Economic Darwinism

Birgitte Sloth  
Institute of Economics  
University of  
Southern Denmark  
Campusvej 55  
DK-5230 Odense C  
DENMARK

Hans Jørgen Whitta-Jacobsen*  
Institute of Economics  
University of Copenhagen  
Studiestraede 6  
DK-1455 Copenhagen K  
DENMARK  
jacob@econ.ku.dk

Abstract

This paper examines the idea of Milton Friedman and others that the foundations for behavioral assumptions usually made in economics should be found in “economic elimination of the unfit”, that is, replacement of (individuals with) behavior that does not accord with principles and concepts such as profit maximization or Nash equilibrium. We consider playing-the-field, symmetric games played recurrently between individuals. Behavior is governed by two forces. One is “Darwinian selection”: if current behavior leads to payoff differences, at least individuals with behavior yielding lowest payoff have strictly positive probability of being replaced by individuals with arbitrary behavior. The other is “mutation”: any individual has at any point in time a strictly positive, very small probability of shifting to an arbitrary behavior. We define a static, intuitively motivated solution concept as well as a solution concept explicitly rooted in a dynamic process capturing these forces. We establish a close relation between the two concepts, which are often equivalent. Using the relatively manageable static solution concept, we identify the economic implications of Darwinian selection and mutation in a number of specific games of economic interest as well as in general games with externalities. A main conclusion is that if externalities are positive (negative), “economic Darwinism” implies even more under- (over-) activity than does Nash equilibrium: not only is economic Darwinism not supportive of Nash equilibrium, it systematically leads to socially worse outcomes.

JEL Classification: C72

Keywords: Darwinian Evolution, Competitive Equilibrium, Oligopoly Games, Evolutionary Stability, Social Welfare.

*Corresponding author
1 Introduction

An interesting view in economics asserts that the foundations for behavioral assumptions such as profit or utility maximization at the individual level or behavior in accordance with Nash equilibrium at the strategic level should be found in evolutionary selection against those who do not behave as assumed. This view, which could be called economic Darwinism, is originally discussed in Alchian (1950), Enke (1951) and Friedman (1953). The idea is that owners and managers of firms (and other economic agents) are guided by simple rules of thumb and markets select against rules (or managers) that lead to “inappropriate” behavior, i.e., to relatively low profits. This happens, for instance, by capital seeking to the more profitable firms. Since inappropriate behavior thus disappears, the economist can safely analyze as if active individuals maximize profits or behave in accordance with Nash equilibrium etc., although no one does so consciously.

This paper investigates explicitly the consequences of economic Darwinism. We consider symmetric games and envisage that such a game is played recurrently by individuals from a population. Any individual participating is locked at a fixed strategy except for random shifts. In each round the game is played once “all against all”, so each individual participating in the game in a specific round is “playing-the-field” as opposed to a situation with “pairwise contests”. Behavior is determined by two forces. One is Darwinian selection: if current behavior leads to payoff differences, at least individuals with behavior yielding lowest payoff have strictly positive probability of being replaced by individuals with arbitrary behavior. The other is mutation: for any individual there is at any point in time a small positive probability that the current behavior is replaced by an arbitrary behavior. We investigate what population behavior results, in particular whether resulting behavior is in accordance with the prediction of Nash equilibrium. For this purpose we formulate two alternative solution concepts, a static one as well as a one explicitly rooted in a dynamic process, each meant to capture the forces of Darwinian selection and mutation.

The static solution concept is closely related to one suggested by Schaffer (1989) and further studied by Possajennikov (2003) and Alós-Ferrer and Ania (2002). The basic idea is that for a situation to be stable with respect to Darwinian selection,
everybody must obtain the same payoff (otherwise those with lowest payoff would tend to die out), and, furthermore, to be stable with respect to mutation, no single player should have a deviation (a mutation) that would lead to a situation where the deviating player obtains strictly more payoff than other, non-deviating players (since such a mutation would be successful and displace behaviors yielding lowest payoffs). A profile of strategies in accordance with these requirements we call a Darwinian equilibrium.\(^1\)

We then consider an explicit dynamic process. In each round of play the actions of the individuals involved in the game are listed in the current “state”. A basic dynamic process (a Markov chain on the state space) captures Darwinian selection: when there are payoff differences, behavior giving maximal payoff is not changed, behavior yielding lowest payoff disappears with positive probability and is replaced by arbitrary new behavior, while behavior giving between minimal and maximal payoff may or may not be displaced (with positive probability). A perturbed process captures Darwinian selection as well as rare mutation: in addition to the transitions between states as governed by the basic process, every individual has a small independent probability of shifting to an arbitrary behavior. When the probability of mutation is small, only particular states with particular patterns of behavior will be observed frequently. These are the so-called stochastically stable states, here also referred to as Darwinianly stable.

A main result establishes the connection between the static solution concept, Darwinian equilibrium, and the dynamic one, Darwinian stability. For a considered (large) class of games, Darwinianly stable states exist, and any such state prescribes behavior as in a Darwinian equilibrium provided such an equilibrium exists. Shortly, Darwinianly stable states are Darwinian equilibria. The relation between Darwinian equilibrium and Darwinian selection dynamics as defined in this paper for finite population, playing-the-field games is similar to the relation between the (static) concept of “evolutionary stable strategies, ESS” and the (dynamic) concept of “replicator dynamics”, respectively, as defined in evolutionary theory based on pairwise contests in large populations, see e.g. Weibull (1995). However, while it is a standard result that evolutionary stable strategies are contained in the set of as-\[^1\]Schaffer (1989) denotes related concepts by “evolutionary biology finite-population ESS” and “symmetric evolutionary equilibrium, SEE”.\[^2\]
ymptotically stable points for the replicator dynamic, we find the other way round that the Darwinianly stable states are included in the set of Darwinian equilibria (if this is non-empty), and inclusion can be strict as we show by an example. The result we obtain is useful for applications because the static solution concept, Darwinian equilibrium, is the more tractable one: if one finds (all the) Darwinian equilibria one has (all) the candidates for Darwinianly stable states and, in particular, if there is a unique Darwinian equilibrium, the two solution concepts coincide.

We utilize this equivalence to characterize Darwinian stability in a number of economic applications. The kind of behavior selected for is typically not in accordance with Nash equilibrium. Hence, one result of relevance for economics is that rationality and profit maximization are not generally supported by economic Darwinism, in so far as one considers Nash equilibrium as the outcome of profit (payoff) maximization by rational individuals in strategic situations.

Rather, behavior supported by Darwinian stability is in accordance with Nash equilibrium of a modified game where payoffs are defined relatively, such that a player’s payoff in the modified game is the excess of his payoff from the original game over the lowest payoff of a player. We show that in typical economic situations where externalities are either overall positive or overall negative, Darwinianly stable outcomes are worse than Nash equilibrium outcomes as regards efficiency, since they involve even more under- or over-activity. This is another and main result of economic relevance.

As mentioned Schaffer (1989) suggests a static evolutionary solution concept very similar to our Darwinian equilibrium. Schaffer studies the implications in Cournot duopoly games and finds equivalence with competitive (price taking) equilibrium.

Vega-Redondo (1997) studies a dynamic imitation-of-success based process with mutations for symmetric, n-player Cournot oligopoly games. The results of Vega-Redondo are generalized and extended by Schenk-Hoppé (2000) who still limits attention to Cournot oligopoly. The process of imitation and mutation considered by Vega-Redondo (and Schenk-Hoppé) differs essentially from ours of Darwinian selection and mutation although both tend to change behavior in direction of actions that currently do not perform too poorly. According to Vega-Redondo’s process good behavior is imitated, whereas according to our process bad behavior dies out
and is replaced by arbitrary behavior. Hence, according to imitation, whenever a type of behavior is revised it is changed into a (better) behavior that is already being used by an individual, so that the set of used behaviors can only shrink except for mutation. According to Darwinian selection vanishing (relatively poorly performing) behavior is replaced by arbitrary behavior, so the set of used strategies can increase even without mutation. In the case of Cournot oligopoly the two processes lead to the same outcome, namely the competitive equilibrium, but as we explain in Section 6.1 the guiding principle behind this result is not ”price equal to marginal costs”, but rather ”relative payoff maximization”. Therefore accordance with competitive equilibrium is not a general implication, but a coincidence for Cournot oligopoly.

Schipper (2003) considers the same type of imitation process as Vega-Redondo and extends the analysis to a class of games more general than Cournot oligopoly, namely games which have an “aggregator property” and are “quasi-submodular”. Like we do, Schipper establishes a connection between a dynamically rooted solution concept, stochastic stability with respect to the process of imitation and mutation, and a static one called “Walrasian equilibrium” by Schipper, also referred to as an equilibrium in “aggregate-taking strategies” by, e.g. Possajennikov (2003). An aggregate-taking strategy is defined by each individual maximizing payoff without taking into account its influence on “aggregate behavior”. There is a close connection between an equilibrium in aggregate-taking strategies and our Darwinian equilibrium for games which do have the aggregator property, as we return to in Section 8. Nevertheless, the general implication of economic Darwinism is not equilibrium in aggregate-taking strategies because (symmetric) games do not have to have the aggregator property.

Our paper also has a relation to the contributions of Blume and Easley (1992), (1999), and (2002), Sandroni (2000), and Becker (2004), in part building on Winther (1964) and (1971) and on Nelson and Winther (1982). All these assume some version of “survival of the fittest” or “elimination of the unfit”, but they consider “markets” whereas we consider “games”. In markets, where agents react to given prices, profit-(payoff-) maximizing behavior is well-defined whereas in games payoff maximizing behavior can only be defined given actions of the opponents. Hence, for markets one can ask if Darwinian dynamics support profit maximizing behavior (as was,
perhaps, the original idea of, e.g., Friedman (1953)), while for games the appropriate question is whether Darwinian dynamics support a strategic equilibrium concept such as Nash equilibrium.\(^2\) The fundamental difference to this strand of literature thus lies in the object considered and in the question asked. Nevertheless there is some agreement in conclusion. For instance, Blume and Easley (2002) do not find accordance between long-run behavior arising from evolution based on “natural selection” and traditional equilibrium based on profit maximization, where we do not find accordance between long-run behavior arising from Darwinian dynamics and Nash equilibrium.

In Section 2 we give our notation for symmetric games, Section 3 contains the definition and a characterization of Darwinian equilibrium, and Section 4 defines the dynamic process and Darwinian Stability. In Section 5 we give our main general result on the relation between Darwinian equilibrium and Darwinian Stability. In Section 6 we use the characterization to identify Darwinianly stable states in different games of economic interest. Section 7 gives our result on the efficiency consequences of economic Darwinism, while Section 8 discusses the relation to games with the aggregator property. Section 9 offers some concluding remarks.

2 Symmetric games

We consider a game with a fixed set of players, \(N = \{1, \ldots, n\}\). The non-empty strategy set of each player \(i\) is \(S \subseteq \mathbb{R}\). The set of strategy profiles is \(\Omega = S^n\). A profile (element in \(\Omega\)) will be denoted by \(\omega\) or \((s_1, \ldots, s_n)\) or \((s_i, s_{-i})\). Payoff functions are \(u_i : \Omega \to \mathbb{R}, i = 1, \ldots, n\).

In an evolutionary context it is natural to focus on symmetric games. We make the very weak symmetry assumption, that if players do the same then they obtain the same payoff:

**Assumption 1.** If \(\omega = (s_1, \ldots, s_n)\) is such that \(s_i = s_j\), then \(u_i(\omega) = u_j(\omega)\).

Assumption 1 only imposes restrictions on payoffs arising from “diagonal play”

\(^2\)As already hinted at, sometimes Nash equilibrium is viewed as resulting from rationality (payoff maximization for given conjectures on opponents), and “common knowledge” of the game and perhaps of the players’ conjectures, see, e.g., Aumann and Brandenburger (1995). In that case, discrepancy between the evolutionary outcome and Nash equilibrium can be interpreted as a lack of evolutionary support for the underlying principle of rationality.
and is clearly much weaker than usual symmetry assumptions imposing restrictions on “off-diagonal” payoffs as well. For our main exposition we only need Assumption 1, but in formulating the economic implications of economic Darwinism in Section 7 we apply the following stronger and more usual symmetry assumption:

Assumption 1’. There is a common payoff function $u : \Omega \rightarrow R$, such that $u_i(s_1, ..., s_i, ..., s_n) = u(s_i, s_{-i})$ for $i = 1, ..., n$, and $u$ has the property that $u(s_i, s_{-i}) = u(s_i, s'_{-i})$ for any permutation $s'_{-i}$ of the strategies in $s_{-i}$.

In words Assumption 1’ says that the payoff obtained by player $i$, when player $i$ uses strategy $s_i$ and the other players use a constellation of strategies $s_{-i}$, depends in the same way on $s_i$ and $s_{-i}$ independently of $i$, and is the same independently of how the strategies of $s_{-i}$ are distributed among the players different from $i$.

In contributions such as Alós-Ferrer and Ania (2002), Possajennikov (2003), Schipper (2003), and Leininger (2004), Assumption 1’ is implied by the assumption that there is a common utility function, $v : \mathbb{R}^2 \rightarrow \mathbb{R}$, and an aggregator function, $A : \Omega \rightarrow \mathbb{R}$, such that $u_i(s_i, s_{-i}) = v(s_i, A(s_i, s_{-i}))$ for all players $i$, and $A$ fulfills $A(\omega) = A(\omega')$ for all permutations $\omega'$ of $\omega$. When the payoff functions have this special form we say that the game has the aggregator property.3 The aggregator property is featured by many economic examples (see Sections 6.1-6.4 below), but Assumption 1’ and in particular Assumption 1 allow for much more.

3 A game has the aggregator property in our terminology if it is an aggregative game in the sense of Corchón (1994) and is symmetric in the sense of Assumption 1’.

3 Darwinian equilibrium

This section formulates a static equilibrium concept meant to capture the elimination of the least fit when mutations occur. The definition is closely related to the one suggested by Schaffer (1989).

A first requirement for evolutionary stability in a Darwinian sense is that all players obtain the same payoff, since if there are payoff differences a behavior yielding lowest payoffs will tend to die out. Second, there should be stability with respect to mutations, such that if a single player (by mutation) shifts behavior then the new behavior will tend to die out because it obtains least payoff in the collection of behavior after the switch. These requirements lead to the following definition.
\textbf{Definition 1} A strategy profile \((s_1^*, \ldots, s_n^*) \in \Omega\) is a Darwinian equilibrium if for all \(i, s_i\) and \(j\): \(u_i(s_i, s_{-i}^*) \leq u_j(s_i, s_{-i}^*)\), and it is a strict Darwinian equilibrium if, furthermore, for all \(i, s_i \neq s_i^*\) and \(j\): \(u_i(s_i, s_{-i}^*) < u_j(s_i, s_{-i}^*)\).

In terms of deviations, \((s_1^*, \ldots, s_n^*) \in \Omega\) is a Darwinian equilibrium if there is no \(i\) and \(s_i\) such that for some \(j\): \(u_i(s_i, s_{-i}^*) > u_j(s_i, s_{-i}^*)\), and it is strict if, furthermore, there is no \(i\) and \(s_i \neq s_i^*\) such that for some \(j\): \(u_i(s_i, s_{-i}^*) \geq u_j(s_i, s_{-i}^*)\).

Note that the definition of a Darwinian equilibrium implies that \(u_i(s_i^*, s_{-i}^*) = u^*\) for all \(i\), since if there are payoff differences then \(u_i(s_i, s_{-i}^*) > u_j(s_i, s_{-i}^*)\) is fulfilled for \(s_i = s_i^*\) for some \(i, j\).

Definition 1 does not require the strategy combination to be symmetric \((s_i^* = s^*\) for all \(i\)), but in an evolutionary context symmetric Darwinian equilibria are of most interest since the aim is to establish support for certain types of “widespread behavior” or “conventions”.

Because of Assumption 1, the requirement for a symmetric Darwinian equilibrium is that no single player has a deviation that gives the player strictly more payoff than all the other players. This gives a useful characterization of symmetric Darwinian equilibria as symmetric Nash equilibria of an appropriately modified game, as also noted by Schaffer (1989) and others. Define \(\hat{u}_i(\omega) \equiv u_i(\omega) - \min_{j \neq i} u_j(\omega)\).

\textbf{Theorem 1} \(\omega^* = (s^*, \ldots, s^*)\) is a (strict) Darwinian equilibrium in the game under consideration if and only if \(\omega^*\) is a (strict) Nash equilibrium of the modified game where the set of players and the strategy set are the same and the payoff functions are \((\hat{u}_i), i = 1, \ldots, n\).

\textit{Proof.} A profile \(\omega^* = (s_1^*, \ldots, s_n^*)\) is a Nash equilibrium of the modified game if and only if for all \(i\) and \(s_i\), \(u_i(s_i^*, s_{-i}^*) - \min_{j \neq i} u_j(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*) - \min_{j \neq i} u_j(s_i, s_{-i}^*)\), and it is strict if furthermore the inequality is strict for all \(s_i \neq s_i^*\). In a symmetric equilibrium, where \(s_i^* = s^*\) for all \(i\), it follows from Assumption 1 that the left hand side in this inequality is zero, so the condition reads: for all \(i\) and \(s_i\), \(\min_{j \neq i} u_j(s_i, s_{-i}^*) \geq u_i(s_i, s_{-i}^*)\) (with strict inequality for \(s_i \neq s_i^*\) in case of strict Nash equilibrium), which is the same as: for all \(i\) and \(s_i\), \(u_i(s_i, s_{-i}^*) \leq u_j(s_i, s_{-i}^*)\) for all \(j\) (with strict inequality for \(s_i \neq s_i^*\) in case of strict Nash equilibrium), which defines (strict) Darwinian equilibrium. \(\square\)
Theorem 1 is useful for demonstrating existence of symmetric Darwinian equilibrium for some games, see Theorem 2 below, and for finding or characterizing Darwinian equilibria as in Sections 6, 7 and 8. Note that for a (common) monotone transformation of the payoff functions \( u_i \), both Darwinian equilibrium of the considered game and Nash equilibrium of the modified game are unchanged. This can be of use when Theorem 1 is applied, as evidenced by the example of Section 6.2.

**Theorem 2** If \( S \) is compact and convex and payoff functions are such that \( u_i(s_i, s_{-i}) - \min_{j \neq i} u_j(s_i, s_{-i}) \) is continuous in \((s_i, s_{-i})\) and quasi concave in \(s_i\) for all \(i\), then a symmetric Darwinian equilibrium exists. If, furthermore, \( u_i(s_i, s_{-i}) - \min_{j \neq i} u_j(s_i, s_{-i}) \) is strictly quasi concave, then all Darwinian equilibria are strict.

**Proof.** For each strategy \( s \in S \) consider the set \( \beta(s) = \arg \max_{s' \in S} \hat{u}_i(s', s, s, ..., s) \), i.e., the set of best replies of a player in the modified game when everybody plays \(s\). By assumption the function \( \hat{u}_i(\cdot, s, s, ..., s) = u_1(\cdot, s, s, ..., s) - \min_{j \neq 1} u_j(\cdot, s, s, ..., s) \) is continuous and quasi concave, implying that \( \beta(s) \) is non-empty and convex. By Berge’s maximum theorem we have that \( \beta(\cdot) \) is upper-hemi-continuous, and then by Kakutani’s fixed point theorem there exists a strategy \( s^* \in S \), such that \( s^* \in \beta(s^*) \), i.e., a symmetric Nash equilibrium of the modified game, which by Theorem 1 is a symmetric Darwinian equilibrium. In case \( u_i(s_i, s_{-i}) - \min_{j \neq i} u_j(s_i, s_{-i}) \) is strictly quasi concave \( \beta(s) \) is in fact a continuous function, and Brouwer’s fixed point theorem is sufficient to ensure existence. Further, if \( u_i(s_i, s_{-i}) - \min_{j \neq i} u_j(s_i, s_{-i}) \) is strictly quasi concave, then \( \max_{s_i' \in S_i} [u_i(s_i', s_{-i}) - \min_{j \neq i} u_j(s_i', s_{-i})] \) has a unique solution for each \(s_{-i}\), such that if \((s_i, s_{-i})\) is a DE, then for all \(s'_i \neq s_i\) it must be the case that \( u_i(s'_i, s_{-i}) - \min_{j \neq i} u_j(s'_i, s_{-i}) < u_i(s_i, s_{-i}) - \min_{j \neq i} u_j(s_i, s_{-i}) = 0 \), such that \( u_i(s'_i, s_{-i}) < \min_{j \neq i} u_j(s'_i, s_{-i}) \), implying that for all \(s'_i \neq s_i\) and all \(j \neq i\) we have \( u_i(s'_i, s_{-i}) < u_j(s'_i, s_{-i}) \) implying that \((s_i, s_{-i})\) is a strict Darwinian equilibrium. \( \square \)

## 4 Darwinian stability

For our explicit dynamic analysis we have, for technical reasons, to assume that the strategy set \( S \) is finite. We assume that the finite game considered is such that “equal payoffs imply equal strategies”:

**Assumption 2.** If \( u_i(s_1, ..., s_n) = u \) for all \(i\), then \( s_1 = \cdots = s_n \).
Given finiteness, Assumption 2 is fulfilled generically: Consider a given finite grid $S$. Assume that the payoff functions $(u_i)$ are such that for some strategy profiles with strategy differences all players get the same payoff. Then almost any (small) perturbation of the collection $(u_i)$ will imply payoff differences for these profiles (also when the perturbation must respect symmetry in the form of Assumption 1 or Assumption 1'). Given finiteness, Assumption 2 is therefore little restrictive, but highly simplifying as will appear.

The $n$-player game considered is assumed to be played recurrently among individuals taking the positions of the game. The outcome in each round is described by a state that lists the actions taken in each of the $n$ player positions. Hence a state has the structure of a strategy profile and will be denoted by $\omega = (s_1, ..., s_n) \in \Omega$. The state space $\Omega$ is finite.

On the state space, $\Omega$, we define a dynamic process capturing the two essential forces of Darwinian selection and mutation.

According to Darwinian selection individuals who perform badly in terms of payoff, the unfit, are sometimes displaced by other individuals who take up a random action. We express Darwinian selection dynamics by a basic (unperturbed) Markov chain on $\Omega$ with a matrix $\Pi^0$ of transition probabilities, $\pi^0_{\omega\omega'}$ ($0 \leq \pi^0_{\omega\omega'} \leq 1$ for all $(\omega, \omega')$, and $\sum_{\omega'} \pi^0_{\omega\omega'} = 1$ for all $\omega$), requiring:

1. If $\omega = (s_1, ..., s_n)$ is such that all players obtain the same payoff ($u_i(\omega) = u_j(\omega)$ for all $i, j$), then $\pi^0_{\omega\omega} = 1$.

2. If $\omega = (s_1, ..., s_n)$ is such that there are payoff differences ($u_i(\omega) \neq u_j(\omega)$ for some $i, j$), then for any state $\omega' = (s'_1, ..., s'_n)$: a) If $\pi^0_{\omega\omega'} > 0$, it must be for all $i \in \arg \max_{i \in N} u_i(\omega)$, that $s'_i = s_i$, and b) if $s'_i = s_i$ for all $i \notin \arg \min_{j \in N} u(s_j, s_{-j})$, then $\pi^0_{\omega\omega'} > 0$.

The restrictions on $\Pi^0$ expressing Darwinian selection are that if everybody gets the same payoff in $\omega$, then one stays in $\omega$, and if there are payoff differences then strategies with highest payoff are not displaced, while strategies with lowest payoff disappear and are taken over by other strategies with positive probability.

We have expressed Darwinian selection directly in terms of restrictions on the transition probabilities, $\pi^0_{\omega\omega'}$, rather than in terms of an explicit selection process.
with implications for the transition probabilities. The following are explicit selection processes that imply our assumptions: Assume that in each round only player positions that “receive a selection draw” can change strategy compared to the previous round, and that the consequence of a selection draw is that an arbitrary strategy is chosen according to a given probability distribution with full support on $S$. One process is defined by everybody with lowest payoff in the last round, and no other, having a given, strictly positive (and independent) probability of receiving a selection draw. Another possibility is that everybody who did not get highest payoff in the last round, and no other, has a given (independent) probability of receiving a selection draw. There are many other possibilities, for instance some where the probability of receiving a selection draw depends negatively on payoff in “softer”, more monotone ways than for the two processes just described.

The mutation process is added as follows: From the state chosen by the Darwinian selection dynamic, $\Pi^0$, there is for each player and for each $s \in S$ (independently) a small probability $\varepsilon > 0$ of switching to $s$. Each such switch is the result of a mutation. This defines a modified Markov chain $\Pi^\varepsilon$, where the transition probability from $\omega$ to $\omega''$ is defined as follows: For a pair of states $\omega', \omega''$, let $k(\omega', \omega'')$ be the number of players behaving differently in the two states, i.e., $k(\omega', \omega'') = \# \{ i | s'_i \neq s''_i \}$. Thus, $k(\omega', \omega'')$ is the number of mutations involved in the transition from $\omega$ to $\omega''$ via $\omega'$. Then, $\pi^\varepsilon_{\omega \omega''} = \sum_{\omega'} \pi^0_{\omega \omega'} \pi^0_{\omega' \omega''} \varepsilon^{k(\omega', \omega'')}$.

It is obvious that $\Pi^\varepsilon$ is ergodic. Furthermore, $\Pi^\varepsilon$ is a regular perturbation of $\Pi^0$ in the sense of Young (1993), i.e., $\Pi^\varepsilon \to \Pi^0$ as $\varepsilon \to 0$, and for each transition $\omega \omega''$ for which $\pi^0_{\omega \omega''} = 0$, there is a well-defined order by which $\pi^\varepsilon_{\omega \omega''} \to 0$ as $\varepsilon \to 0$. Here this order is given by the minimal number of mutations required to go from $\omega$ to $\omega''$, i.e., by $\min_{\omega', \pi^0_{\omega \omega'} > 0} k(\omega', \omega'')$. This number (which is equal to zero when $\pi^0_{\omega \omega''} > 0$) is referred to as the resistance in the transition from $\omega$ to $\omega''$.

In the following we state some standard results from the theory of Markov chains needed for our purposes. The results mentioned can all be found in Young (1993)

---

4The mutation process is that each individual mutates (independently) with probability $\varepsilon \cdot \#S$, and in case of mutation a random strategy is picked from $S$ according to the uniform probability distribution. We could, however, be much more general without affecting results. For instance different players could mutate with different probabilities and different strategies could have different probabilities when mutations occur.

5It is a finite Markov chain and irreducible (since from any state one can go to any other state in one step by appropriate mutation), and aperiodic (since for any state one can stay in that state by appropriate mutation).
or Freidlin and Wentzel (1984), see also Young (1998).

Since $\Pi^\varepsilon$ is ergodic it has a unique invariant (stationary) distribution $\mu^\varepsilon$, i.e., a probability distribution over $\Omega$ fulfilling $\mu^\varepsilon \Pi^\varepsilon = \mu^\varepsilon$. In the long run the relative frequencies by which states are visited converge with probability one to the probabilities of $\mu^\varepsilon$. Since our interest is in the process with both Darwinian selection and mutation for small mutation probability, we will be interested in the limit distribution, $\mu^0 = \lim_{\varepsilon \to 0} \mu^\varepsilon$. This limit distribution exists and is invariant for $\Pi^0$ ($\mu^0 \Pi^0 = \mu^0$). Thus, the states that have strictly positive probability according to $\mu^0$ are the only states that will be observed frequently for (very) small mutation probability. Usually these states are termed stochastically stable states. In our context a slight renaming is appropriate:

**Definition 2** A state $\omega$ is Darwinianly stable if and only if $\mu^0(\omega) > 0$.

It follows from the existence of $\mu^0$ that Darwinianly stable states exist. Stochastically (Darwinianly) stable states are characterized as follows:

An absorbing set is a subset $M \subset \Omega$ of states that is closed with respect to finite chains of transitions with positive probability according to $\Pi^0$, that is, for all $\omega \in M$, $\omega' \notin M$, one has $\pi^0_{\omega \omega'} = 0$, and for all $\omega, \omega' \in M$, there are states $\omega_1, ..., \omega_m \in M$, such that $\pi^0_{\omega \omega_1} > 0$, $\pi^0_{\omega_1 \omega_2} > 0$, ..., $\pi^0_{\omega_m \omega'} > 0$. In our case, the absorbing sets are particularly simple, being exactly all the singleton sets of form $\{(s, s, ..., s)\}$, where $s$ is an arbitrary strategy in $S$:

**Lemma 1** For any $s \in S$, the set $\{(s, s, ..., s)\}$ is absorbing, and there are no other absorbing sets.

**Proof.** In a state $\omega = (s, s, ..., s)$ all players obtain the same payoff by Assumption 1. Hence $\pi^0_{\omega \omega} = 1$, and $\{\omega\}$ is absorbing.

Consider a state $\omega = (s_1, ..., s_n)$, where players do not all use the same strategy. By Assumption 2, they do not all obtain the same payoff. With positive probability according to $\Pi^0$, all the players with minimal payoff switch to one of the strategies in $\{s_i\}_{i \in N}$ that did not yield minimal payoff, so in the resulting state fewer different strategies are used. If not all players use the same strategy in the resulting state, then the argument can be repeated. It therefore has positive probability according to $\Pi^0$ that in a finite number of rounds a state of the form $(s, s, ..., s)$ is reached,
and such a state is absorbing. Hence, no state with action differences can be in an absorbing set and therefore no other absorbing sets than the singleton sets of form \{(s, s, ..., s)\} exist. □

We will talk simply of absorbing states (not sets). Since an invariant distribution for \(\Pi^0\) can only attach positive probability to absorbing states it follows that only absorbing states can be stochastically stable. This leaves us with many candidates, but stochastically stable states can be further characterized.

Above we defined the resistance in a transition \(\omega \rightarrow \omega''\) from one state to another as the integer number \(\min_{\omega': \pi^0(\omega, \omega') > 0} k(\omega', \omega'')\), i.e., as the minimal number of mutations required to go from \(\omega\) to \(\omega''\) in one step. Define the resistance in a transition from one absorbing state, \(\omega\), to another, \(\omega''\), as the minimal total resistance (minimal number of mutations) required to go from \(\omega\) to \(\omega''\), possibly indirectly over other states \(\omega'\) which do not have to be absorbing.

For any absorbing state \(\omega\), define an \(\omega\)-tree as a directed graph on the set of absorbing states, such that for any absorbing state \(\omega' \neq \omega\) there is exactly one path in the graph leading from \(\omega'\) to \(\omega\). If there are \(k\) absorbing states, then any \(\omega\)-tree contains \(k - 1\) arcs. For any given \(\omega\)-tree define its total resistance as the sum of all the resistances over the arcs in the tree. For any absorbing state \(\omega\) define the stochastic potential as the minimal total resistance over all \(\omega\)-trees.

Young (1993) proves that the Darwinianly (stochastically) stable states are exactly the absorbing states with minimal stochastic potential. We use this in the next section to establish a very close connection between Darwinian equilibrium and Darwinian stability.

## 5 Main general result

In this section we (still) impose everywhere Assumption 1, finiteness of \(S\), and Assumption 2. In a Darwinian equilibrium all players get the same payoff. By Assumption 2 then, any Darwinian equilibrium must be symmetric, i.e., of the form \((s^*, ..., s^*)\). Furthermore, any Darwinian equilibrium must be strict, since a deviating player cannot obtain the same payoff as the other players and must therefore obtain strictly less. Given a Darwinian equilibrium, \((s^*, ..., s^*)\), we say that the absorbing state \(\omega^* \equiv (s^*, ..., s^*)\) corresponds to the Darwinian equilibrium.
**Theorem 3** If a Darwinian equilibrium exists, then any Darwinianly stable state corresponds to a Darwinian equilibrium.

By assumption there is at least one Darwinian equilibrium, so there is at least one state, \( \omega^* = (s^*,...,s^*) \), corresponding to a Darwinian equilibrium. The proof proceeds by showing, in two lemmas, that the resistance in the transition from such an \( \omega^* \) to any other absorbing state is at least two, while from any absorbing state, \( \omega = (s,...,s) \), that does not correspond to a Darwinian equilibrium, the resistance in the transition from \( \omega \) to any state \( \omega^* \) that does correspond to a Darwinian equilibrium is one.

**Lemma 2** The resistance in a transition from an absorbing state, \( \omega^* \), corresponding to a Darwinian equilibrium to any other absorbing state, \( \omega \), is at least two.

**Proof.** We show that it takes at least two mutations to go from an absorbing state \( \omega^* = (s^*,...,s^*) \) corresponding to a Darwinian equilibrium to a different absorbing state. If there is only one mutation from \( \omega^* \), then the resulting state has (at most) one player using a strategy \( s \neq s^* \), while all the remaining players still use \( s^* \). By Assumption 1, all the players using \( s^* \) get the same payoff. Furthermore, since \( (s^*,...,s^*) \) is a Darwinian equilibrium, and different strategies are used, it follows (from Assumption 2) that the player who plays \( s \) gets strictly less payoff than the other players who get the same. So, according to \( \Pi^0 \) it has probability one that in the next state all the other players still play \( s^* \), while the last player plays some \( s' \), and it has positive probability that \( s' = s^* \). If \( s' \neq s^* \), then the same happens again with positive probability of reaching \( \omega^* \) in the next step. As long as \( \omega^* \) has not yet been reached there is in each round positive probability of reaching \( \omega^* \) in the next round. Let the minimal of these probabilities (over all states where the last player has not yet come to play \( s^* \)) be \( \pi^\text{min} \). Then over \( T \) rounds, the probability of not reaching \( \omega^* \) is at most \((1 - \pi^\text{min})T\), which goes to zero as \( T \) goes to infinity. It thus has probability one according to \( \Pi^0 \) to eventually return to \( \omega^* \). One cannot get from \( \omega^* \) to another absorbing state by just one mutation, so the resistance from \( \omega^* \) to any other absorbing state \( \omega \) is at least two. □

**Lemma 3** The resistance in a transition from an absorbing state, \( \omega \), not corresponding to a Darwinian equilibrium to any \( \omega^* \) corresponding to a Darwinian equilibrium is one.
Proof. We show that from an absorbing state \( \omega = (s, ..., s) \) not corresponding to a Darwinian equilibrium, there is a single mutation leading to a state from which it has positive probability according to \( \Pi^0 \) to go to any \( \omega^* = (s^*, ..., s^*) \) corresponding to a Darwinian equilibrium. Since \((s, ..., s)\) is not a Darwinian equilibrium, there is a player position \( i \) and an action \( s'_{i} \), such that \( u_i(s'_i, s_{-i}) > u_j(s'_i, s_{-i}) \) for all \( j \neq i \) (Assumption 1 ensures that all \( j \neq i \) get the same payoff). Assume that player \( i \) mutates and plays \( s'_i \). Now with positive probability according to \( \Pi^0 \), the process reaches in the next round a state \((s'_i, s^*, ..., s^*)\), where player 1 still plays \( s'_i \) while all other players play the strategy \( s^* \) of the Darwinian equilibrium in question. In this new state the players different from \( i \) all get the same payoff (Assumption 1) and player \( i \) obtains strictly less (because \((s^*, ..., s^*)\) is a Darwinian equilibrium, and by Assumption 2 there cannot be payoff equality when different strategies are used). With positive probability according to \( \Pi^0 \) the process will therefore in the next round reach the state \((s^*, ..., s^*)\). This means that the resistance from \( \omega \) to any \( \omega^* \) is one. \( \square \)

Proof of Theorem 3. Consider an \( \omega \)-tree, where \( \omega \) does not correspond to a Darwinian equilibrium. Let \( \omega^* \) be a an absorbing state that does correspond to a Darwinian equilibrium. Change the \( \omega \)-tree in the following way: remove the transition out of \( \omega^* \) (by Lemma 2 this has resistance at least 2), and add a transition from \( \omega \) to \( \omega^* \) (by Lemma 3 this has resistance 1). Thereby an \( \omega^* \)-tree with strictly lower resistance than the \( \omega \)-tree has been constructed. This shows that the stochastic potential of \( \omega^* \) is strictly lower than the stochastic potential of \( \omega \), so that \( \omega \) cannot be stochastically stable. Thus, any stochastically stable state must correspond to a Darwinian equilibrium. \( \square \)

Theorem 3 does not say that in games with more than one Darwinian equilibrium all states corresponding to Darwinian equilibria are stochastically stable.\footnote{Section 6.5 below provides an example of a game with several Darwinian equilibria out of which exactly one is Darwinianly stable.} For games with a unique Darwinian equilibrium we have:

Corollary 1 If the game has a unique Darwinian equilibrium \((s^*, ..., s^*)\), then the unique Darwinianly stable state is \( \omega^* = (s^*, ..., s^*) \).
This characterization of Darwinian stability in games with a unique Darwinian equilibrium is useful. If, in a specific game, we find exactly one symmetric Darwinian equilibrium, then this is the only stable situation according to the dynamics of Darwinian selection and mutation.

6 Examples

6.1 Cournot oligopoly

Consider a symmetric, quantity-setting Cournot oligopoly, where each of \( n \) firms has cost function \( C(q_i) \), where \( q_i \in S = [0, \infty) \) is the production of firm \( i \). The inverse demand curve is \( P(Q) \). The payoff functions are:

\[
 u_i(q_1, ..., q_n) = P\left(\sum_{h=1}^{n} q_h\right) q_i - C\left(q_i\right).
\]

This game has the aggregator property with \( \sum_{h=1}^{n} q_h \) as the aggregator. We assume differentiability with \( P' < 0 \), \( C' > 0 \), and \( C'' \geq 0 \). Furthermore we assume \( P(\infty) < C'(0) < P(0) \), which means that a unique “competitive equilibrium”, \( q_i = q^c \) for all \( i \) and \( P(nq^c) = C'(q^c) \), exists.

Consider a symmetric situation where everybody sells the quantity \( q \). For any oligopolist \( i \), the change in profit from increasing production a little is \( \partial u_i/\partial q_i = P'(nq)q + P(nq) - C'(q) \), and a requirement for (interior) symmetric Nash equilibrium is that \( \partial u_i/\partial q_i \) is zero. The effect on the profits of the other oligopolists \( j \), from \( i \) increasing production is \( \partial u_j/\partial q_i = P'(nq)q \), and a requirement for Darwinian equilibrium (Theorem 2) is that \( \partial u_i/\partial q_i - \partial u_j/\partial q_i = 0 \), or \( P(nq) = C'(q) \). Hence, a symmetric Darwinian equilibrium has to be a competitive equilibrium. Obviously, under an appropriate assumption of concavity, if no oligopolist has an infinitesimal deviation that gives him more profits than others, no oligopolist will have a large such deviation either. Hence, under such a condition the game has a unique symmetric Darwinian equilibrium and this is the competitive equilibrium.

For a version of the game where the strategy set is finite and \( q^c \in S \) and Assumption 2 is fulfilled, the unique Darwinian equilibrium is the symmetric one: \( q_i = q^c \) for all \( i \). The state \( \omega^c = (q^c, ..., q^c) \) is then by Theorem 3 the unique Darwinianly stable state.
The result, that Darwinian equilibrium and Darwinian stability lead to the competitive equilibrium in Cournot oligopoly, is related to the results of Schaffer (1989), Vega-Redondo (1997) and Schenk-Hoppé (2000). Accordance with competitive equilibrium is not, however, a general implication of economic Darwinism (and not of imitation dynamics with mutation as studied by the latter two references either). To see why, the intuition for our result is important.

Start from a Cournot-Nash equilibrium given by \( q_i = q^{ne} \) for all \( i \), and \( \partial u_i / \partial q_i = 0 \), or \( P'(nq^{ne})q + P(nq^{ne}) = C'(q^{ne}) \). Here the price, \( P(nq^{ne}) \), is above marginal cost, \( C'(q^{ne}) \). A unilateral increase of production from \( q^{ne} \) by one oligopolist will (of course) imply a loss of profit for the deviator, but the other oligopolists will lose more (strictly at \( q^{ne} \) the deviating player’s loss is of second order while the others’ are of first order). What happens is that all oligopolists’ payoffs are reduced to the same extent by a decreased price, but as long as price is above marginal cost the deviator is partly compensated by the positive marginal profit earned on the increased production (given the price). Hence, all the way up to the competitive equilibrium where price is equal to marginal cost, increases in production will imply increased relative profit for the deviator, so only at the competitive equilibrium no increases in relative profit can be obtained. This intuition should reveal that “price equal to marginal cost” is not the guiding principle here, rather it is “relative payoff maximization”. It just happens for the particular game of Cournot oligopoly that the possibilities for relative payoff increases are exhausted exactly when price equals marginal cost.

Our next example shows that price equal to marginal cost is not a general implication of economic Darwinism.

### 6.2 Differentiated product Bertrand oligopoly

Consider a symmetric, differentiated product, price setting oligopoly (of monopolistic competition). In accordance with optimization by a representative household with a CES utility function the demand function for each product type \( i \) is:

\[
D_i (p_1, ..., p_n) = \left( \frac{p_i}{P} \right)^{-\sigma} \frac{Y}{P}, \quad \sigma > 1,
\]
where $Y$ is the budget of the household, and $P$ is the price index:

$$P(p_1, ..., p_n) = \left( \frac{1}{N} \sum_{h=1}^{n} p_h^{1-\sigma} \right)^{\frac{1}{1-\sigma}}.$$

Each firm is the sole producer of one product type and sets its price $p_i \geq 0$. The payoff functions are:

$$v_i(p_i, p_{-i}) = \left( \frac{p_i}{P} \right)^{-\sigma} Y \left( p_i - c \right),$$

where $c$ is a constant unit cost. This game also has the aggregator property with the price index $P(p_1, ..., p_n)$ taking the place of the aggregator. The set of Darwinian equilibria is insensitive to the monotone transformation:

$$u_i(p_i, p_{-i}) \equiv \log v_i(p_i, p_{-i}) = -\sigma (\log p_i - \log P) + \log Y - \log P + \log (p_i - c).$$

Now,

$$\frac{\partial u_i}{\partial p_i} = -\sigma \left( \frac{1}{p_i} - \frac{\partial \log P}{\partial p_i} \right) - \frac{\partial \log P}{\partial p_i} + \frac{1}{p_i - c},$$

while for $j \neq i$:

$$\frac{\partial u_j}{\partial p_i} = -\sigma \left( -\frac{\partial \log P}{\partial p_i} \right) - \frac{\partial \log P}{\partial p_i}.$$

In the requirement for (interior) Darwinian equilibrium, $\partial u_i/\partial p_i - \partial u_j/\partial p_i = 0$, the terms involving $P$ cancel (so we do not use the exact formula for $P$ in the computation), leading to:

$$p_i = p^m = \frac{\sigma}{\sigma - 1} c \text{ for all } i.$$

This is the unique Darwinian equilibrium and hence, in an appropriate finite version of the game, the unique Darwinianly stable state is $(p^m, ..., p^m)$.

In this example Darwinian evolutionary forces do not lead to price equal to marginal cost. Rather $p^m$ is the “monopoly price” that results if the competitors optimize neglecting their influence on the price index, $P$, an assumption often made in macroeconomic applications and usually defended by $n$ being large. The evolutionary rationale given here does not rely on the number of competitors.

### 6.3 A public good game

Each of $n$ individuals has to decide on an amount $x_i \geq 0$ to contribute to the procurement of a public good. As a function of the total contribution, $X = \sum_{h=1}^{n} x_h$, 

17
a certain amount of the good is provided giving each individual a utility (in money terms) of \( G(X) \). The payoff functions are thus:

\[
u_i(x_1, \ldots, x_n) = G \left( \sum_{h=1}^{n} x_h \right) - x_i,
\]

where the aggregator property is again fulfilled. We assume differentiability and that \( G' > 0 \) and \( G'' \leq 0 \). Furthermore we assume that \( G'(\infty) < 1/n < G'(0) \), which implies that there is a unique symmetric social optimum, \( x_i = x^o > 0 \) for all \( i \), where \( G'(nx^o) = 1/n \) (the point where \( G(nx) - x \) is maximized).

From a symmetric situation, \( x_i = x \) for all \( i \), the change in payoff for player \( i \) from an infinitesimal increase in contribution is \( \partial u_i/\partial x_i = G'(nx) - 1 \). If \( G'(0) > 1 \), there is a unique symmetric Nash equilibrium, where each individual \( i \) contributes positively, \( x_i = x^{ne} > 0 \), and \( G'(nx^{ne}) = 1 \). Hence \( G'(nx^{ne}) > G'(nx^o) \), implying \( x^{ne} < x^o \). If \( G'(0) < 1 \), the unique Nash equilibrium is \( x_i = x^{ne} = 0 \) for all \( i \), again implying \( x^{ne} < x^o \). Positive contributions arise in Nash equilibrium if marginal products and utilities are sufficiently large for low levels of procurement.

The influence on payoffs of other players \( j \), when \( i \) increases his contribution, is \( \partial u_j/\partial x_i = G'(nx) \), so \( \partial u_i/\partial x_i - \partial u_j/\partial x_i = -1 \). This means that the unique symmetric Darwinian equilibrium is \( x_i = 0 \) for all \( i \). (Of course, in the considered game an increase in contribution can only affect other players’ payoffs more positively than own payoff, so one wouldn’t have to differentiate to reach this conclusion, but it is again instructive to see how the influence on the aggregator cancels).

In an appropriate finite version of the game where \( 0 \in S \) and Assumption 2 is fulfilled, the unique Darwinian stable state also has \( x_i = 0 \) for all \( i \). In this game economic Darwinism leads to an outcome, which is worse with respect to efficiency than the Nash equilibrium, which in turn is worse than the social optimum.

### 6.4 Tragedy of the commons

Each of \( n \) fishermen decides on a number of fishing hours, \( z_i \geq 0 \). The total catch will be \( F(Z) \geq 0 \), where \( Z = \sum_{h=1}^{n} z_h \). It is assumed that \( F' > 0 \) and \( F'' < 0 \), so average product or catch per hour, \( F(Z)/Z \), is decreasing and larger than marginal product, \( F'(Z) \), implying \( F'(Z)Z/F(Z) < 1 \) Each fisherman gets his own catch.
The payoff functions are thus:

\[ u_i(z_1, ..., z_n) = z_i \frac{F(\sum_{h=1}^{n} z_h)}{\sum_{h=1}^{n} z_h} - C(z_i), \]

where \( C \) is a cost function, \( C' > 0, C'' \geq 0 \). In the following we assume that the first order conditions are necessary and sufficient for maximum whenever stated.

Hence, the condition for a symmetric social optimum, \( z_i = z^o \) for all \( i \) (maximizing \( F(nz)/n - C(z) \)), is \( F'(nz^o) = C''(z^o) \).

Starting at \( z_i = z \) for all \( i \), one has:

\[
\frac{\partial u_i}{\partial z_i} = F'(nz) \left( \frac{1}{n} + \frac{1}{F'(nz)n_z} \frac{n - 1}{n} \right) - C'(z) \\
= \frac{F(nz)}{n_z} - \frac{1}{n} \frac{F(nz)}{n_z} \left( 1 - \frac{F'(nz)n_z}{F(nz)} \right) - C'(z),
\]

and the condition for a symmetric Nash equilibrium, \( z_i = z^{ne} \) for all \( i \), is \( \partial u_i / \partial z_i = 0 \). From the two expressions for \( \partial u_i / \partial z_i \) one can see that \( \partial u_i / \partial z_i = 0 \) implies \( F'(nz^{ne}) < C'(z^{ne}) < F(nz^{ne})/(nz^{ne}) \). Hence, from the first inequality and \( F'(nz^o) = C''(z^o) \), one has \( z^o < z^{ne} \).

Still starting at \( z_i = z \) for all \( i \), one has:

\[
\frac{\partial u_j}{\partial z_i} = -\frac{1}{n} \frac{F(nz)}{n_z} \left( 1 - \frac{F'(nz)n_z}{F(nz)} \right) \text{ for } i \neq j,
\]

so \( \partial u_i / \partial z_i - \partial u_j / \partial z_i = F(nz)/n_z - C'(z) \). Hence the condition for a symmetric Darwinian equilibrium, \( z_i = z^{de} \) for all \( i \), is: \( F(nz^{de})/(nz^{de}) = C'(z^{de}) \). It then follows from \( C'(z^{ne}) < F(nz^{ne})/(nz^{ne}) \) (the second inequality above), that \( z^{ne} < z^{de} \).

In an appropriate finite version of the game, where \( z^{de} \in S \) and Assumption 2 is fulfilled, the state \( (z^{de}, ..., z^{de}) \) is the only Darwinian stable state. Economic Darwinism thus leads to even more overfishing than Nash equilibrium.

In our examples so far the outcome in Darwinian equilibrium corresponds to “optimization neglecting one’s influence on the aggregator". This is not the general implication of economic Darwinism, however, since there are games that do not have the aggregator property.
6.5 A game without the aggregator property

Each of $n \geq 4$ firms produces a commodity, and production depends on how much costly effort, $e_i$, the firm exerts. The firms are not competitors in a usual sense, but a productive externality (learning spillover) implies that a firm’s production depends positively on the production of the firm among the other firms that has the highest production (and thus exerts the most effort). The payoff functions are:

$$u_i(e_1, ..., e_n) = (e_i + a \max_{j \neq i} e_j)p - \frac{1}{2}e_i^2,$$

where $a \in [0, 1]$, and the price $p > 0$, are parameters, and it is required that $e_i \in [0, \infty]$. This game does not have the aggregator property.

Independently of the effort choices of other firms the best reply of any one firm $i$, given by $\partial u_i / \partial e_i = 0$, is $e_i = p$. Hence, the unique Nash equilibrium is the symmetric one, $e_i = p$ for all $i$.

The Nash equilibrium is not efficient. The best symmetric outcome where all firms use the strategy $e^o$ is found by:

$$\frac{\partial}{\partial e} \left[ (e + ae)p - \frac{1}{2}e^2 \right] = (1 + a)p - e = 0,$$

implying $e^o = (1 + a)p$. In Nash equilibrium too little effort is exerted, because each firm does not take the positive externality of its effort into account. Since only the maximal effort (among others) matters for the productive externality, the best symmetric outcome, $e_i = e^o$ for all $i$, is not overall efficient: a higher total payoff can be obtained in asymmetric situations where only one or two firms exert a lot of effort, but implementation of such a social optimum would require side payments.

To identify the Darwinian equilibria in this game we cannot rely on simple first order conditions, because $u_j$ is not differentiable with respect to $e_i$ for $j \neq i$ at symmetric points. However, right and left hand derivatives do exist and can be used to find the Darwinian equilibria:

$$\left( \frac{\partial u_i}{\partial e_i} - \frac{\partial u_j}{\partial e_i} \right)_{\partial e_i > 0} = p - e_i - a, \quad \left( \frac{\partial u_i}{\partial e_i} - \frac{\partial u_j}{\partial e_i} \right)_{\partial e_i < 0} = p - e_i.$$

The conditions for an interior Darwinian equilibrium are $\left( \frac{\partial u_i}{\partial e_i} - \frac{\partial u_j}{\partial e_i} \right)_{\partial e_i > 0} \leq 0$ and $\left( \frac{\partial u_i}{\partial e_i} - \frac{\partial u_j}{\partial e_i} \right)_{\partial e_i < 0} \geq 0$, which are fulfilled for $p - a \leq e_i \leq p$. Thus, there is a
continuum of symmetric Darwinian equilibria, $e_i = e^{de}$ for all $i$, where $p - a \leq e^{de} \leq p$, and the Nash equilibrium is among the Darwinian equilibria.

In an appropriate finite version of the game, where $p \in S$ and Assumption 2 is fulfilled, the unique Darwinianly stable state is $\omega^* = (p, ..., p)$ corresponding to the Darwinian equilibrium that is also Nash equilibrium. To see this first note that by Theorem 3, only states corresponding to Darwinian equilibria can be Darwinianly stable. Consider any state $(e^{de}, ..., e^{de})$ where $e^{de}$ is a Darwinian equilibrium and $e^{de} \neq p$. From this state simultaneous mutation by two players to $p$ will result in a state $\omega'$ where the mutants get the Nash-equilibrium payoff (since they can utilize the learning spill-over from each other), whereas the non-mutants get a strictly lower payoff (facing the same learning spill-over as the mutants, but exerting too little effort). Thus there will be positive probability according to the unperturbed process to go from $\omega'$ to $\omega^*$. From $\omega^*$, on the other hand, two mutations do not suffice to escape, since any two mutants will obtain strictly lower payoff than the remaining (at least two) players who still play $p$: two mutants who both play $e < p$, will (as the non-mutants) utilize the spill-over from $p$, but their effort levels will be inferior, etc.

This example illustrates the close relation between the static concept, Darwinian equilibrium, and the dynamically rooted concept, Darwinian stability, in a situation where the first does not coincide with “optimization neglecting one’s influence on the aggregator” (since the game does not have an aggregator).

7 Welfare implications of economic Darwinism

The examples of Section 6 indicate that with positive externalities economic Darwinism generally implies socially “too small contributions” to a degree at least as bad, and most often worse than, Nash equilibrium, while with negative externalities economic Darwinism leads to at least as much “over-activity” as Nash equilibrium. This is indeed a general property as we will now show. We think that this section points to the economically most important general implication of economic Darwinism.

We impose stronger assumptions than in Section 2, which is natural for our present purpose. First we assume that the strategy set, $S$, is a convex and closed
subset of $\mathbb{R}_+$, and that the payoff functions, $u_i$, are differentiable (until we “finitize” the game for equivalence between Darwinian equilibrium and Darwinianly stable states).

Second, we impose the stronger symmetry assumption, Assumption 1’. Note that all examples of Section 6 fulfill Assumption 1’. Without such a stronger symmetry assumption, there would be no reason to be concerned particularly with symmetric social optima below.

Third, we consider situations where externalities are overall positive, $\partial u_i / \partial s_h > 0$ for all $i \neq h$. Results for the case of overall negative externalities, $\partial u_i / \partial s_h < 0$ for all $i \neq h$, follow analogously.

Fourth, we assume that each of the concepts symmetric social optimum, symmetric Nash equilibrium and symmetric Darwinian equilibrium is unique and that the latter two are interior and fully characterized by first order conditions (considered below) and furthermore, these properties are insensitive to small perturbations of payoff functions (the latter property being fulfilled generically).

Welfare, $W(s)$, at a common strategy, $s$, is the common utility of all players at the strategy profile $(s, ..., s)$, that is, $W(s) \equiv u_i(s, ..., s)$ for any choice of $i$. The symmetric social optimum is: $s_i = s^o$ for all $i$, where $s^o$ maximizes $W(s)$. We allow for $s^o$ being non-interior (in the Bertrand example of Section 6.2, for instance, the social optimum is $p = \infty$).

Our final assumption is that $W$ is concave. The examples in Section 6.1 through 6.4 are all compatible with our assumptions.\footnote{In the Cournot example of Section 6.1, the welfare function, $W(q) = P(q)q - C(q)$, is concave under the additional assumption, $P''(Q)Q/P'(Q) > -2$. All other assumptions mentioned are fulfilled as the examples stand.}

The derivative of $W(s)$ is:

$$W'(s) = \sum_{h=1}^{n} \frac{\partial u_i}{\partial s_h}(s, ..., s),$$

where, from the assumed symmetry, $i$ can again be any player.

Define the “marginal product”, $m(s) \equiv \frac{\partial u_i}{\partial s_i}(s, ..., s)$, from a symmetric profile. Again, $i$ can be any player in this definition. The unique symmetric Nash equilibrium, $s_i = s^{ne}$ for all $i$, is given by $m(s^{ne}) = 0$. From our assumptions, $m(s)$ has to be strictly decreasing at $s^{ne}$. (First, $m(s)$ has to intersect strictly with the $s$-axis
at $s = s^{ne}$, since otherwise a small perturbation of payoff functions could imply non-existence of Nash equilibrium. Second, $m(s)$ must be decreasing by the second order condition).

Since, at the symmetric Nash equilibrium, $m(s^{ne}) = \frac{\partial u_i}{\partial s_i}(s^{ne},...,s^{ne}) = 0$, one has:

$$W'(s^{ne}) = \sum_{h \neq i} \frac{\partial u_i}{\partial s_h}(s^{ne},...,s^{ne}),$$

which is strictly positive because of positive externalities. Hence, $s^o > s^{ne}$, and obviously $W(s^{ne}) < W(s^o)$. This is just a restatement of the well-known result that with positive externalities, Nash equilibrium implies socially too little activity.

A symmetric Darwinian equilibrium, $s_i = s^{de}$ for all $i$, is given by the first order condition: $\frac{\partial u_i}{\partial s_i}(s^{de},...,s^{de}) - \frac{\partial u_i}{\partial s_i}(s^{de},...,s^{de}) = 0$, where the first term is $m(s^{de},...,s^{de})$, and, again because of symmetry, $i$ and $j$ can be any two (different) players. Hence,

$$m(s^{de}) = \frac{\partial u_i}{\partial s_i}(s^{de},...,s^{de}).$$

Positive externalities thus imply $m(s^{de}) > 0$. Since $m(s^{ne}) = 0$, and $m(s)$ is decreasing at $s = s^{ne}$ and $m(s)$ only has the one intersection with the axis at $s^{ne}$, one has that $s^{de} < s^{ne}$. Hence, the Darwinian equilibrium is further away from the social optimum than the Nash equilibrium implying, from the concavity of $W$, that $W(s^{de}) < W(s^{ne})$.

Imposing an appropriate further assumption of finiteness where all of $s^o$, $s^{ne}$ and $s^{de}$ are in $S$, and Assumption 2 is fulfilled, symmetric social optimum, symmetric Nash equilibrium, and symmetric Darwinian equilibrium are unchanged and the latter coincide with the unique Darwinianly stable state. This suffices for:

**Theorem 4** Under the assumptions of this section, if externalities are overall positive (negative), symmetric social optimum implies “higher” (“lower”) actions and higher welfare than does symmetric Nash equilibrium, and symmetric Nash equilibrium implies “higher” (“lower”) actions and higher welfare than do Darwinian equilibrium and Darwinian stability.
8 Games with the aggregator property

For games with the aggregator property, Alós-Ferrer and Ania (2002) and Possajennikov (2003) define an “aggregate-taking strategy” as the strategy of a symmetric profile where each player has maximized his payoff neglecting his influence on the aggregator $A$, that is, as a strategy $s^w$ such that for all $i, s_i \in S$: $v(s^w, A(s^w, ..., s^w)) \geq v(s_i, A(s^w, ..., s^w))$. Schipper (2003) uses the term (symmetric) “Walrasian equilibrium” for the same concept.

In games with the aggregator property, where both symmetric Darwinian equilibria and aggregate-taking strategies are characterized by first order conditions, they must coincide. This follows since the first order condition for $s^*$ being an aggregate-taking strategy is:

$$\frac{\partial}{\partial s} [v(s, A(s^*))] = v_1 = 0,$$

while $(s^*, ..., s^*)$ is a Darwinian equilibrium if and only if $s^*$ maximizes $u_i(s, s^*_{-i}) - \min_{j \neq i} u_j(s, s^*_j)$ with respect to $s_i$ for each $i$, this maximization problem being the same as:

$$\max_{s \in S} [v(s, A(s, s^*_{-i})) - v(s^*, A(s, s^*_{-i}))],$$

for which the first order condition is:

$$\frac{\partial}{\partial s} [v(s, A(s, s^*_{-i})) - v(s^*, A(s, s^*_{-i}))] = v_1 + v_2 \frac{\partial A}{\partial s} - v_2 \frac{\partial A}{\partial s} = v_1 = 0.$$

Schipper (2003) considers the dynamic process of imitation and mutations first considered by Vega-Redondo (1997). As explained in the introduction, this process is related to, but essentially different from our process of Darwinian selection and mutation. Schipper shows that if the game is quasi-submodular and has a unique aggregate-taking strategy, then there is a unique stochastically stable state and this corresponds to the aggregate-taking strategy (an extension of Vega-Redondo’s result for Cournot oligopoly). For the proof of Schipper’s result, submodularity is required because submodularity ensures that if all players play the same non-aggregate-taking strategy, then a player can increase his payoff by playing the aggregate-taking strategy, and then, without further mutations, it is possible by imitation to go to
a state where everybody plays the aggregate-taking strategy. For our process an assumption like submodularity is not required, since Darwinian selection allows for new strategies to emerge (without mutation).

Our results thus apply also for games that are not submodular. This is an important extension because many games of economic interest are not submodular, e.g., the price-setting oligopoly with strategic substitutes considered in Section 6.2. Furthermore, our results apply for symmetric games without the aggregator property.

9 Concluding remarks

In the presence of overall positive or negative externalities, outcomes arising from Darwinian evolutionary stability are even worse than Nash equilibrium outcomes.

It is not just, as with Nash equilibrium, that each individual’s behavior is insensitive to the way it affects other people’s payoffs. For the outcomes of economic Darwinism it is true that even if an increase in “effort” would benefit the individual undertaking it, the increase will not be undertaken unless it improves the individual’s relative position, that is, unless it benefits the individual more than it benefits other individuals. And even if it hurts the individual, it will be undertaken if it hurts other people more. This gives an increased tendency (as compared to Nash equilibrium behavior) to contributing too little in the presence of positive spillovers, and to exploiting too much in the presence of negative spillovers.

It is often argued that social institutions such as legal systems or customary norms and conventions are rooted in the fact that traditional selfish (Nash equilibrium) behavior would create socially too bad outcomes in standard social environments. In so far as observed behavior is more guided by Darwinian evolution than by selfish “rationality”, the argument in favor of social institutions stands even stronger.

The closer we are to the jungle, the more we need laws and conventions to restrict individual behavior.


10 References


