When will payoff maximization survive?

An indirect evolutionary analysis[*]

Werner Güth[1] and Bezalel Peleg[2]

[1] Max Planck Institute for Research into Economic Systems, Strategic Interaction Group, Kahlaische Straße 10, 07745 Jena, (e-mail: gueth@mpiew-jena.mpg.de)
[2] Center for Rationality and Interactive Decision Theory, The Hebrew University of Jerusalem, Feldman Building, Givat Ram, 91904 Jerusalem, Israel

Abstract. Survival of payoff maximization is the usual as if-justification for assuming rational economic agents. An indirect evolutionary analysis allows for stimuli which are not directly related to reproductive success although they affect behavior. One first determines the solution for all possible constellations of stimuli, and then the evolutionarily stable stimuli. Our general analysis confirms the special results of former studies that payoff maximization in case of commonly known stimuli requires either that own success does not depend on other's behavior or that other's behavior is not influenced by own stimuli. When stimuli are private information, one can derive similar necessary conditions.

Key words: (Indirect) evolution – Endogenous preferences – Rationality – Incomplete information

JEL classification: B4, C7, D0, D7, L1

1 Introduction

As in evolutionary biology, evolutionary game theory (see the survey of Hammerstein and Selten, 1994) assumes genetically determined behavior and tries to determine the evolutionarily stable genotype or behavior. Whereas this makes sense for primitive organisms such as plants, the assumption of genetically determined behavior is impossible for more highly developed species since they live in complex environments and therefore face far too many different choice problems.

The basic idea of the indirect evolutionary approach (Güth and Yaari, 1992), is to allow for an indirect dependence of behavior on genetically determined stimuli (in general "stimuli" can be any individual aspect determining the rules of the game; one prominent example is when this aspect defines individual preferences). More specifically, it is assumed that genetically determined stimuli define a game one has to solve in order to derive how behavior depends on these genetically determined stimuli - we refer to this as the first step of an indirect evolutionary analysis. Although we rely on rationality, as do most of the previous studies, it would be even more important to apply psychologically more convincing ideas when solving the game (see Güth and Kliemt, 1998; Güth, Kareev and Kliemt, 2001).

By inserting the solution into the material payoff function, one then knows how (reproductive) success depends indirectly on genetically determined stimuli via the solution of the game.[1] This defines an evolutionary game whose strategies are the genetically determined stimuli. As in usual evolutionary game theory, the second step of an indirect evolutionary analysis requires us to determine the evolutionarily stable strategy or stimuli. If stimuli are private information, behavior can, of course, depend only on the beliefs about the stimuli of others.

Previous studies (e.g. Bester and Güth, 1998; Güth, 1998; Güth and Huck, 1997; Güth and Kliemt, 2000; Possajennikov, 2000; Schaffer, 1989) more or less rigorously proved the following type of results for the case of commonly known stimuli: Survival of the fittest (in the sense of maximizing reproductive success or payoff) results if

(I.i) own success does not depend on other players' behavior

or

(I.ii) own stimuli do not influence other players' behavior.

Here it may suffice if these conditions are true only locally.[2] Furthermore, it has been argued that (I.ii) also covers the result of privately known stimuli that cannot be signaled at all.

Payoff maximization, e.g. in the sense of profit maximization on markets, represents the orthodox rationality assumption in economics. This has either been stated as a behavioral assumption, or justified by an as if-hypothesis in the sense that, in markets, only profit maximizing behavior can survive. This latter claim is rigorously analyzed when exploring the evolutionary stability of payoff or profit maximization. As already indicated above, this claim, i.e. the orthodox rationality assumption, can only be maintained under special circumstances, indicated by (I.i) and (I.ii).

---

[1] By this we rely on the usual fitness measure of evolutionary game theory (see, e.g., Weibull, 1995).

[2] Examples where (I.i) applies globally are competitive markets, and for (I.ii) games without signaling of stimuli which are private information. Some of the specific examples, listed above, rely on local conditions and (mostly) do not actually derive the results for privately known stimuli by explicitly solving Bayesian games.

Other more recent attempts, which focus on the polar case of complete type information and which mainly discuss whether individualistic (one cares only for the own material success) or interdependent preferences are evolutionarily stable (one cares also for the material success of others), are Kockesen, Ok, and Sethi (1997) and Ok and Vega-Redondo (1999). Any attempt to prove generally what has been previously learned only for specific examples will have to be limited to the class of games for which the claim can be shown to hold. In our case we rely on the necessary conditions of interior extrema under suitable differentiability assumptions. For easy exposition, we also proceed as if the solution (on the first step of the indirect evolutionary analysis) is unique.[3] This allows an unambiguous definition of the evolutionary game (on the second step of the indirect evolutionary analysis).

Actually the assumption of uniqueness is closely related to the one of differentiability. To define the evolutionary game unambiguously one needs in general only an association between the constellation of stimuli and an equilibrium of the game resulting from this constellation. One way to establish such an association is to apply the theory of equilibrium selection (e.g. Harsanyi and Selten, 1988). Once such an association is constructed, one can proceed as if this equilibrium would be the only one. If, however, the mapping from stimuli to solutions is constructed by selecting one of many equilibria, the selection rule will have to be discontinuous. This explains why uniqueness is required by our differentiability requirements. When dealing with situations where own stimuli are private information we, however, show that selection rules might be differentiable (Sect. 6).

As in typical evolutionary game theory (see again Weibull, 1995), the evolutionary games will be symmetric 2-person games although not necessarily with complete information. Evolutionarily stable strategies (ESS) for such games are symmetric Nash-equilibria satisfying an additional stability condition, namely that an ESS is better against an alternative best reply to the ESS than the alternative. Since we rely on differentiability, an alternative best reply will usually not exist. This explains why the additional stability requirement can usually be neglected. Our main motivation is not to prove the existence of an ESS, but rather to explore its implications whenever it exists.[4]

We first discuss in Section 2 the simplest case where both the commonly known stimuli and the strategies are one-dimensional. A simple example of pollution by production is used in Section 3 to illustrate our general findings. In Section 4 it is shown how the results can be generalized for multi-dimensional situations as well as for more refined solution concepts. Section 5 introduces a narrow class of Bayesian games for which symmetric Bayesian equilibria exist. Then, in Section 6 the evolutionarily stable distribution is analysed before, finally, Section 7 concludes with an extension of the analysis of the example to privately known stimuli.

---

[3] In economics, a continuous setting (see, for instance, the classic market models as reviewed by Tirole, 1988) usually implies unique equilibria.

[4] If no ESS exists one can rely on familiar coarsening ideas like neutral stability (Maynard Smith, 1982) or limit evolutionary stability (Selten, 1983 and 1988).

Quite often the results of indirect evolution resemble those of the literature on strategic commitment in interactive decision making, e.g. in agency relationships (see, for instance, van Damme and Hurkens, 1996, who also review some of this literature). There are, however, important differences between the two approaches: Strategic commitment relies on an overall game model, while in indirect evolution no decision maker has to be aware of the evolutionary forces, i.e. the modelling tasks are very different. As will be demonstrated below, indirect evolution allows us to distinguish between utility and (reproductive) success what in strategic delegation requires different agents, e.g. by modelling a firm as a team of a principal and an agent when analysing strategic commitment of competing firms (see Dufwenberg and Güth, 1999, who try to elaborate and illustrate in more detail the differences between indirect evolution and strategic commitment).

2 The one-dimensional case with commonly known stimuli

Let S, with $S \subset \mathbb{R}$, be a closed interval with a non-empty interior. An element s_i of S is a strategy or a form of behavior of player i. Similarly, let M with $M \subset \mathbb{R}$ be a closed interval whose non-empty interior contains 0. The elements $m_i \in M$ are the genetically determined stimuli which, together with the chosen strategies, determine the payoff $H_1(s_1, s_2; m_1, m_2)$ of the symmetric 2 person-game

(II.1) $G(m_1, m_2) = (S; H_1(s_1, s_2; m_1, m_2))$

for all $m_1, m_2 \in M$ where, of course, symmetry implies that player 2's payoff function is determined by

(II.2) $H_2(s_1, s_2; m_1, m_2) = H_1(s_2, s_1; m_2, m_1)$.

We assume that $H_1(s_1, s_2; m_1, m_2)$ is continuous in all its arguments and quasi-concave in s_1 for all $s_2 \in S$ and $m_1, m_2 \in M$. Furthermore, let the constellation $(m_1, m_2) \in M \times M$ be common knowledge. From these assumptions (see, for instance, van Damme, 1987) it follows

Remark 1. For all $m_1, m_2 \in M$ *game* $G(m_1, m_2)$ has at least one equilibrium

$$s^*(m_1, m_2) = \left(s_1^*(m_1, m_2), s_2^*(m_1, m_2)\right).$$

An obvious implication of the symmetry of $G(m_1, m_2)$ in the sense of (II.2) is

Lemma 2. *For all $m_1, m_2 \in M$ and all equilibria $s^* = \left(s_1^*, s_2^*\right)$ of $G(m_1, m_2)$ the strategy vector $\left(s_2^*, s_1^*\right)$ is an equilibrium of $G(m_2, m_1)$.*

What will be needed in the following is a mapping which associates one equilibrium $s^*(m_1, m_2) = \left(s_1^*(m_1, m_2), s_2^*(m_1, m_2)\right)$ with each game $G(m_1, m_2)$. If there exists only one equilibrium for all games $G(m_1, m_2)$, the mapping $s^*(m_1, m_2)$ is already determined. Otherwise it must be derived by applying some theory of equilibrium selection (e.g. Harsanyi and Selten, 1988) or some suitable criteria

(see Güth and Nitzan, 1997, for an example). If the mapping $s^*(m_1, m_2)$ is established, one can proceed as if $s^*(m_1, m_2)$ is the only equilibrium of $G(m_1, m_2)$. Such a mapping would, however, be discontinuous. Thus we could not rely on differentiability when solving the evolutionary game. We circumvent this problem by requiring

Uniqueness. *For all $m_1, m_2 \in M$ the game $G(m_1, m_2)$ has only one equilibrium $s^*(m_1, m_2) = \left(s_1^*(m_1, m_2), s_2^*(m_1, m_2)\right)$ to which we refer as the solution of $G(m_1, m_2)$.*

Uniqueness of $s^*(m_1, m_2)$ for all $m_1, m_2 \in M$ and upper hemicontinuity of the set of equilibria (Fudenberg and Tirole, 1991) implies

Remark 3. The solution strategies $s_i^*(m_1, m_2)$ are continuous in $m_1, m_2 \in M$.

Since the second step of indirect evolutionary analysis, namely the derivation of evolutionarily stable stimuli, will be based on differentiability, it is important to investigate the differentiability of the solution strategy $s_i^*(m_1, m_2)$. In order to do so we introduce the following Assumption.

Interiority. *For all $m_1, m_2 \in M$ the Nash equilibria of $G(m_1, m_2)$ are interior points of $S \times S$.*

Lemma 4. *Assume that in a neighborhood of the graph of $s^*(m_1, m_2)$ $H_1 = H_1(s_1, s_2; m_1, m_2)$ has the following properties:*

(i)   *$H_1(s_1, s_2; m_1, m_2)$ is concave in s_1 for every $s_2 \in S$ and $m_1, m_2 \in M$;*
(ii) *$H_1(s_1, s_2; m_1, m_2)$ is three times continuously differentiable in s_1, s_2, m_1 and m_2;*
(iii) *the Jacobian*

$$
\begin{vmatrix}
\dfrac{\partial^2}{\partial s_1^2} H_1 & \dfrac{\partial^2}{\partial s_1 \partial s_2} H_1 \\[3mm]
\dfrac{\partial^2}{\partial s_1 \partial s_2} H_2 & \dfrac{\partial^2}{\partial s_2^2} H_2
\end{vmatrix} \neq 0
$$

   for all $s_1, s_2 \in int(S)$ and $m_1, m_2 \in int(M)$.

Under the foregoing assumptions the (symmetric) strategies $\left(s_1^*(m_1, m_2), s_2^*(m_1, m_2)\right)$ are twice continuously differentiable in m_1 and m_2 (on $int(M)$).

Proof. Let $m_1, m_2 \in int(M)$. By Uniqueness, Interiority, (i), and (ii), the equilibrium $\left(s_1^*(m_1, m_2), s_2^*(m_1, m_2)\right)$ is the unique (interior) solution of the following equations:

$$
\begin{aligned}
\frac{\partial}{\partial s_1} H_1(s_1, s_2; m_1, m_2) &= 0 \\[2mm]
\frac{\partial}{\partial s_2} H_2(s_1, s_2; m_1, m_2) &= 0
\end{aligned}
$$

Now, by (ii), (iii), and the Implicit Functions Theorem, $s_i^*(m_1, m_2)$ for $i = 1, 2$ is twice continuously differentiable. $\square$

Instead of stating the assumptions rather globally, as in Lemma 4, one can prove also a local version of Lemma 4 which requires (i), (ii) and (iii) only for some neighborhood of the graph of the (vectorial) function $s^*(m_1, m_2)$. Another result is

Lemma 5. $s_1^*(m_1, m_2) = s_2^*(m_2, m_1)$ *for all* $m_1, m_2 \in M$.

Proof. Let $\left(s_1^*(m_1, m_2), s_2^*(m_1, m_2)\right)$ be the solution of $G(m_1, m_2)$. According to Lemma 2, the strategy vector $\left(s_2^*(m_1, m_2), s_1^*(m_1, m_2)\right)$ is an equilibrium of $G(m_2, m_1)$. Thus uniqueness of the solution for $G(m_2, m_1)$ implies $s_2^*(m_2, m_1) = s_1^*(m_1, m_2)$.   □

The first step of an indirect evolutionary analysis amounts to computing $s^*(m_1, m_2) = \left(s_1^*(m_1, m_2), s_2^*(m_1, m_2)\right)$ for all games $G(m_1, m_2)$ with $m_1, m_2 \in M$. Having completed this task, the second step starts with the definition of the evolutionary game

(II.3) $\Gamma = (M; M; R_1(m_1, m_2), R_2(m_1, m_2))$
(II.4) $R_i(m_1, m_2) = H_i(s^*(m_1, m_2); \lambda m_1, \lambda m_2)$ for $i = 1, 2$.

Here λ measures how stimuli are directly related to (reproductive) success $R_i(m_1, m_2)$ where we allow only for the two extreme relationships $\lambda = 0$, i.e. stimuli influence success only indirectly via $s^*(m_1, m_2)$, and $\lambda = 1$ where stimuli are directly related to (reproductive) success.[5]

In case of $\lambda = 1$, payoff in the games $G(m_1, m_2)$ and success in the evolutionary game Γ are the same. It may be surprising that $\lambda = 1$ does not usually imply the survival of the fittest: Equilibrium in games $G(m_1, m_2)$ requires a best response in the sense of

$$\frac{\partial}{\partial s_1} H_1\left(s_1^*(m_1, m_2), s_2^*(m_1, m_2); m_1, m_2\right) = 0$$

which does not guarantee

$$\frac{\partial}{\partial m_1} R_1(m_1, m_2) = 0$$

for $m_1 = 0 = m_2$ and $\lambda = 1$ (see equation (II.5') below). Actually this may render the case $\lambda = 1$ more rather than less difficult or even trivial. In applications, one often will want to allow for differences between motivation and success, i.e. one is interested in situations where $\lambda \neq 1$. In the following, the case $\lambda = 1$ will therefore be largely neglected.

For $\lambda = 1$, utility and (reproductive) success coincide; they differ in the case of $\lambda = 0$. Indirect evolution allows for both cases. One can, of course, also assume other values of λ, e.g. by considering $\lambda \in [0, 1]$ instead of $\lambda \in \{0, 1\}$.

---

[5] As Güth and Yaari (1992) most applications rely on $\lambda = 0$, i.e. assume that stimuli have no direct material implications. In our view, this is an interesting border case but a rather unlikely one. In most situations emotional states like anger, love etc. have direct material impact. It is, for instance, generally accepted that production of adrenaline is costly in terms of life expectation.

The case $\lambda = 0$ justifies speaking of **indirect** evolution: Since m-values do not determine success $R_i\,(m_1, m_2)$ directly, all their influence on success is indirect via the induced strategic behavior $s^*\,(m_1, m_2)$. Only through behavior can m_1 and m_2, respectively, determine success and thus the survival of behavioral dispositions captured by the m_i-values, $i = 1, 2$.

Thus the set M of possible, genetically determined stimuli is the strategy set of both players $i = 1, 2$. The reproductive success $R_i\,(m_1, m_2)$ measures how player i with stimuli m_i fares when encountering another player $j\,(\neq i)$ with stimuli m_j. One can describe Γ also by $\Gamma = (M; R_1\,(m_1, m_2))$ due to

Lemma 6. Γ *is symmetric, i.e.* $R_1\,(m_1, m_2) = R_2\,(m_2, m_1)$ *for all* $m_1, m_2 \in M$.

Proof.
$$
\begin{aligned}
R_1\,(m_1, m_2) \quad &= H_1\left(s_1^*\,(m_1, m_2)\,, s_2^*\,(m_1, m_2)\,; \lambda m_1, \lambda m_2\right) \\
&= H_2\left(s_2^*\,(m_1, m_2)\,, s_1^*\,(m_1, m_2)\,; \lambda m_2, \lambda m_1\right) \text{ due to II.2} \\
&= H_2\left(s_1^*\,(m_2, m_1)\,, s_2^*\,(m_2, m_1)\,; \lambda m_2, \lambda m_1\right) \text{ due to Lemma 5} \\
&= R_2\,(m_2, m_1) \qquad\qquad\qquad\qquad\qquad\qquad\qquad \square
\end{aligned}
$$

For the purpose of the study at hand it suffices to rely on static concepts of evolutionary stability such as the one of an evolutionarily stable strategy / ESS [see Maynard Smith and Price (1973) as well as Maynard Smith (1982), and Hammerstein and Selten (1994) and Weibull (1995) for surveys]. In our context an evolutionarily stable strategy / ESS $m^* \in M$ must satisfy

(i) $R_1\,(m^*, m^*) \geq R_1\,(m, m^*)$ for all $m \in M$

  and

(ii) $R\,(m^*, m) > R\,(m, m)$ for all $m \in M$ with $R\,(m^*, m^*) = R\,(m, m^*)$.

A necessary condition[6] for an interior $m^* \in M$ to be an evolutionarily stable strategy (ESS) of the evolutionary game $\Gamma = (M; R_1\,(m_1, m_2))$ is condition

(II.5) $\quad \dfrac{\partial}{\partial m_1} R_1(m^*, m^*) = 0$

where the differentiability of $R_1(\cdot)$ follows from the differentiability of $s_i^*\,(m_1, m_2)$ according to Lemma 4. If (II.5) does not hold, m^* does not qualify as a local maximum of $R_1\,(m, m^*)$ over all $m \in M$ so that an m^*-monomorphic population could be successfully invaded. Because of

---

[6] If the extremum is (locally) unique also the additional stability requirement of an (locally defined) ESS is fulfilled.

$$\frac{\partial}{\partial m_1} R_1 (m_1, m_2)$$

$$= \frac{\partial}{\partial m_1} H_1 \left(s^* (m_1, m_2) ; \lambda m_1, \lambda m_2 \right)$$

(II.6)
$$= \frac{\partial}{\partial s_1} H_1 \left(s_1^* (m_1, m_2), s_2^* (m_1, m_2) ; \lambda m_1, \lambda m_2 \right) \frac{\partial}{\partial m_1} s_1^* (m_1, m_2)$$

$$+ \frac{\partial}{\partial s_2} H_1 \left(s_1^* (m_1, m_2), s_2^* (m_1, m_2) ; \lambda m_1, \lambda m_2 \right) \frac{\partial}{\partial m_1} s_2^* (m_1, m_2)$$

$$+ \lambda \frac{\partial}{\partial m_1} H_1 \left(s^* (m_1, m_2) ; \lambda m_1, \lambda m_2 \right)$$

condition (II.5) is equivalent to

$$\frac{\partial}{\partial s_1} H_1 \left(s_1^* (m^*, m^*), s_2^* (m^*, m^*) ; \lambda m^*, \lambda m^* \right) \frac{\partial}{\partial m_1} s_1^* (m^*, m^*)$$

(II.5')
$$+ \frac{\partial}{\partial s_2} H_1 \left(s_1^* (m^*, m^*), s_2^* (m^*, m^*) ; \lambda m^*, \lambda m^* \right) \frac{\partial}{\partial m_1} s_2^* (m^*, m^*)$$

$$+ \lambda \frac{\partial}{\partial m_1} H_1 \left(s^* (m^*, m^*) ; \lambda m^*, \lambda m^* \right) = 0.$$

Let us call $m \in M$ critical if condition (II.5') or (II.5) holds.

Given $\lambda = 0$, according to the definition (II.4) of $R_1 (m_1, m_2)$, player 1's own parameter m_1 influences his fitness $R_1 (m_1, m_2)$ only indirectly via the solution $s^* (m_1, m_2)$ of the game $G (m_1, m_2)$ determined by m_1 and m_2. This justifies our interpretation that only in the case of an evolutionarily stable $m_1^* = 0$ does the survival of the fittest result: According to (II.4), if $m_1^* = 0$ maximizes $H_1 (s^* (m_1, m_2) ; \lambda m_1, \lambda m_2)$ this is equivalent to maximizing player 1's fitness $R_1 (m_1, m_2)$. In the following we want to explore the conditions for the survival of the fittest, i.e. for an ESS $m^* = 0$ based on $\lambda = 0$:

For $m^* = 0$ the equilibrium $s^* (m^*, m^*)$ of $G (m^*, m^*)$ implies

(II.7) $\frac{\partial}{\partial s_1} H_1 \left(s_1^* (0, 0), s_2^* (0, 0) ; 0, 0 \right) = 0$

so that (II.5') for $m^* = 0$ and $\lambda = 0$ simply means

(II.8) $\frac{\partial}{\partial s_2} H_1 \left(s_1^* (0, 0), s_2^* (0, 0) ; 0, 0 \right) \frac{\partial}{\partial m_1} s_2^* (0, 0) = 0.$

Since (II.8) is a necessary condition for $m^* = 0$ to be an equilibrium of Γ, it is also necessary for an ESS $m^* = 0$ of Γ. Thus for $\lambda = 0$ there are two requirements guaranteeing the necessary condition (II.8) for an ESS $m^* = 0$ of Γ, namely[7]

(II.9) $\frac{\partial}{\partial s_2} H_1 \left(s_1^* (0, 0), s_2^* (0, 0) ; 0, 0 \right) = 0$

or

---

[7] It would suffice to require these two conditions only locally if strategy variation and mutation can occur only locally, i.e. if only strategies s_i near to $s_i^* (0, 0)$ are feasible and in the case of $m^* = 0$ only mutants m close to $m^* = 0$ are to be expected, and if the derivatives vanish only generically (at one point) so that local uniqueness is guaranteed.

(II.10) $\quad \dfrac{\partial}{\partial m_1} s_2^* (0, 0) = 0.$

Equation (II.9) means that player 1's reproductive success does not depend on player 2's behavior. Typical examples of such situations are social, but non-strategic environments such as two isolated monopolists or, when leaving the narrow range of duopoly markets, competitive markets where a seller's success does not depend on the behavior of any individual coseller. Condition (II.9) thus confirms our initial claim (I.i) in a general framework.

Similarly, condition (II.10) justifies (I.ii). It says that the other player's behavior is - at least locally - not influenced by own stimuli. More specifically: If m_1 would change and if, according to our assumptions, player 2 would be aware of it, player 2's equilibrium strategy would remain constant.

In previous studies it has been argued (Bester and Güth, 1998; Güth and Huck, 1997; Güth, 1998; Possajennikov, 2000; Schaffer, 1989) that condition (II.10), albeit being a result for the case of known stimuli, already sheds light on situations where stimuli are private information. The argument is simply that privately known stimuli should guarantee that own stimuli cannot influence other players' behavior. If only player 1 is aware of m_1, a change of m_1 cannot induce a change of s_2, so that condition (II.10) applies. In our view, such conclusions should be substantiated by an explicit analysis of games in which stimuli are private information.[8] The conjecture will be rigorously proved for a narrow but reasonable class of beliefs concerning the other players' stimuli.

For the case of commonly known stimuli and $\lambda = 0$ our results can be summarized by

Theorem 7. *Let $\Gamma = (M; R_1(m_1, m_2))$ be the evolutionary game defined by (II.3) and (II.4) with the help of the solutions $s^*(m_1, m_2)$ of games $G(m_1, m_2)$ with $m_1, m_2 \in M$. For the survival of the fittest, i.e. for $m^* = 0 \in M$ in case of $\lambda = 0$ to be evolutionarily stable, it is necessary that equation (II.9) holds, confirming (I.i), or that condition (II.10) is true what justifies (I.ii).*

Up to now we only investigated the necessary condition for an interior ESS $m^* \in M$. According to our differentiability assumptions, $\frac{\partial^2}{\partial m_1^2} R_1(m_1, m_2)$ exists (actually requirement (ii) of Lemma 4 is only needed for this). If for an interior $m^* \in M$ one also would have

(II.11) $\quad \dfrac{\partial^2}{\partial m_1^2} R_1\left(m_1, m^*\right) < 0$ for all $m_1 \in M,$

the only best reply to m^* in Γ would be $m_1 = m^*$. Thus (II.5) and (II.11) for $m^* \in M$ are sufficient to prove that $m^* \in M$ is an ESS of Γ (see Hammerstein and Selten, 1994). Whether (II.11) holds for m^* depends, of course, on the mathematical structure of the model under consideration.

---

[8] Privately known stimuli are just another extreme possibility [for an indirect evolutionary analysis where stimuli can be more or less reliably signaled see Güth, Güth and Kliemt (forthcoming); Güth and Kliemt (1994, 2000); and Güth, Kliemt, and Peleg (2000)]

3 An example

To illustrate our general results we consider a simple heterogeneous market with complementary products and individual demand functions

(III.1) $x_i\,(s_1, s_2) = 1 - s_i - \alpha s_j$ for $i, j = 1, 2$ and $i \neq j$.

Here $x_i\,(s_1, s_2)$ denotes seller i's sales amount and s_1, s_2 their respective sales prices. The parameter $\alpha\,(\geq 0)$ describes how closely sellers 1 and 2 are interrelated, i.e. for $\alpha = 0$ condition (II.9), substantiating our claim (I.i), should hold.

For the payoffs we assume

(III.2) $H_1\,(s_1, s_2; m_1, m_2) = s_1 x_1\,(s_1, s_2) - m_1 x_1\,(s_1, s_2)$.

In the case of $m_1 > 0$, one can interpret the term $m_1 x_1\,(s_1, s_2)$ as expressing player 1's concern about environmental damage caused by his production activities. Similarly, $m_1 = 0$ stands for no environmental concern and $m_1 < 0$ for enjoying polluting. For $\lambda = 0$ (reproductive) success on the market depends only on the actual profit $s_1 x_1\,(s_1, s_2)$. This justifies characterization of an evolutionarily stable $m^* = 0$ as the survival of the fittest. We thus have $S \subset \mathbb{R}$ and $M \subset \mathbb{R}$ with $0 \in M$. By assuming that both, S and M, are closed intervals with non-empty interiors, the other assumptions of the previous section are satisfied, too. Later we will specify further conditions to guarantee the economic non-negativity constraints.

Disregarding boundary solutions (which will anyhow be eliminated by our more specific restrictions for S, M, and α) one obtains the solution $s^*\,(m_1, m_2) = \left(s_1^*\,(m_1, m_2), s_2^*\,(m_1, m_2)\right)$ given by

(III.3) $s_1^* = s_1^*\,(m_1, m_2) = \dfrac{2 - \alpha + 2m_1 - \alpha m_2}{4 - \alpha^2}$

(III.4) $s_2^* = s_2^*\,(m_1, m_2) = \dfrac{2 - \alpha + 2m_2 - \alpha m_1}{4 - \alpha^2}$.

Since

(III.5)
$$\begin{aligned} x_1\left(s_1^*, s_2^*\right) &= \tfrac{4 - \alpha^2 - (1+\alpha)(2-\alpha) - 2m_1 + \alpha m_2 - 2\alpha m_2 + \alpha^2 m_1}{4 - \alpha^2} \\ &= \frac{2 - \alpha - \left(2 - \alpha^2\right) m_1 - \alpha m_2}{4 - \alpha^2}, \end{aligned}$$

reproductive success is given by

(III.6) $R_1\,(m_1, m_2) = \dfrac{\left(2 - \alpha + 2m_1 - \alpha m_2\right)\left(2 - \alpha - \left(2 - \alpha^2\right) m_1 - \alpha m_2\right)}{\left(4 - \alpha^2\right)^2}$

$$-\lambda m_1 \cdot \frac{2 - \alpha - \left(2 - \alpha^2\right) \cdot m_1 - \alpha \cdot m_2}{4 - \alpha^2},$$

with $\lambda \in \{0, 1\}$. To guarantee non-negativity of prices and sales amounts we assume

(III.7) $0 \leq \alpha < \frac{1}{2}$

(III.8) $M = \left[-\frac{1}{3}, \frac{1}{3}\right]$,

and

(III.9) $S = \left[0, \frac{2}{3}\right]$,

what completes the description of the example.

Now, we want to derive the necessary conditions (II.9) or (II.10) for the evolutionary stability of $m^* = 0 \in M$. Since for $\lambda = 0$ one has

(III.10) $\quad \dfrac{\partial}{\partial s_2} H_1 \left(s^*(0,0); 0, 0\right) = s_1^*(0,0) \dfrac{\partial}{\partial s_2} x_1 \left(s^*(0,0)\right)$

$$= -\alpha \frac{2-\alpha}{4-\alpha^2} = -\frac{\alpha}{2+\alpha} \; ,$$

condition (II.9) requires

(III.11) $\alpha = 0$,

i.e. that the two sellers are actually monopolists serving two isolated markets, as required in our initial claim (I.i). Since

(III.12) $\quad \dfrac{\partial}{\partial m_1} s_2^* (m_1, m_2) = \dfrac{-\alpha}{4-\alpha^2}$

for all $m_1, m_2 \in M$, condition (II.10) also implies (III.11). Thus for $m^* = 0 \in M$ to be evolutionarily stable, one needs $\alpha = 0$; i.e. with commonly known stimuli and $\lambda = 0$, the survival of the fittest can only be expected when the two sellers are completely unrelated. This means that two related sellers, i.e. with $\alpha > 0$, will not in general neglect the environmental aspects of their production activities, i.e. the evolutionarily stable $m^* \in M$ will usually satisfy $m^* \neq 0$. The sign of m^* will determine whether this will actually reduce pollution, which according to (III.1) and (III.3) or (III.4) requires $m^* > 0$.

4 Generalizing the case of commonly known stimuli

When $S \subset E^p$ and $M \subset \mathbb{R}^q$ such that S has a non-empty interior and is convex and compact, and M is convex and compact with $0 \in \text{int}(M)$, and when all our other assumptions are generalized to situations with $p > 1$ or $q > 1$ accordingly, the generalization of our results in Section 2 is straightforward. For the sake of simplicity this will only be done for the case $\lambda = 0$ when stimuli influence success only indirectly. Let us, as before, denote by $s_i = \left(s_i^1, ..., s_i^p\right)$ and $m_i = \left(m_i^1, ..., m_i^q\right)$ an arbitrary element of S and M, respectively. Instead of (II.6) in the multi-dimensional framework one has

$$
\begin{aligned}
\text{(IV.1)} \quad \nabla_1 R_1 (m_1, m_2) &= \left(\frac{\partial}{\partial m_1^1} R_1 (m_1, m_2), \dots, \frac{\partial}{\partial m_1^q} R_1 (m_1, m_2) \right) \\
&= \nabla_1 H_1 \left(s^* (m_1, m_2); 0, 0 \right) * \Box_1 s_1^* (m_1, m_2) \\
&\quad + \nabla_2 H_1 \left(s^* (m_1, m_2); 0, 0 \right) * \Box_1 s_2^* (m_1, m_2)
\end{aligned}
$$

where
$$
\nabla_1 H_1 \left(s^* (m_1, m_2); 0, 0 \right)
$$
$$
= \left(\frac{\partial}{\partial s_1^1} H_1 \left(s^* (m_1, m_2); 0, 0 \right), \dots, \frac{\partial}{\partial s_1^p} H_1 \left(s^* (m_1, m_2); 0, 0 \right) \right)
$$
$$
\nabla_2 H_1 \left(s^* (m_1, m_2); 0, 0 \right)
$$
$$
= \left(\frac{\partial}{\partial s_2^1} H_1 \left(s^* (m_1, m_2); 0, 0 \right), \dots, \frac{\partial}{\partial s_2^p} H_1 \left(s^* (m_1, m_2); 0, 0 \right) \right)
$$
$$
\Box_1 s_1^* (m_1, m_2) = \left(\frac{\partial}{\partial m_1^i} \left(s_1^j \right)^* (m_1, m_2) \right) \quad j = 1, \dots, p; i = 1, \dots, q
$$
$$
\Box_1 s_2^* (m_1, m_2) = \left(\frac{\partial}{\partial m_1^i} \left(s_2^j \right)^* (m_1, m_2) \right) \quad j = 1, \dots, p; i = 1, \dots, q
$$

and "$*$" stands for multiplying a $p \times q$ matrix with a p-vector. With the help of this notation the equivalent condition for (II.5') can be written as

$$
\begin{aligned}
\text{(IV.2)} \quad & \nabla_1 H_1 \left(s^* (m_1, m_2); 0, 0 \right) * \Box_1 s_1^* (m_1, m_2) \\
& + \nabla_2 H_1 \left(s^* (m_1, m_2); 0, 0 \right) * \Box_1 s_2^* (m_1, m_2) = 0.
\end{aligned}
$$

As for the case of $q = 1$ evolutionary stability of stimuli $m^* = 0 \in M$ can be interpreted as the survival of the fittest. Since in the equilibrium $s^* (0, 0)$ of $G (0, 0)$ the condition

$$
\text{(IV.3)} \quad \nabla_1 H_1 (s^* (0, 0); 0, 0) = 0
$$

holds, condition (IV.2), when $m^* = 0$, assumes the simpler form

$$
\text{(IV.4)} \quad \nabla_2 H_1 (s^* (0, 0); 0, 0) * \Box_1 s_2^* (0, 0) = 0
$$

corresponding to equation (II.8) in the one-dimensional situation. Technically (IV.4) is the product of a p-vector and $p \times q$-matrix and thus a q-vector whose components all have to vanish. In general, there are many ways in which

$$
\sum_{j=1}^{p} \frac{\partial}{\partial s_2^j} H_1 \left(s^* (m_1, m_2); 0, 0 \right) \frac{\partial}{\partial m_1^i} \left(s_2^j \right)^* (m_1, m_2) = 0
$$

can be guaranteed for $i = 1, \dots, q$. A sufficient but not necessary condition for (IV.4) is

$$
\text{(IV.5)} \quad \nabla_2 H_1 (s^* (0, 0); 0, 0) = 0 \text{ or } \Box_1 s_2^* (0, 0) = 0.
$$

As (II.9) condition (IV.5) justifies the claim (I.i), respectively (I.ii), namely, that the survival of the fittest ($m^* = 0 \in M$) can be expected if a player's payoff does not depend on any of the other's choices, respectively if a player's behavior (equilibrium choice) does not depend on any component of the other player's stimuli.

The strategy set S may be the set of mixed strategies for a finite strategic game $\left(\widehat{S}; H_1\left(\cdot, \cdot; m_1, m_2\right)\right)$, i.e. every $s_i \in S$ is a probability distribution over the finite set $\widehat{S}$. One can then refer to refinements such as perfect (Selten, 1975) or proper (Myerson, 1978) equilibria which are defined for finite games. Of course, the other assumptions concerning $H_1\left(\cdot, \cdot; m_1, m_2\right)$ have to be satisfied, too, but they do not contradict the interpretation that the games $G\left(m_1, m_2\right)$ are the mixed extensions of some finite games.

If one applies our uniqueness assumption to more refined equilibrium notions, provided they are well-defined and exist, all our analysis remains valid. A careful examination even reveals that the uniqueness assumption may be relaxed in the following way:

There exists a single-valued selection $\varphi : M \times M \to S \times S$ associating an equilibrium solution $\varphi\left(m_1, m_2\right) = \left(\varphi_1\left(m_1, m_2\right), \varphi_2\left(m_1, m_2\right)\right)$ with each game $G\left(m_1, m_2\right)$ which is twice continuously differentiable and satisfies the symmetry requirement $\left(\varphi_1\left(m_1, m_2\right), \varphi_2\left(m_1, m_2\right)\right) = \left(\varphi_2\left(m_2, m_1\right), \varphi_1\left(m_2, m_1\right)\right)$.

In spite of the general problems with continuous equilibrium selection (see Harsanyi and Selten, 1988) the differentiability assumption may appear reasonable when considering certain families of games $G\left(m_1, m_2\right)$. Of course, $\varphi\left(\cdot\right)$ may also select more refined equilibria when these are defined. The game $G\left(m_1, m_2\right)$ may also be seen as the normal form of an extensive game $g\left(m_1, m_2\right)$. In such a case the selection $\varphi\left(m_1, m_2\right)$ may depend on the original extensive game $g\left(m_1, m_2\right)$, e.g. $\varphi\left(m_1, m_2\right)$ may select the unique subgame perfect equilibrium (Selten, 1975) of $g\left(m_1, m_2\right)$. However, one has to assume that all the properties of the selection $\varphi\left(\cdot\right)$ are satisfied.

When studying sequential games a typical problem is the non-existence of evolutionarily stable strategies. In such games, certain stimuli might guide the behavior in proper subgames which, however, may not be reached at all (see Güth and Kliemt, 2000, for a simple example). Such phenomena typically imply that no evolutionarily stable strategy exists since evolutionary forces cannot drive the evolution of stimuli in unreached information sets.

For the purpose of the study at hand, this problem, however, causes no harm since non-existence of ESS typically results from the non-uniqueness of best replies to a supposedly stable $m \in M$. We are not so much concerned with conditions guaranteeing the existence of an ESS or of a coarsening of an ESS (see, for instance, Selten, 1988), but with the necessary condition for $m^* = 0$ to be evolutionarily stable in the sense of (II.8) or (IV.4). Any coarsening of the ESS-concept will also have to satisfy these conditions so that our results will be true regardless of whether the ESS-concept or one of its coarsenings has to be applied.

5 Bayesian equilibria when stimuli are private information

As for the case of commonly known stimuli in Section 2, we focus on one dimensional actions and stimuli, i.e. $S \subset \mathbb{R}$ and $M \subset \mathbb{R}$. More specifically, we consider the class of symmetric Bayesian games

(V.1) $G = \left(S, M; H_1 \left(s_1, s_2; m_1, m_2 \right); p \left(\cdot \mid m \right) \right)$

where $S, M \subset \mathbb{R}$ are closed intervals with non-empty interiors and 0 is an interior value of M. The payoff function $H_1 \left(s_1, s_2; m_1, m_2 \right)$ is assumed to be three times continuously differentiable in all its arguments and satisfies

(V.2) $H_1 \left(s_1, s_2; m_1, m_2 \right) = H_2 \left(s_2, s_1; m_2, m_1 \right)$

for all $s_1, s_2 \in S$ and $m_1, m_2 \in M$. The probability $p_i \left(\widehat{m} \mid \widetilde{m} \right)$ measures how likely it is for player i's type $\widetilde{m}$ that player $j \left(\neq i \right)$ is of type $\widehat{m}$. Symmetry of beliefs concerning other's stimuli requires

(V.3) $p_1 \left(\widehat{m} \mid \widetilde{m} \right) = p \left(\widehat{m} \mid \widetilde{m} \right) = p_2 \left(\widehat{m} \mid \widetilde{m} \right)$

for $\widehat{m}, \widetilde{m} \in M$.

   A symmetric Bayesian equilibrium of game G is a Borel-measurable function $s^* \left(\cdot \right) = s_1^* \left(\cdot \right) = s_2^* \left(\cdot \right)$ from M to S such that

(V.4)
$$\int_M H_1 \left(s^* \left(m_1 \right), s^* \left(m_2 \right); m_1, m_2 \right) dp \left(m_2 \mid m_1 \right)$$
$$\geq \int_M H_1 \left(s_1, s^* \left(m_2 \right); m_1, m_2 \right) dp \left(m_2 \mid m_1 \right)$$

holds for all $s_1 \in S$ and all $m_1 \in M$. When proving the existence of symmetric Bayesian equilibria for games G we rely on type independent or free beliefs $p \left(\cdot \mid m \right)$ satisfying

(V.5) $p \left(\widehat{m} \mid \widetilde{m} \right) = p \left(\widehat{m} \right)$ for all $\widehat{m}, \widetilde{m} \in M$.

   Further restrictions are that there exists δ such that

(V.6) $\dfrac{\partial^2}{\partial s_1^2} H_1 \left(s_1, s_2; m_1, m_2 \right) \leq \delta / |M| < 0$ for all $s_1, s_2 \in S; m_1, m_2 \in M$,

where $|M|$ is the length of M. From the continuity of $\frac{\partial^2}{\partial s_1 \partial m_1} H_1 \left(\cdot \right)$ follows that there exists a positive constant K with

(V.7) $\max \left\{ \left| \dfrac{\partial^2}{\partial s_1 \partial m_1} H_1 \left(s_1, s_2; m_1, m_2 \right) \right| / |\delta| : s_1, s_2 \in S; m_1, m_2 \in M \right\}$

$\leq K / |M|$.

If one only allows for strategies $s \left(\cdot \right) : M \to S$ such that

(V.8) $| s (\widehat{m}) - s (\widetilde{m}) | \leq K | \widehat{m} - \widetilde{m} |$ for all $\widehat{m}, \widetilde{m} \in M$,

then the set $\mathfrak{S}$ of possible strategies $s (\cdot)$ is convex, uniformly bounded and uniformly continuous. According to Ascoli's Theorem (see, for instance, Arrow and Intriligator, 1981) the set $\mathfrak{S}$ is therefore a compact subset of the set of continuous strategies $s (\cdot)$ in game G.

Theorem 8. *Given the assumptions of games G there exists a symmetric Bayesian equilibrium for all games G if all best replies are interior values of S.*

Proof. (see Appendix).   □

Another existence proof of Bayesian equilibria for infinite games with incomplete information can be found in Milgrom and Weber (1985). Here we do not prove mere existence, but the existence of **differentiable** Bayesian equilibria. To the best of our knowledge our result is not implied by the existing literature.

6 On the survival of the fittest
when stimuli are private information

Let π be the true population distribution over M. As for games with commonly known stimuli we define the evolutionary game by anticipating the solution $s^* (\cdot) : M \to S$, i.e. the Bayesian equilibrium derived above which does not have to be unique. More specifically, let

(VI.1) $G (\pi) = (S, M; H_1 (s_1, s_2; m_1, m_2); \pi (\cdot))$

be the Bayesian game with $p (\cdot) \equiv \pi (\cdot)$, i.e. beliefs are determined by the true population distribution. Let $(\pi; s^*)$ denote a population density π defining the symmetric Bayesian game $G (\pi)$ of which s^* is a symmetric Bayesian equilibrium. With the help of this notation the evolutionary Bayesian game depends on the true population density $\pi (\cdot)$ over M as follows:

(VI.2) $\Gamma (\pi; s^*) = \left(M; M; R_1 (\widehat{m}, \widetilde{m}; m_1, m_2); \pi \right)$

for all distributions $\pi (\cdot)$ over M, with

(VI.3) $R_1 (\widehat{m}, \widetilde{m}; m_1, m_2) = H_1 (s^* (\widehat{m} \mid \pi), s^* (\widetilde{m} \mid \pi); \lambda m_1, \lambda m_2)$

The notation $s^* (m \mid \pi)$ is supposed to indicate that the equilibrium depends on the population distribution π (it is not meant to indicate a conditional probability). In the evolutionary game $\Gamma (\pi; s^*)$ the set M of possible stimuli serves both as an action space (players announce stimuli $\widehat{m}$ and $\widetilde{m}$, respectively) and as a type space (players are of certain stimuli types m_1 and m_2, respectively). The true population density $\pi (\cdot)$ determines a player's beliefs concerning the other player's stimuli. As before, λ can assume two values, $\lambda = 0$ and $\lambda = 1$: While in case of $\lambda = 0$ stimuli influence (reproductive) success only indirectly via the solution behavior $s_i^* (m_i \mid \pi)$, in case of $\lambda = 1$ utility equals (reproductive) success.

In direct evolutionary game theory, e.g. for the hawk and dove-game, the carrier of an ESS may contain several strategies. In a Bayesian setting this corresponds to distributions $\pi(\cdot)$ over M which are not just one point- or Dirac-measures. Such a distribution $\pi(\cdot)$ would be stable[9] if all stimuli types are matched onto themselves. To allow for polymorphic distributions as stable results of evolutionary processes[10] we rely on

Definition 9. *A distribution $\pi(\cdot)$ over M is evolutionarily stable if* $id: M \to M$ *with* $id(m) = m$ *for all* $m \in M$ *is an evolutionarily stable strategy of some* $\Gamma(\pi; s^*)$ *where* s^* *is a symmetric Bayesian equilibrium of* $G(\pi)$.

For $\lambda = 1$, where utility maximization corresponds to maximizing fitness, we obtain

Proposition 10. *If* $\lambda = 1$, *then* id *satisfies the first order condition for an evolutionarily stable strategy of each* $\Gamma(\pi; s^*)$.

When interpreting Proposition 10, one should recall that the evolutionary game $\Gamma(\pi; s^*)$ with action space M for both players is a direct mechanism and that the first order condition for an evolutionarily stable strategy $s^*(\cdot) = id$ means that the mechanism $\Gamma(\pi; s^*)$ is incentive compatible. Thus Proposition 10 is a special case of the revelation principle (Myerson, 1979) so that a proof is not needed. In case of $\lambda = 1$ different m-types will rely on different announcements and thus on different actions since their success functions differ. What Proposition 10 thus says is that every type will finally behave as maximizing his own fitness. In other words: The distribution of actions must resemble the distribution of types. Any distribution over m-types with this property is evolutionarily stable according to Definition 9.

For $\lambda = 0$ the necessary condition is

$$(\text{VI.4}) \quad \int_M \frac{\partial}{\partial s_1} H_1\left(s^*\left(m_1 \mid \pi\right), s^*\left(m_2 \mid \pi\right); 0, 0\right) \frac{d}{dm_1} s^*\left(m_1 \mid \pi\right) d\pi\left(m_2\right) = 0$$

for all $m_1 \in M$. So we can use the one point solution $m_1^* = 0 \in M$, which confirms our initial claim (I.ii) that in the case of privately known stimuli, in the sense defined above, only the fittest will survive, i.e. for $\lambda = 0$ only the stimuli $m^* = 0$. Our results are summarized by

Theorem 11. *Assume that stimuli are private information and cannot be signaled. If* $\lambda = 1$ *then the (first order) necessary condition for the survival of the fittest is always satisfied (see Proposition 10). If* $\lambda = 0$ *then (VI.4) is this necessary condition.*

---

[9] This requirement, of course, captures only the symmetric Nash-equilibrium property and no additional stability requirements like an ESS whose mathematical analogue would raise serious technicalities.

[10] For examples of stable polymorphisms see again Güth, Güth, and Kliemt (forthcoming); Güth and Kliemt (1994, 2000), and Güth, Kliemt, and Peleg (2000), which are all based on the game of trust.

7 The example with stimuli being private information

The necessary condition for maximizing

$$(\text{VII.1}) \quad \int_M (s_1 - \lambda m_1) x_1 (s_1, s_2 (m_2)) \, d\pi (m_2)$$

is

$$(\text{VII.2}) \quad \int_M x_1 (s_1, s_2 (m_2)) \, d\pi (m_2) = \int_M (\lambda m_1 - s_1) \tfrac{\partial}{\partial s_1} x_1 (s_1, s_2 (m_2)) \, d\pi (m_2).$$

Thus a symmetric Bayesian equilibrium requires

$$(\text{VII.3}) \quad \int_M \left(1 - s^* (m) - \alpha s^* (\widetilde{m})\right) d\pi (\widetilde{m}) = s^* (m) - \lambda m$$

what implies $s^* (m) = \tfrac{\lambda}{2} m + \gamma$ for some constant γ and thus

$$(\text{VII.4}) \quad s^* (m) = \frac{\lambda}{2} m + \frac{1}{2 + \alpha} - \frac{\lambda}{2 (2 + \alpha)} \mu_\pi$$

where μ_π is the mean value of m with respect to the distribution $\pi (\cdot)$. According to Proposition 10, any distribution $\pi (\cdot)$ over M is evolutionarily stable for $\lambda = 1$ so that no condition has to be imposed. For $\lambda = 0$ equation (VII.4) implies that

$$(\text{VII.5}) \quad \frac{\partial}{\partial m} s^* (m) = 0.$$

Thus a change of one's own stimuli neither influences the other's nor the own behavior. For $\lambda = 0$ only $m^* = 0$ can be evolutionarily stable, since maximizing $R_1 (\cdot)$ is only equivalent to maximizing $H_1 (\cdot)$ if $m = 0$.

8 Conclusions

According to traditional evolutionary analysis (see Hammerstein and Selten, 1994), behavior evolves in such a way that it is optimally adjusted to the population behavior. This has usually been described as the survival of the fittest. This follows from the definition of evolutionary stability, e.g. in the sense of evolutionarily stable strategies which are best replies to themselves.

Indirect evolution studies the way in which stimuli evolve which, in turn, determine behavior. The manner in which behavior depends on stimuli has been derived by applying game theory. Inserting this dependency yields an evolutionary game with stimuli as strategies to which one, as in traditional analysis, can apply concepts of evolutionary stability. We then have asked whether the behavior, implied by the evolutionarily stable stimuli, is optimally adjusted to the population behavior as in traditional evolutionary analysis.

For all situations satisfying our - admittedly - strong differentiability requirements, the results confirm the intuition suggested by previous applications: If stimuli are common knowledge, the fittest behavior only survives when own success does not depend on other's behavior or when other's behavior does not react to own stimuli. In the case of privately known stimuli, results depend on whether stimuli are directly ($\lambda = 1$) or only indirectly ($\lambda = 0$) related to success as described by Theorem 11.

We have mostly (see the discussion of condition (II.11) for an exception) confined ourselves to investigating the (first order) necessary conditions of evolutionary stability. When asking for existence or sufficient conditions for evolutionary stability one should recall that evolutionarily stable strategies might not exist. Actually this inspired some coarsenings of evolutionary stability such as neutral evolutionary stability (condition (ii) of an ESS is a weak inequality instead of a strict one, see Maynard Smith, 1982) or limit evolutionary stability (the ESS-conditions can be justified by imposing arbitrary, but small trembles, see Selten, 1988). As previous examples, especially of sequential games (e.g. Güth and Kliemt, 1998), have shown these ideas usually allow derivation of evolutionarily stable configurations. A general proof of this would, however, require a thorough mathematical investigation which would overburden our analysis.

Let us discuss also why we have relied on rather strong assumptions of differentiability. To guarantee existence of (interior) equilibria one usually imposes either convexity (see van Damme, 1987) or differentiability assumptions as in the present study. Recall, however, that indirect evolution requires two steps of equilibrium analysis - one first derives the equilibrium for all type constellations and then the equilibrium constellations of types or behavioral dispositions. While convexity, imposed for the first step, would not be inherited by the evolutionary game, strong enough differentiability assumptions can be inherited in the form of weaker ones as long as one does not rely on equilibrium selection. This explains why in the first part of our study we have relied on admittedly strong differentiability assumptions and also on the uniqueness of equilibria.

Appendix: Proof of Theorem 8

We first construct a best reply mapping $\beta(\cdot): \mathfrak{S} \to \mathfrak{S}$ which is continuous with respect to the maximum norm for continuous strategies $s(\cdot): M \to S$ and then prove that the requirements of the Brouwer-Schauder-Tychonoff-Theorem are satisfied.

(i)   We want to show that for any $s(\cdot) \in \mathfrak{S}$ the best reply $\beta(s)$ is also contained in $\mathfrak{S}$. For all $\widehat{s} \in S$ and $m \in M$ define the expected payoff of the stimuli type m for strategy $\widehat{s}$ by

$$(V.9) \quad F(\widehat{s}, m) = \int_M H_1(\widehat{s}, s(m_2); m, m_2) \, dp(m_2).$$

Because of our assumptions F is three times continuously differentiable in both its arguments. Furthermore,

$$(\text{V.10}) \quad \frac{\partial^2}{\partial s_1^2} F\left(\widehat{s}, m\right) = \int_M \frac{\partial^2}{\partial s_1^2} H_1\left(\widehat{s}, s\left(m_2\right); m, m_2\right) dp\left(m_2\right) \leq \delta < 0,$$

so the best response $\beta\left(s\left(\cdot\right)\right)$ to $s\left(\cdot\right)$ is implicitly given by

$$(\text{V.11}) \quad \frac{\partial}{\partial s_1} F\left(\beta\left(m\right), m\right) = 0 \text{ for all } m \in M$$

due to our assumption that all best replies $\beta\left(m\right)$ are interior values of S. By the Implicit Functions Theorem one has

$$(\text{V.12}) \quad \frac{d}{dm} \beta\left(m\right) = -\frac{\frac{\partial^2}{\partial s_1 \partial m} F\left(\beta\left(m\right), m\right)}{\frac{\partial^2}{\partial s_1^2} F\left(\beta\left(m\right), m\right)}.$$

Because of $\left| \frac{\partial^2}{\partial s_1^2} F\left(\beta\left(m\right), m\right) \right| \geq \left| \delta \right|$ one obtains

$$(\text{V.13}) \quad \left| \frac{d}{dm} \beta\left(m\right) \right| \leq \frac{1}{\left| \delta \right|} \left| \frac{\partial^2}{\partial s_1 \partial m} F\left(\beta\left(m\right), m\right) \right| \leq K.$$

Hence $\beta\left(m\right)$ satisfies condition (V.8), i.e. $\beta\left(m\right) \in \mathfrak{S}$. A corollary of part (i) of the proof is that a symmetric Bayesian equilibrium $s\left(\cdot\right) \in \mathfrak{S}$ of G is twice continuously differentiable.

(ii)   Let us rewrite the function $F\left(\widehat{s}, m\right)$, defined in (V.9), as

$$(\text{V.14}) \quad F\left(\beta\left(m\right), m; s\left(\cdot\right)\right) = \int_M H_1\left(\beta\left(m\right), s\left(m_2\right); m, m_2\right) dp\left(m_2\right)$$

where, as before, $\beta\left(m\right)$ is the best reply to $s\left(\cdot\right)$ for the m-type of player 1. For any continuous strategy $s\left(\cdot\right) : M \to S$ let

$$(\text{V.15}) \quad \left\| s\left(\cdot\right) \right\| = \max\left\{ \left| s\left(m\right) \right| : m \in M \right\}$$

be the (maximum) norm of $s\left(\cdot\right)$. Let $M\left(\widehat{s}, \widetilde{s}\right) = \left\{ m \in M : \widehat{s}\left(m\right) \neq \widetilde{s}\left(m\right) \right\}$. Since

$$\left| F\left(\beta\left(m\right), m; \widehat{s}\left(\cdot\right)\right) - F\left(\beta\left(m\right), m; \widetilde{s}\left(\cdot\right)\right) \right|$$
$$= \left| \int_{M\left(\widehat{s}, \widetilde{s}\right)} \frac{H_1\left(\beta\left(m\right), \widehat{s}\left(m_2\right); m, m_2\right) - H_1\left(\beta\left(m\right), \widetilde{s}\left(m_2\right); m, m_2\right)}{\widetilde{s}\left(m_2\right) - \widehat{s}\left(m_2\right)} \left(\widetilde{s}\left(m_2\right) - \widehat{s}\left(m_2\right)\right) dp\left(m_2\right) \right|$$

$$\leq \max\left\{ \left| \frac{\partial}{\partial s_2} H_1\left(\beta\left(m\right), s_2; m, m_2\right) \right| : \beta\left(m\right), s_2 \in S, m, m_2 \in M \right\} \left\| \widehat{s}\left(\cdot\right) - \widetilde{s}\left(\cdot\right) \right\|$$

because of the Mean Value Theorem, the function $F\left(\beta\left(m\right), m; s\left(\cdot\right)\right)$, defined in (V.14), is Lipschitz-continuous in $s\left(\cdot\right) \in \mathfrak{S}$. Similarly, all its first and second derivatives are Lipschitz-continuous in $s\left(\cdot\right) \in \mathfrak{S}$. Because of (V.10) and (V.12) also $\beta\left(\cdot\right)$, i.e. the best reply function $\beta\left(\cdot\right) : M \to S$ against $s\left(\cdot\right)$, is Lipschitz-continuous in $s\left(\cdot\right)$ and therefore continuous on $\mathfrak{S}$. Thus the existence of a symmetric Bayesian equilibrium $s\left(\cdot\right)$ with $\beta\left(m\right) = s\left(m\right)$ for all $m \in M$ follows from

the Brouwer-Schauder-Tychonoff-Theorem (see, for instance, Aliprantis and Border, 1994).

References

Aliprantis CD, Border KC (1994) Infinite dimensional analysis. Springer, Berlin Heidelberg New York

Arrow KJ, Intriligator MD (1981) Handbook of mathematical economics, Vol I. North-Holland, Amsterdam

Bester H, Güth W (1998) Is altruism evolutionarily stable?, Journal of Economic Behavior and Organization 34(2): 193–209

Dufwenberg M, Güth W (1999) Indirect evolution versus strategic delegation: A comparison of two approaches to explaining economic institutions. European Journal of Political Economy 15: 281–295

Fudenberg D, Tirole J (1991) Game theory. MIT Press, Cambridge, MA

Güth W (1998) Are rational cost expectations evolutionarily stable? ifo-Studien – Zeitschrift für empirische Wirtschaftsforschung 1: 13

Güth W, Güth S, Kliemt H (2002) The dynamics of trustworthiness among the few. Japanese Economic Review (forthcoming)

Güth W, Huck S (1997) A new justification of monopolistic competition. Economic Letters 57: 177–182

Güth W, Kareev Y, Kliemt H (2001): How to play randomly without random generator – The case of maximin players. Homo oeconomicus (forthcoming)

Güth W, Kliemt H (1994): Competition or co-operation – On the evolutionary economics of trust, exploitation and moral attitudes. Metroeconomica 45: 155–187

Güth W, Kliemt H (1998) The indirect evolutionary approach: Bridging the gap between rationality and adaptation. Rationality and Society 10 (3): 377–399

Güth W, Kliemt H (2000) Evolutionarily stable co-operative commitments. Theory and Decision 49: 19–221

Güth W, Kliemt H, Peleg B (2000) Co-evolution of preferences and information in simple games of trust. German Economic Review 1(1): 83–110

Güth W, Nitzan S (1997) The evolutionary stability of moral objections to free riding. Economics and Politics 9: 133–149

Güth W, Yaari M (1992) An evolutionary approach to explain reciprocal behavior in a simple strategic game. In: Witt U (ed) Explaining process and change – Approaches to evolutionary economics, pp 23–34. The University of Michigan Press, Ann Arbor, MI

Harsanyi JC, Selten R (1988) A general theory of equilibrium selection in games. MIT Press, Cambridge, MA

Hammerstein P, Selten R (1994) Game theory and evolutionary biology. In: Aumann RJ, Hart S (eds) Handbook of game theory, Vol 2. North-Holland, Amsterdam

Kockesen J, Ok EA, Sethi R (1997) On the strategic advantage of negatively interdependent preferences. C. V. Starr Center Working Paper, pp 97-34. New York University, New York, NY

Maynard Smith J (1982) Evolution and the theory of games. Cambridge University Press, Cambridge, MA

Maynard Smith J, Price GR (1973) The logic of animal conflict. Nature 246: 15–18

Milgrom P, Weber R (1985) Distributional strategies for games with incomplete information. Mathematics of Operations Research 10: 619-632

Myerson R (1978) Refinements of the Nash equilibrium concept. International Journal of Game Theory 7: 73–80

Myerson R (1979) Incentive compatibility and the bargaining problem. Econometrica 47: 61–73

Ok EA, Vega-Redondo F (1999) On the evolution of individualistic preferences: Complete versus incomplete information scenarios. Mimeo

Possajennikov A (2000) On evolutionary stability of spiteful preferences. Journal of Economic Behavior and Organization 42: 125–129

Schaffer M (1989) Are profit-maximisers the best survivors? Journal of Economic Behavior and Organization 12: 29–45

Selten R (1975) Reexamination of the perfectness concept for equilibrium points in extensive games. International Journal of Game Theory 4: 25–55

Selten R (1983) Evolutionary stability in extensive two-person games. Mathematical Social Sciences 5 (3): 269–363

Selten R (1988) Evolutionary stability in extensive two-person games – Corrections and further development. Mathematical Social Sciences 16 (3): 223–266

Tirole J (1988): The theory of industrial organization. MIT Press, Cambridge, MA

van Damme E, Hurkens S (1996) Endogenous Stackelberg Leadership. CentER Discussion Paper, Nr. 96115, Tilburg University, Tilburg

van Damme E (1987) Stability and perfection of Nash equilibria. Springer, Berlin Heidelberg New York

Weibull JW (1995) Evolutionary game theory. MIT Press, Cambridge, MA