THE RESEARCH INSTITUTE OF INDUSTRIAL ECONOMICS

WORKING PAPER No. 489, 1997

EVOLUTION, RATIONALITY AND EQUILIBRIUM IN GAMES

BY JÖRGEN W. WEIBULL

Evolution, Rationality and Equilibrium in Games

Jörgen W. Weibull*
Stockholm School of Economics and I.U.I.

September 30, 1997

Abstract

Evolutionary game theory studies the robustness of strategy profiles and sets of strategy profiles with respect to evolutionary forces in games played repeatedly in large populations of boundedly rational agents. The approach is macro oriented in the sense of focusing on the strategy distribution in the interacting population(s). Some main features of this approach are here outlined, and connections with learning models and standard notions of game-theoretic rationality and equilibrium are discussed. Some desiderata and results for robust long-run predictions are considered. Doc: *eea.tex.*

1 Introduction

The usual rationalistic interpretation of non-cooperative game theory assumes that the game is played exactly once by perfectly rational players. The game, along with the players' rationality and the predicted equilibrium, is assumed to be common or mutual knowledge, see e.g. Tan and Werlang (1988) and Aumann and Brandenburger (1995). By contrast, the evolutionary interpretation assumes that the game is played many times by boundedly rational players who are randomly drawn from large populations and who have little or no information about the game.

The hallmark of an evolutionary model is that it combines two processes: one *selection process* that favor some varieties over others, and one process

*This essay was presented in the session "Learning and evolution, where do we stand?" at the European Economic Association Annual Congress in Toulouse, September 1, 1997. I thank Maria Saez Marti for helpful comments.

that creates this variety, to be called the *mutation process*. In the applications to be discussed here, the varieties in question are strategies in a game.[1] The emphasis will be on the selection aspect, while mutations will be accounted for only indirectly by way of stability analysis.

A class of selection dynamics is introduced in section 2. In section 3 these dynamics are related to certain learning models. Connections with standard notions of rationality and equilibrium in games are discussed in section 4, and section 5 elaborates on desiderata and results for structurally robust predictions in games. Section 6 concludes.

2 Selection Dynamics

Consider a finite n-player game. Imagine that there are n large populations, one for each of the n player positions in the game. Every now and then n-tuples of individuals, one from each population, are randomly drawn to play the game. All individuals play pure strategies. A *population state* $x = (x_1, ..., x_n)$ consists of n population distributions, one for each player population, where each component x_{ih} of the i'th distribution x_i is the fraction of individuals in that population who play pure strategy h. Hence, a population state x is formally identical with a mixed-strategy profile in the game, and the population distribution x_i is formally identical with a mixed strategy for player position i. The set of mixed-strategy profiles in the game is the Cartesian product of the mixed-strategy simplexes associated with the player positions in the game - a polyhedron.

We describe the evolution of the population distribution x by way of a system of (autonomous) ordinary differential equations of the form

$$\dot{x}_{ih} = g_{ih}(x)x_{ih} \ , \tag{1}$$

where $g_{ih}(x)$ is the *growth rate* of pure strategy h in population i.[2] Such dynamics will be referred to as *selection* dynamics since pure strategies that are not initially used will remain unused forever ($x_{ih}(0) = 0 \ \Rightarrow \ x_{ih}(t) = 0 \ \forall t > 0$).

[1]The objects that are selected and mutated are whole strategies, not just the local strategies at information sets (of an underlying extensive-form game) that are reached in the play of the game. For an analysis of evolutionary processes of the latter type, see Nöldeke and Samuelson (1993).

[2]Such a system of ordinary differential equations have well-defined solutions if each growth-rate function g_i is Lipschitz continuous and keeps the sum $\sum_h x_{ih} g_{ih}(x)$ constantly equal to zero, so that population shares always sum to one.

The focus will here be on three classes of such selection dynamics. The first class are the *payoff-positive* dynamics, where all pure strategies that earn above average have positive growth rates and all pure strategies that earn below average have negative growth rates. The second class are the *convex-monotone* dynamics, where a pure *or mixed* strategy has a higher growth rate than a pure strategy if the first earns a higher payoff than the second. The third class are the *weakly payoff-positive* dynamics, where at least some pure strategy that earns above average has a positive growth rate when such a strategy exists. The convex monotone dynamics and the payoff positive dynamics are both sub-classes of the weakly payoff-positive dynamics.

The most widely used selection dynamics in evolutionary game theory is the replicator dynamics, first developed by Taylor and Jonker (1978) for the case of a single population playing a symmetric two-player game. Extensions to the case of n populations playing an n-player game were developed by Taylor (1979) and Maynard Smith (1982). The following version is due to Taylor (1979):[3]

$$\dot{x}_{ih} = \left[u_i(e_i^h, x_{-i}) - u_i(x)\right] x_{ih} \tag{2}$$

Here $u_i(x)$ denotes the (expected) payoff to player position i when strategy profile x is played, and $u_i(e_i^h, x_{-i})$ is the (expected) payoff to pure strategy h when played in position i of the game against strategy profile x. In population terms, $u_i(e_i^h, x_{-i})$ is the payoff to individuals in population i who play pure strategy h, and $u_i(x)$ is the average payoff in that population. Hence the population growth rate of each pure strategy is here proportional to its payoff advantage - the difference between its payoff and the average payoff in its player population. Clearly the replicator dynamics belongs to all three classes of selection dynamics described above.

3 Learning Models and Selection Dynamics

One can divide learning models into three broad categories, belief-based learning, reinforcement (or stimulus-response) learning, and learning by imitation. It has recently been shown that the replicator dynamics and certain other weakly payoff-positive selection dynamics can be viewed as approximations for learning models in the last two categories.

[3]In Maynard Smith (1982) the right-hand side is divided by the average payoff, $u_i(x)$, which is then presumed to be positive.

3.1 Individual learning by reinforcement

A model used in the psychology literature on learning is the so-called reinforcement model due to Bush and Mosteller (1951). Börgers and Sarin (1997) analyze Cross' (1973) version of Bush's and Mosteller's (1951) model. Börgers and Sarin consider two players who repeatedly play a finite two-player game. The players use mixed strategies and after every round each player adjusts her choice probabilities according to her realized payoff. Börgers and Sarin show that if the number n of rounds played before any given finite time t is increased, and the reinforcement feed-back from payoffs to choice probabilities is made accordingly smoother, then in the limit, as $n \to \infty$, their stochastic process places unit probability on the state that the replicator dynamics would have reached at time t.[4]

3.2 Social learning by imitation

Gale, Binmore and Samuelson (1995) provide a simple model of social learning in large populations of individuals playing pure strategies. Each individual maintains an aspiration level concerning the payoff to be earned in the game. At discrete times a small population share of individuals, drawn at random, compare their current payoffs with their aspiration levels. If an individual's realized payoff falls below her aspiration level then she imitates a randomly drawn individual in her player population. It follows that if the population distribution of aspiration levels is rectangular over some interval containing all possible payoff values, then the probability for imitation is linearly decreasing in the expected payoff to the individual's current strategy. Based on this observation, the authors show that this process is approximated by the replicator dynamics over bounded time intervals as the number of rounds per time unit is increased and the population share of reviewing individuals is accordingly decreased.

Alternatively, one may view individual strategy adaptation as a stochastic process in continuous time. Suppose that every now and then each individual in every player population gets an impulse to revise her strategy choice. If these impulses arrive according to statistically independent Poisson processes, then the aggregate process is also a Poisson process. For large populations one may approximate the aggregate process by deterministic flows. Björnerstedt and Weibull (1996) study a number of such models, where revising individuals imitate other individuals in their own player popu-

[4] This does not imply, however, that the asymptotic behaviors of the stochastic and deterministic processes coincide. Indeed, Börgers and Sarin show that, unlike the replicator dynamics, their stochastic process is eventually absorbed in a pure-strategy profile.

lation, and show that a number of weakly payoff-positive selection dynamics, including the replicator dynamics, may be so derived.

Schlag (1997) analyses the question what imitation rules an individual *should* choose, when she now and then has the opportunity to imitate another individual in the same player position but is otherwise constrained by severe restrictions on information and memory. He finds that if the individual wants a learning rule that is *payoff increasing* in all stationary environments, then the individual should always imitate (not experiment) when changing strategy. Moreover, she should only imitate individuals whose payoff realizations are better than her own, and do this with a probability that is proportional to this payoff difference. Schlag shows that for large populations the induced stochastic process can be approximated by (a discrete-time version of) the replicator dynamics.

4 Rationality and Equilibrium

4.1 Rationality

A basic rationality postulate in non-cooperative game theory is that players never use pure strategies that are strictly dominated. This postulate requires no knowledge of other players' preferences or behavior. A more stringent rationality postulate is that players never use pure strategies that are iteratively strictly dominated. In addition, this postulate requires that all players know each others payoffs, that they know that they know each others payoffs, etc. up to a finite level of mutual knowledge (see e.g. Tan and Werlang (1988)).

Hofbauer and Weibull (1996) show that if initially all pure strategies in the game are present in the player populations, then all iteratively strictly dominated pure strategies vanish over time in any (finite) n-player game and in any *convex-monotone* dynamics. Moreover, we show that convex monotonicity is essentially necessary for the elimination of iteratively strictly dominated pure strategies in all such games. The first of these two results generalizes results in Akin (1980) and Samuelson and Zhang (1992). Cabrales (1996) develops a stochastic version of the replicator dynamics and provides conditions under which iteratively strictly dominated strategies still get wiped out in the long run.

In sum: in a certain class of selection dynamics individuals will in the long run behave as if they were rational and as if this rationality were mutual knowledge. Strictly dominated strategies may survive forever in other types of selection dynamics.

4.2 Nash equilibrium

It is easily verified that every Nash equilibrium, viewed as a population state, constitutes a stationary state in the replicator dynamics. Moreover, not all stationary states are Nash equilibria. However, it turns out that all stationary states which are not Nash equilibria are dynamically unstable. The stability criterion used here is that of *Lyapunov stability*, essentially requiring that no small perturbation of the population state can lead it away.[5] This result is due to Bomze (1986), and can be shown to hold for any weakly payoff-positive dynamics in any (finite) n-player game: Every Lyapunov stable population state in any such dynamics constitutes a Nash equilibrium (Weibull (1995)). Hence, if we require Lyapunov stability, then we in fact ask for a *refinement* of the Nash equilibrium concept. In comparison with standard non-cooperative refinements, such as "trembling hand" perfection, it turns out that Lyapunov stability in the replicator dynamics has more cutting power against mixed-strategy Nash equilibria, less cutting power against weakly dominated strategies, and more cutting power in cheap-talk games.

In sum: in any dynamically stable population state, in a fairly wide class of selection dynamics, individuals behave as if they expected this population state and played optimally given this expectation.

5 Robust Long-Run Predictions

Lyapunov stability is relevant for predictions in the medium-run but may be insufficient for robust long-run predictions: if a predicted population state x is slightly perturbed, then the dynamics should not only not lead the state away, but it should carry it back toward the predicted state x.[6] Otherwise a sequence of small shocks can add up over time and eventually initiate a motion far away from the predicted state x. Second, a robust prediction should not be too sensitive to the dynamics - "nearby" dynamics (vector fields) should have nearby predictions - a robustness property that Lyapunov stability does not have. Finally, it seems desirable that predictions be valid for a whole range of potentially relevant dynamics, including the (Taylor and Maynard Smith n-population versions of the) replicator dynamics. After all, we do not know what is the "right" dynamics, but we do know that the replicator dynamics is an approximation of a variety of empirically relevant learning processes.

[5] Formally: A state x is *Lyapunov stable* if every neighborhood B of x contains a neighborhood A of x such that if the initial state is in A, then all future states are in B.

[6] It can of course not reach x in finite time, though.

The first two desiderata are met if we require *asymptotic* rather than Lyapunov stability, i.e., the further stability property that the dynamics should bring the population state back after any sufficiently small perturbation.[7] In view of the third desideratum we are thus lead to a search for mixed-strategy profiles which, as population states, are asymptotically stable in some relevant class of selection dynamics containing the replicator dynamics. However, it turns out that many games of economic interest have no strategy profile with this property: a population state is asymptotically stable in the Taylor n-population replicator dynamics if and only if it constitutes a strict Nash equilibrium. In order to obtain robustness we are hence forced to sacrifice precision and look for asymptotically stable *sets* of mixed-strategy profiles.[8]

One mathematically simple class of sets of mixed-strategy profiles are the *faces* of the polyhedron of mixed-strategy profiles. A face X is a product set of mixed-strategy profiles where each component set X_i consists of all mixed strategies with support in some *subset* of the pure strategies available to player position i. Recall that the pure-strategy *best-reply* correspondence assigns to each strategy profile x and to each player position i its set $\beta_i(x)$ of best pure-strategy replies to x. Likewise, the pure-strategy *better-reply* correspondence assigns to each strategy profile x and to each player position i its set $\gamma_i(x)$ of weakly better pure replies, i.e., all pure strategies that do not earn less than average in their player population against x. In particular, if the strategy profile x is a *strict* Nash equilibrium then it is a pure-strategy profile, and, as a singleton face, contains its best and better replies. More generally: a face X is said to be *closed under best (better) replies* if it contains all its pure best (better) replies (Basu and Weibull (1991), Ritzberger and Weibull (1995)).

The relevance of closure under better replies for evolutionary dynamics is that if a face X has this property then it is asymptotically stable in *all* payoff positive selection dynamics, and conversely, if a face X is asymptotically stable in *some* payoff-positive dynamics, then it is closed under better replies. Faces that are closed under weakly better replies are thus robust long-run predictors for all payoff positive selection dynamics. Indeed such faces are asymptotically stable in any selection dynamics that assigns negative growth rates to pure strategies that earn below average in their player position - a property held by all convex monotone selection dynamics. Moreover, any face that is closed under best replies (and all faces that are closed under better replies are such) contains a strategically stable set in the sense

[7]A state x (or closed set X of states) is *asymptotically stable* if it is Lyapunov stable and is contained in an open set C such that the solution through any initial state in C converges to x (or to X).

[8]The following discussion is based to a large extent on Ritzberger and Weibull (1995).

of Kohlberg and Mertens (1986).[9] Thus there is a set-inclusive connection between asymptotically stable faces and powerful refinements of the Nash equilibrium concept.

It may here be noted that there is also a connection with the stochastic evolutionary models discussed in the companion paper by Young (1997). With evolution in discrete time, a finite population and best-reply selection against samples from finite records of past play, Young establishes convergence with probability one to some minimal face closed under *best* replies (related results are obtained in Hurkens (1995) and Sanchirico (1996)).

The following equations define a one-dimensional parametric family of convex-monotone selection dynamics that range from the replicator dynamics to the best-reply dynamics:

$$\dot{x}_{ih} = \frac{x_{ih} \exp\left[\sigma u_i(e_i^h, x_{-i})\right]}{\sum_k x_{ik} \exp\left[\sigma u_i(e_i^k, x_{-i})\right]} - x_{ih} \tag{3}$$

Here $\sigma > 0$ is a parameter that determines how strongly the best replies are selected for. As $\sigma \to \infty$, interior solution orbits converge to those of the best-reply dynamics, and as $\sigma \to 0$ the solution orbits converge to those of the replicator dynamics (see Hofbauer and Weibull (1996)). Since every face X that is closed under better replies is asymptotically stable in any convex-monotone dynamics, such faces are robust predictors also for dynamics of the form (3). Moreover, these dynamics are formally similar to the smooth best-reply dynamics in equation (?) in Fudenberg and Levine (1997). Potential connections between the long-run properties of these two classes of dynamics remain to be analyzed.

6 Concluding comments

Theoretical research on processes of evolution and learning in games may now have reached a point where an integration of diverse approaches is within reach. Moreover, researchers may soon be able to confront theoretical dynamic models with evidence from the laboratories of experimental game theorists. In some years' time this may provide game theory with an empirical basis on which to build dynamic models of evolution and learning in games.

[9] A *strategically stable* set in the sense of Kohlberg and Mertens (1986) is a minimal set of Nash equilibria with the property that if players "tremble" slightly, the so perturbed game has at least one nearby Nash equilibrium.

References

[1] Akin E. (1980): "Domination or equilibrium", *Mathematical Biosciences* 50, 239-250.

[2] Aumann R. and A. Brandenburger (1995): "Epistemic conditions for Nash equilibrium", *Econometrica* 63, 1161-1180.

[3] Basu K. and J. Weibull (1991): "Strategy subsets closed under rational behavior", *Economics Letters* 36,141-146.

[4] Björnerstedt J. and J. Weibull (1996): "Nash equilibrium and evolution by imitation", in K. Arrow et al (eds.), *The Rational Foundations of Economic Behaviour*, Macmillan.

[5] Bomze I. (1986): "Non-cooperative two-person games in biology: a classification", *International Journal of Game Theory* 15, 31-57.

[6] Börgers T. and R. Sarin (1997): "Learning through reinforcement and replicator dynamics", forthcoming, *Journal of Economic Theory.*

[7] Bush R. and F. Mosteller (1951): "A mathematical model for simple learning", *Psychological Review* 58, 313-323.

[8] Cabrales A. (1996): "Stochastic replicator dynamics", Universitat Pompeu Fabra, mimeo.

[9] Cross J. (1973): "A stochastic learning model of economic behavior", *Quarterly Journal of Economics* 87, 239-266.

[10] Fudenberg D. and D. Levine (1997): "Learning in games", forthcoming, *European Economic Review*, Papers and Proceedings.

[11] Gale J., K. Binmore and L. Samuelson (1995): "Learning to be imperfect: the ultimatum game", *Games and Economic Behavior* 8, 56-90.

[12] Hofbauer J. and J. Weibull (1996): "Evolutionary selection against dominated strategies", *Journal of Economic Theory* 71, 558-573.

[13] Hurkens S. (1995): "Learning by forgetful players", *Games and Economic Behavior* 11, 304-329.

[14] Kohlberg E. and J.-F. Mertens (1986): "On the strategic stability of equilibria", *Econometrica* 54, 1003-1037.

[15] Maynard Smith J. (1982): *Evolution and the Theory of Games*, Cambridge University Press (Cambridge UK).

[16] Nash J. (1950): "Non-cooperative games", PhD thesis, Princeton University.

[17] Nöldeke G. and L. Samuelson (1993): "An evolutionary analysis of backward and forward induction", *Games and Economic Behavior* 5, 225-254.

[18] Ritzberger K. and J. Weibull (1995): "Evolutionary selection in normal-form games", *Econometrica* 63, 1371-1400.

[19] Samuelson L. and J. Zhang (1992): "Evolutionary stability in asymmetric games", *Journal of Economic Theory* 57, 363-391.

[20] Sanchirico C. (1996): "A probabilistic model of learning in games", *Econometrica* 64, 1375-1393.

[21] Schlag K. (1997): "Why imitate, and if so, how? A bounded rational approach to multi-armed bandits", forthcoming, *Journal of Economic Theory*.

[22] Tan T. and S.R. Werlang (1988): "The Bayesian foundations of solution concepts of games", *Journal of Economic Theory* 45, 370-391.

[23] Taylor P. (1979), "Evolutionarily stable strategies with two types of player", *Journal of Applied Probability* 16, 76-83.

[24] Taylor P. and L. Jonker (1978), "Evolutionary stable strategies and game dynamics", *Mathematical Biosciences* 40, 145-156.

[25] Weibull J. (1995): *Evolutionary Game Theory*, MIT Press (Cambridge, MA).

[26] Young P. (1997): "Individual learning and social rationality", forthcoming, *European Economic Review*, Papers and Proceedings.