two altruists, two selfish types, or a selfish type in the role of 1 and an altruist in the role of 2 are matched, their SCE strategies and conjectures are consistent with all these assumptions. Consider the match between an altruistic player 1 and a selfish player 2. Given this match, the SCE path will be (A, d). After A, the selfish player 2, who (initially) believes in the rationality of player 1, is allowed to hold arbitrary (conditional) beliefs about the subsequent move of player 1. That is to say, the conditional belief that D' is very likely (i.e., $\mu_{2,\sigma}(AA'|A) < 1/3$) is consistent with the (initial) belief of player 2 in the rationality of player 1. Given this conditional belief and the rationality of player 2, d follows A. By continuing to play in the same way and observing ex post *only the actions* played, selfish types learn the correct frequencies of actions A and D. However, they are unable to infer that A' follows A with probability one. But if we were to assume that players can observe ex post *also the preference types* of their co-player, the SCE profile of strategies described above would not be rationalizable. Indeed, if preference types are observed ex post, the selfish types could infer the correct frequencies of strategies from the rationality assumption and the observed frequency of preference types. This is why we adopted the general version of SCE proposed by Dekel et al. (2004) applied to extensive-form games and not the original notion of SCE introduced by Fudenberg and Levine (1993a, 1993b). In fact, in the latter version, the ex post information structure is such that players can observe *both* the moves of the opponents *and* nature. But if this is the case, then ex post observability of preference types is implied.

What is crucial in our example is that selfish individuals *do not learn* the exact frequencies of their co-players' strategies but continue to expect selfish behavior. Because of their incorrect, but confirmed beliefs, they stick

[14] Rubinstein and Wolinsky (1994) introduce the concept of *rationalizable conjectural equilibrium* for static games; Dekel et al. (1999) produce a version of this concept for extensive-form games (*rationalizable self-confirming equilibrium*). See also Esponda (2010), who provides an extension of Rubinstein and Wolinsky's (1994) notion of rationalizable conjectural equilibrium to games with incomplete information.

to their selfish behavior so that, depending on the value of q, they might perform worse than altruistic individuals.

Let us analyze how the share of altruists in the population evolves. Recall that the equation of the replicator dynamics is the following: $\frac{dq}{dt} = [\overline{m}_\alpha(q) - \overline{m}(q)]\, q$. We need to compute $\overline{m}_\alpha(q)$, the average SCE payoff of an altruistic type, averaged across the roles he can take and across the preference types he can face. When an altruist faces another altruist (and this occurs with probability q), with probability $\frac{1}{2}$ he takes the role of player 1 and obtains 2, whereas with probability $\frac{1}{2}$ he takes the role of player 2 and obtains 4. When an altruist faces a selfish agent (which occurs with probability $1-q$), he obtains 0 whatever his role. Hence, the average payoff in the sub-population of altruists is: $\overline{m}_\alpha(q) = (\frac{1}{2}2 + \frac{1}{2}4)q = 3q$. A selfish agent faces an altruist with probability q and obtains 1 when he takes player 1's role, whereas he obtains 2 when he takes player 2's role. With probability $1-q$ he faces another selfish agent and obtains 1 when he takes player 1's role and 0 when he takes player 2's role. Thus, the average payoff of a selfish type is: $\overline{m}_\sigma(q) = (\frac{1}{2}1 + \frac{1}{2}2)q + (\frac{1}{2}1 + \frac{1}{2}0)(1-q)$. Consequently, we can compute the average payoff in the population, that is: $\overline{m}(q) = 2q^2 + \frac{1}{2}q + \frac{1}{2}$. The dynamics of the share of altruists in the population is represented by the following equation: $\frac{dq}{dt} = [\overline{m}_\alpha(q) - \overline{m}(q)]\, q = \left[\frac{5}{2}q^2 - 2q^3 - \frac{1}{2}q\right]$. The picture below represents the phase diagram for this dynamics.

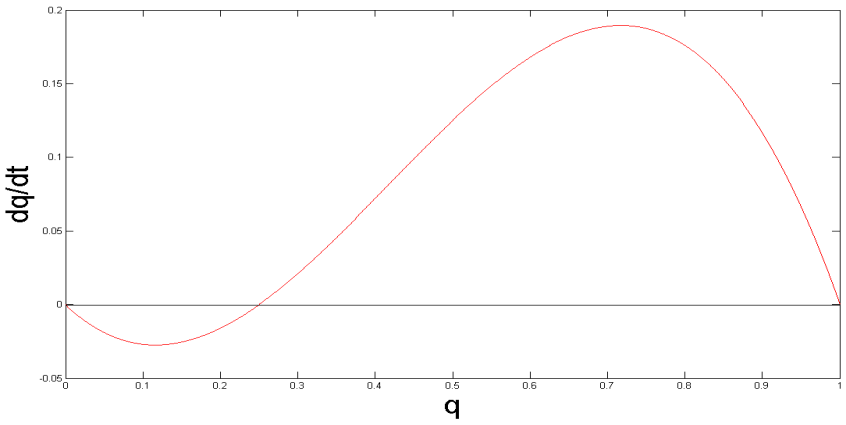

Fig. 2. Phase diagram for the dynamics of the population share of altruists.

Observe that if we have a heterogeneous population where the measure of altruists is high enough, i.e., $q > \frac{1}{4}$ individuals with altruistic preferences perform better than individuals with selfish preferences, and they are able to invade the whole population. Indeed, altruistic types fare well when they meet other altruists because they reach the efficient outcome and obtain a high payoff. If this match occurs sufficiently often, they reproduce faster than selfish types (who never reach the efficient outcome) and end up being the only surviving preference type. More interestingly, Figure 2 and the reasoning above tell us that *a population of all altruists is not vulnerable with respect to the injections of a small share of selfish individuals.*

It is worth noting that we did not make any explicit assumption on *experimentation*. We implicitly assumed that agents in every round maximize their current (expected) utility, given their actual beliefs, and behave *myopically*, which implies that they do not implement any dynamically optimal strategy to learn opponents' behavior. Even if we allow for *active* learning, it is not granted that agents learn the correct probabilities of opponents' strategies so that the steady state is a BNE. Suppose we have a population of all altruists and we inject few selfish types. It might be objected that if selfish agents in player 2's role often experiment with action a, they might realize that player 2's strategy actually is to play A' after a. A similar reasoning would apply to selfish agents in player 1's role: if they often experiment with action A, they will eventually realize that action a is more likely than expected. However, it is not clear how much patient they should be (i.e., how high their discount factor should be) in order to experiment frequently enough with a (currently) suboptimal action and learn the correct probabilities of their opponents' strategies.[15] Moreover, the system of beliefs of selfish agents, which supports the SCE described above, actually attaches positive probability to actions A' and a. Neverthless, given the heterogeneity of the population, whenever one selfish agent is matched with another, he observes exactly what he expects. Hence, even with little experimentation, we could justify the SCE described above as a steady state. If, for example, a selfish agent in the role of player 1 observes a when experimenting with A, he could reason that the other player also experimented or simply made a mistake.

Notice that the SCE we considered is played *throughout* the evolutionary

[15] For optimal experimentation in extensive-form games, see Fudenberg and Levine (1993b). They show that the steady states of an extensive-form game approximate Nash equilibria *if* lifetimes go to infinity more quickly than the discount factor tends to one.

dynamics. Indeed, if players stick to their equilibrium strategies, the initial system of beliefs of selfish agents is never contradicted by *ex post* observations. However, notice also that, if, in a population of all altruists, we inject selfish agents endowed with another system of beliefs, it might happen that a profile of behavioral strategies which is a SCE in one state of the evolutionary dynamics is not a SCE in another state. For example, if we inject selfish types with beliefs $\mu_{1,\sigma}(a) \geq 1/3$ and $\mu_{2,\sigma}(AA'|A) \geq 1/3$, which support a SCE where selfish players play respectively AD' and a, these types perform better than altruists and their measure increases until the state $q = 1/3 - \varepsilon$ is reached. In this state, a switch to another SCE needs to occur since the initial system of beliefs of selfish agents could not be confirmed. However, in order to characterize *every* possible evolutionary dynamics, given the initial system of beliefs, we would need to introduce a specific criterion to select a SCE in every state (and to switch from one SCE to another). In particular, we would need to be more explicit on the way agents *form* their (initial) beliefs and on how these beliefs *evolve* together with q. Such a general model would be a worthwhile topic for future research.

3 Conclusions

The literature on the evolution of preferences has produced a common result: conditional on playing some (Bayes) Nash equilibrium in every state of the evolutionary dynamics (when the preferences are fixed), only selfish individuals survive *if* the preference types are not observable. We have shown that by adopting the weaker solution concept of SCE, which allows for heterogenous beliefs across preference types, there is room for altruistic preferences to evolve *even if* preferences are not observable. Hence, adopting self-confirming equilibrium as a rule to pin down behavior in the stages of the long-term evolutionary dynamics is promising, and, most remarkably, it has a natural learning foundation.

It is worth pointing out that we selected a *particular* self-confirming equilibrium and that, of course, there may be many other SCE's for the same ex post information structure. However, we allowed for a very large set of initial beliefs imposing some restrictions only to the system of beliefs of selfish types. The restrictions we imposed are quite plausible nonetheless: selfish players believe that the opponent will behave in the same way as they would if they were in his shoes. The implicit assumption is that they use introspection to

form their beliefs.

Moreover, we assumed that in every state of the evolutionary dynamics, players would play a self-confirming equilibrium *as if* they learned to play it. That is, we did not explicitly model the short-term learning dynamics that would lead to players playing that specific self-confirming equilibrium. By virtue of the fact that a self-confirming equilibrium can be typically learned in an adaptive way, it would be instructive to analyze and describe the short-term behavioral adaptation. Such an analysis would more extensively support the choice of self-confirming equilibrium as a rule for mapping preferences into behavior in the relevant state of the long-term evolutive process.

References

[1] Battigalli, P., 1987. Comportamento Razionale ed Equilibrio nei Giochi e nelle Situazioni Sociali. Unpublished dissertation, Bocconi University, Milan.

[2] Battigalli, P., Guaitoli, D., 1997. Conjectural Equilibria and Rationalizability in a Game with Incomplete Information, in: Battigalli, P., Montesano, A., Panunzi, F. (Eds), Decisions, Games and Markets. Kluwer Academic Publishers, Dordrecht, pp. 97-124.

[3] Battigalli, P., Di Tillio, A., Grillo, E., 2008. Interactive Epistemology and Solution Concepts for Games with Asymmetric Information. Working Paper 340, IGIER, Bocconi University, Milan.

[4] Bester, H., Güth, W., 1998. Is Altruism Evolutionarily Stable? Journal of Economic Behavior and Organization, 34, 193-209.

[5] Binmore, K., 1987. Modeling Rational Players: Part 1. Economics and Philosophy, 3, 179-214.

[6] Dekel, E., Fudenberg, D., Levine, D.K., 1999. Payoff Information and Self-Confirming Equilibrium. Journal of Economic Theory, 89, 165-185.

[7] Dekel, E., Fudenberg, D., Levine, D.K., 2004. Learning to Play Bayesian Games. Games and Economic Behavior, 46, 282-303.

[8] Dekel, E., Fudenberg, D., Morris, S., 2007. Interim Correlated Rationalizability. Theoretical Economics, 2, 15–40.

[9] Dekel, E., Ely, J.C., Ylankaya, O., 2007. Evolution of Preferences. Review of Economic Studies, 2007, 74(3), 685-704.

[10] Ely, J.C., Pesky, M., 2006. Hierarchies of Belief and Interim Rationalizability. Theoretical Economics, 1, 19-65.

[11] Esponda, I., 2010. Rationalizable Conjectural Equilibrium. Mimeo, New York University, Stern School of Business.

[12] Fudenberg, D., Kreps, D.M., 1994. Learning in Extensive-Form Games, II: Experimentation and Nash Equilibrium. Mimeo, Stanford University.

[13] Fudenberg, D., Kreps, D.M., 1995. Learning in Extensive-Form Games, I: Self-Confirming Equilibria. Games and Economic Behavior, 8, 20-55.

[14] Fudenberg, D., Levine, D.K., 1993a. Self-Confirming Equilibrium. Econometrica, 61, 523-545.

[15] Fudenberg, D., Levine, D.K., 1993b. Steady State Learning and Nash Equilibrium. Econometrica, 61, 547-573.

[16] Güth, W., Kliemt, H., 1998. The Indirect Evolutionary Approach: Bridging the Gap between Rationality and Adaptation. Rationality and Society, 10(3), 377-399.

[17] Güth, W., Yaari, M., 1992. Explaining Reciprocal Behavior in a Simple Strategic Game, in: Witt, U. (Ed.), Explaining Process and Change-Approaches to Evolutionary Economics. University of Michigan Press, pp. 23-24.

[18] Güth, W., 1995. An Evolutionary Approach to Explaining Cooperative Behavior by Reciprocal Incentives. International Journal of Game Theory, 24, 323-344.

[19] Ely, J.C., Ylankaya, O., 2001. Nash Equilibrium and the Evolution of Preferences. Journal of Economic Theory, 97(2), 255-272.

[20] Herold, F., Kuzmics, C., 2008. Evolution of Preferences Under Perfect Observability: Almost Anything is Stable. Mimeo, Northwestern University, Kellogg School of Management.

[21] Herold, F., Kuzmics, C. 2009. Evolutionary Stability of Discrimination Under Observability. Games and Economic Behavior, 67(2), 542-551.

[22] Ok, E., Vega Redondo, F., 2001. On the Evolution of Individualistic Preferences: An Incomplete Information Scenario. Journal of Economic Theory, 97, 231-254.

[23] Rubinstein, A., Wolinsky, A., 1994. Rationalizable Conjectural Equilibrium: Between Nash and Rationalizability. Games and Economic Behavior, 6, 299-311.

[24] Samuelson, L., 2001. Introduction to the Evolution of Preferences. Journal of Economic Theory, 97, 225-230.