DISCUSSION PAPERS
Department of Economics
University of Copenhagen

99-01

Evolutionary Learning in Signalling Games

Hans Jørgen Jacobsen, Mogens Jensen,
Birgitte Sloth

Evolutionary Learning in Signalling Games[*]

Hans Jørgen Jacobsen, Mogens Jensen,[†] and Birgitte Sloth

Institute of Economics, University of Copenhagen

Studiestæde 6, DK-1455 Copenhagen K, DENMARK

Email: mogens.jensen@econ.ku.dk

March 1998

This version: April 8th, 1999

Abstract

We study equilibrium selection by evolutionary learning in monotone signalling games. The learning process is a development of that introduced by Young for static games extended to deal with incomplete information and sequential moves; it thus involves stochastic trembles. For vanishing trembles the process gives rise to strong selection among sequential equilibria. If the game has separating equilibria, then in the long run only play according to a specific separating equilibrium, the so-called Riley equilibrium, will be observed frequently. This selection is stronger than, and only partly in accordance with, traditional selection based on restictions on "out-of-equilibrium" beliefs.

1. Introduction

We define an evolutionary learning process in the spirit of Young (1993a) for monotone two-type signalling games and analyze it with respect to equilibrium selection. Monotonicity is

a property of most signalling games studied in economics. Our results are strongly in favor of separating (sequential) equilibria: whenever a monotone signalling game has separating equilibria, evolutionary learning selects the particular one known as the Riley equilibrium.

Many important economic situations involving asymmetric information and sequential moves are naturally formalized as signalling games. Particular games of interest often have many, and qualitatively very different, sequential or perfect Bayesian equilibria. This has led to refinements of sequential equilibrium based on restrictions on "out-of-equilibrium" beliefs such as the Intuitive Criterion suggested by Cho and Kreps (1987), and the Divinity concepts of Banks and Sobel (1987), etc.

Another, and very differently motivated, device for equilibrium selection is the evolutionary approach suggested by Canning (1992a), Young (1993a), and Kandori, Malaith, and Rob (1993), henceforth KMR. For static games of complete information these authors define evolutionary and/or learning processes by assuming some kind of boundedly rational learning behavior and small stochastic disturbances. They show that such processes naturally take the mathematical form of Markov chains, and that results on so-called stochastically stable states can be used to obtain strong equilibrium selection. For specific strategic form games the selection has been characterized, e.g., for 2×2 coordination games, Young (1993a) and KMR, and for the Nash demand game, Young (1993b).

We develop here the approach of Young (1993a) to signalling games, and thus suggest how Young's definition of a learning process can be extended to games with incomplete (asymmetric) information and sequential moves. As Young, we envisage that a game, here a signalling game, is played recurrently between two pools of players. We assume that the players keep records of past events *for each signal separately,* that is, the senders hold for each possible signal a record of responses observed the last m times that particular signal was sent, and the receivers have for each possible signal a record of which types sent the signal the last m times it was sent. In actual play the players choose best responses to samples from their records except for a tremble: with a small probability they take an arbitrary action.

In each round a single sender-receiver pair plays the game once. The sender is drawn at random from the pool of senders, such that the sender's type is determined according to a prior distribution. We assume that the receivers do not know the prior, and therefore receivers cannot use the prior when choosing a response to a signal. The only information available to a receiver about the senders' type is what is kept in the records. Having observed a signal a receiver draws a sample from the record concerning that signal and

2

uses the relative frequencies of types in the sample as belief about the sender's type. This procedure leads to what we call "drift in in-equilibrium beliefs". If for a lot of rounds the two types of senders both play the same signal as best response, then the receivers' beliefs about which type sent this signal, will fluctuate due to the random draws of senders. From time to time the belief will drift far away from the prior distribution, and this may change the best response of the receiver after this signal, even though the senders did not change their strategy. Drift in in-equilibrium beliefs is not due to trembles; it concerns beliefs after signals that are currently possible best responses for senders. Beliefs after other signals will only change when these signals are played by tremble. Such changes in beliefs we call "drift in out-of-equilibrium beliefs". It is basically the combination of drift in in-equilibrium and in out-of-equilibrium beliefs that creates the strong support for the Riley equilibrium that we obtain.

The evolutionary learning process is studied for signalling games with two possible types for the senders and one structural assumption, Monotonicity, imposed. A signalling game is called monotone if the possible actions of the receivers can be ordered such that both of two requirements are fulfilled: (i) Senders do, independently of type and signal, prefer "higher" responses to "lower". (ii) The higher probability receivers attach to a specific of the two types, the higher actions they prefer (the specific type is called the "high" one). As said, most signalling games studied in the literature are indeed two type, monotone games; this is true, e.g., for formalizations of entry deterrence situations, Milgrom and Roberts (1982), of job market signalling situations, Cho and Kreps (1987), of credit rationing or insurance situations, Hellwig (1987), etc.

As in the related literature our process takes the form of a Markov chain and by using results on stochastically stable states we can show that for small trembling probability a particular behavior will emerge as the only one observed frequently in the long run.

The signalling games studied in the literature usually have many separating sequential equilibria (and beside that also some pooling and/or semi-separating ones). Our main selection result is that for monotone signalling games, whenever there are separating sequential equilibria, the evolutionary learning process will support the particular one called the Riley equilibrium. The Riley equilibrium is the one which among all separating equilibria is best for the high type of sender (the one where the high type separates away from the low type at minimal cost).

If the game has no separating equilibrium, the behavior that emerges in the long run is that senders of high type send a specific signal all the time, namely the one that is best

for a sender of high type when all signals are responded to as if they came from the low type. Call this signal x_H^*. The low type will shift between x_H^* and the signal x_L^* that is best for a sender of low type when all signals are responded to as if they came from a low type. This behavior *may* correspond to a pooling sequential equilibrium, e.g., if $x_H^* = x_L^*$. As will also be explained, if there is indeed a pooling equilibrium in which both types of the sender choose x_H^* (but $x_H^* \neq x_L^*$), then it may well be that most of the time both types of the sender will play x_H^* according to our learning process. So, if there is no separating equilibrium our process may select in favor of a particular pooling sequential equilibrium, but this has to be one where both types of the sender choose the particular signal x_H^*, which of course requires that such a pooling equilibrium exists.

In a related paper Nöldeke and Samuleson (1997) consider an evolutionary process for Spence-type signalling games. Their process is an extension of the process for strategic form games studied by Kandori, Mailath, and Rob (1993). As we explain in detail in Subsection 3.3 below, such a process does not give rise to drift in in-equilibrium beliefs, which is the main reason why our results differ from those of Nöldeke and Samuleson; we get stronger support for separating equilibria than they do. Canning (1992b) is also closely related to our paper with respect to the definition of the learning process, but Canning considers a different class of games, namely so-called common interest signalling games (language games).

In Section 2 we describe signalling games and their equilibria formally and state the assumption of monotonicity. Section 3 defines our evolutionary learning process, and discusses in Subsection 3.3 important aspects of it. We attempt here to give the intuition for our main result, and describe the major forces driving it. Section 4 derives our results formally, and Section 5 compares our selection to that obtained by refinement concepts based on restrictions on out-of-equilibrium belief, and gives an example demonstrating the importance of the monotonicity assumption. All proofs are given in Appendix B.

2. Signalling Games

Signalling games, as introduced by Cho and Kreps (1987), are played between two players, the sender and the receiver. First the sender chooses a signal $x \in X$. The receiver observes the sender's signal and then chooses an action in response $a \in A$. The sender has some private information, a type $t \in \{L, H\}$. Payoffs depend on the sender's type, the signal sent, and the response given; the sender's payoff is $u(t, x, a)$, and the receiver's is $v(t, x, a)$, both are von Neumann-Morgenstern utilities. Finally, there is a prior probability $\pi \in \,]0, 1[$

for the sender being of type H. In traditional equilibrium analysis π is interpreted as the receiver's subjective probability assessment of the sender's type and assumed to be common knowledge among the players. For the learning process defined below the prior is an objective probability determining the actual type of the sender in each round.

A (pure) strategy of the sender is a function $x(\cdot)$, giving for each type $t = L, H$, a specific signal $x(t) \in X$. A strategy of the receiver is a function $a(\cdot)$, giving for each signal $x \in X$, the action $a(x) \in A$, that the receiver will play conditional on signal x being observed. The updated belief of the receiver is a function $\mu(\cdot)$, giving for each signal x, the probability $\mu(x) \in [0, 1]$, that the receiver assigns to type H, conditional on signal x being observed.

The actions that give highest expected payoff for a receiver who has observed signal x and holds belief μ, are actions a that solve $\max_{a \in A} [\mu v(H, x, a) + (1 - \mu)v(L, x, a)]$. Call the set of all such actions $BR(\mu, x)$.

A sequential (or perfect Bayesian) equilibrium, SE, is an assessment $(\mu(\cdot), x(\cdot), a(\cdot))$, such that,

(a) The sender's strategy is optimal against the receiver's strategy. For $t = L, H$, the signal $x(t)$ is a solution to $\max_{x \in X} u(t, x, a(x))$,

(b) The receiver's strategy is optimal against each signal and the belief held at that signal. For all $x \in X$, the action $a(x)$ is in $BR(\mu(x), x)$, and

(c) The receiver's beliefs are derived from the sender's strategy by Bayes' rule when possible. If the equilibrium is separating, that is, if $x(L) = x_L \neq x(H) = x_H$, then $\mu(x_L) = 0$, and $\mu(x_H) = 1$, and if it is pooling, that is, if $x(L) = x(H) = x'$, then $\mu(x') = \pi$. There are no other restrictions on beliefs.

The following assumption spells out what signalling games are usually all about,

A1 (Monotonicity). There is an order $<$ on A such that:

(i) Higher responses are better for the sender: For $t = L, H$ and for all $x \in X$, if $a < \hat{a}$, then $u(t, x, a) < u(t, x, \hat{a})$.

(ii) Higher belief on the high type (H) makes higher actions preferable for the receiver: For all $x \in X$, if $\mu < \bar{\mu}$, and $a \in BR(\mu, x)$, and $\hat{a} \in BR(\bar{\mu}, x) \backslash BR(\mu, x)$, then $a < \hat{a}$.[1]

---

[1]It is for convenience we call a signalling game monotone when it fulfils *both* of (i) and (ii). Cho and Sobel (1990)'s assumption (A1') is basically (i), whereas their (A3) is (ii), and they call a game monotonic if it fulfils just (A1').

In most applications, what is important for the sender is to obtain a response in a certain direction, which is here called "higher". The mean to achieve this end is to try to give the impression that one is of a specific type, which is here called high. This is all that A1 says. For instance, for Spence job market signalling games as stated in Cho and Kreps (1987) and more generally in Fudenberg and Tirole (1991), for entry deterrence games as studied in Milgrom and Roberts (1982),[2] and for banking, credit rationing games (or equivalently, insurance market games) as surveyed in Hellwig (1987), monotonicity is indeed fulfilled.

We assume (unrestrictively) that $BR(0, x)$ and $BR(1, x)$ are singletons for all x, and let their single elements be denoted by $\underline{a}(x)$ and $\bar{a}(x)$ respectively, so $\bar{a}(x)$ is simply the solution to $\max_{a \in A} v(H, x, a)$, etc. By A1, $\bar{a}(x)$ is the highest action the receiver will possibly take after observing signal x, while $\underline{a}(x)$ is the lowest. The actions a, for which $\underline{a}(x) \leq a \leq \bar{a}(x)$, according to the order of A1, we call the rational actions after signal x.

The two particular signals x_L^* and x_H^* are defined as the best signals for a sender of type L and H respectively, under the assumption that all signals are responded to as if they came from type L for sure, that is, x_H^* is the solution in x to $\max_{x \in X} u(H, x, \underline{a}(x))$, which we (unrestrictively) assume to be unique, and similarly x_L^* solves $\max_{x \in X} u(L, x, \underline{a}(x))$. For some signalling games of economic interest $x_L^* = x_H^*$; this is for instance the case for Cho and Kreps (1987)'s formulation of the Spence job market signalling game (where education does not affect productivity). Note that x_t^* is an insurance level strategy giving the lower payoff limit $u(t, x_t^*, \underline{a}(x_t^*))$ for the sender of type $t = L, H$.

The signal x_H^* is of importance for the selection obtained below. This signal may be different from x_H^π, the best signal for type H given that every signal x will be responded to as if it came from a type H with probability π, that is, by $a \in BR(\pi, x)$, where the latter is unrestrictively assumed to be a singleton. Generally, one could define x_H^μ as a signal that is best for a sender of type H, when every signal x is responded to by an action in $BR(\mu, x)$. For many signalling games of interest, $x_H^\mu = x_H^*$ for all μ. This follows, for instance, if the payoff of the receiver only depends on his own action and the sender's type, $v(t, x, a) = v_1(t, a)$, and the sender's payoff function has the separable form $u(t, x, a) = u_1(t, x) + u_2(t, a)$. These conditions imply that $BR(\mu, x)$ only depends on μ. So, for given μ, the sender's signal does not influence the receiver's best response, and therefore x_H^μ simply comes from maximizing $u_1(H, x)$ with respect to x. When we obtain below selection among pooling equilibria in favor of those where both types choose the signal x_H^*, this is thus not necessarily in contradiction to selection in favor of x_H^π.

---

[2] Strictly speaking, the game studied in Milgrom and Roberts is not a signalling game, but it can be reduced to an equivalent game that is, and this equivalent game fulfils A1.

In a separating SE, where $x(L) = x_L$, and $x(H) = x_H$, one must, since $\mu(x_L) = 0$, and $\mu(x_H) = 1$, have $a(x_L) = \underline{a}(x_L)$, and $a(x_H) = \overline{a}(x_H)$. Since type L gets the worst possible response, he only plays optimally if $x_L = x_L^*$. Further, the equilibrium payoff $u(L, x_L^*, \underline{a}(x_L^*))$ of type L must be at least as high as the payoff that could be obtained by deviating to x_H (thereby obtaining the response $\overline{a}(x_H)$), and the equilibrium payoff $u(H, x_H, \overline{a}(x_H))$ of type H must be weakly greater than what would at least be obtained from deviating to x_H^*,

$$u(L, x_L^*, \underline{a}(x_L^*)) \geq u(L, x_H, \overline{a}(x_H)). \tag{2.1}$$

$$u(H, x_H, \overline{a}(x_H)) \geq u(H, x_H^*, \underline{a}(x_H^*)). \tag{2.2}$$

On the other hand, if a signal x_H fulfills (2.1) and (2.2), then at least one separating SE with $x(L) = x_L^*$, and $x(H) = x_H$ exists, namely the one where beliefs $\mu(x)$ are zero for all signals $x \neq x_H$, implying the bad response $a(x) = \underline{a}(x)$.

Let X^S be the set of separating signals for type H, that is, the set of signals x_H fulfilling (2.1) and (2.2). Note that the set X^S does not depend on the prior π. If X^S is non-empty, one defines the so-called Riley signal r, as the specific signal that, among all signals x in X^S, maximizes $u(H, x, \overline{a}(x))$, so it is the signal of a sender of type H in the separating equilibrium that is best for type H, the Riley equilibrium.

In a pooling SE with $x(L) = x(H) = x'$, one has $\mu(x') = \pi$, and therefore x' must be responded to by an action $a(x')$ in $BR(\pi, x')$. So, higher π pulls in the direction of a higher equilibrium response $a(x')$. In equilibrium the sender must, independently of type, get weakly higher payoff than what could at least be obtained by deviating to the insurance level signal, so it must be possible to choose $a(x')$ in $BR(\pi, x')$ such that,

$$u(L, x', a(x')) \geq u(L, x_L^*, \underline{a}(x_L^*)). \tag{2.3}$$

$$u(H, x', a(x')) \geq u(H, x_H^*, \underline{a}(x_H^*)). \tag{2.4}$$

On the other hand, if (2.3) and (2.4) are fulfilled for some x', then at least one pooling equilibrium using this signal exists, namely one where the receiver's beliefs $\mu(x)$ are zero for all $x \neq x'$, implying $a(x) = \underline{a}(x)$. We denote by X^P the set of signals x' fulfilling (2.3) and (2.4) for a possible choice of $a(x')$ in $BR(\pi, x')$; thus X^P is the set of pooling signals. Note that this set depends on π. A higher π gives the possibility of choosing a higher $a(x')$, and hence increases the left hand sides above. Therefore, the set X^P is weakly increasing in π, and it may well be empty for small, and non-empty for larger, values of π.

A specific signalling game may have no separating or pooling equilibrium, in which case it does have semi-separating SE where type H plays a signal x_H for sure, and type L mixes

between x_L^* and x_H. In particular, there will be one such equilibrium where $x_H = x_H^*$, which is also of importance for the equilibrium selection obtained below. However, for most games studied in applications, not only are there pure strategy equilibria, but there are *separating* SE, that is, X^S is non-empty.

3. The Evolutionary Learning Process

In order to use certain results from the theory of Markov chains we will have to assume that the sets X and A are finite; let n and p be their cardinalities, respectively. Having assumed finiteness the following assumption holds generically, and implies as already assumed above, that $BR(0, x)$ and $BR(1, x)$ are singletons, and that x_t^* is uniquely defined.

A2 For both types $t = L, H$:

(i) If $(x, a) \neq (\hat{x}, \hat{a})$, then $u(t, x, a) \neq u(t, \hat{x}, \hat{a})$.

(ii) For all $x \in X$: If $a \neq \hat{a}$, then $v(t, x, a) \neq v(t, x, \hat{a})$.

The first part of A2 implies that the sender always has a unique best response given a pure strategy of the receiver. The second part says that given the sender's type and signal the payoff to the receiver is different for any two actions in A, implying that for any x, $BR(\mu, x)$ is a singleton for almost all $\mu \in [0, 1]$, and especially for $\mu = 0$, $\mu = 1$, and (generically) for $\mu = \pi$.

Consider two large pools of players playing a signalling game recurrently. One pool contains senders and is subdivided into types, the other contains receivers. In each round of play a sender and a receiver are picked randomly and independently with small probability for any individual player to be picked. The probability of picking a sender of type H is π. The picked players play the game and afterwards, before the next round, everybody observes the type of the sender as well as the signal and the action taken in the round.

The players of each pool hold a common record of past opponent play. For each $x \in X$, let g^x be the senders' common record of the actions taken by the receivers the last m times signal x was sent by senders. The first element of g^x is the response by the receiver the m'th last time a sender played x, the last element is the response by the receiver the very last time a sender played x. Thus g^x is an m-vector, where each element is an action a from A. Similarly, h^x is the receivers' common record of the senders' types in the last m rounds where x was chosen, so h^x is an m-vector, and each element is a type, L or H. In the following we will always assume that for each x, the record g^x contains only rational

actions, i.e., each element a in g^x fulfills $\underline{a}(x) \le a \le \bar{a}(x)$. This amounts to assuming that receivers never, not even by mistake, take an action that could not be best for any belief about the sender.[3]

The state of the learning process is given by the currently held records $\{(g^x)_{x \in X}, (h^x)_{x \in X}\}$ of past play. The state space Ω is the set of all possible records, which is finite. From one round to the next the state is updated as follows. If the signal played in the round was x', the picked sender's type was t, and the picked receiver's response was a, then all records g^x and h^x for which $x \ne x'$ are unchanged, while the first element in each of $h^{x'}$ and $g^{x'}$ is deleted, and a t is inserted as last element in $h^{x'}$, and an a is inserted as last element in $g^{x'}$. A state $\bar{\omega} = (\bar{g}^x, \bar{h}^x)$ is a successor of $\omega = (g^x, h^x)$, if $\bar{\omega}$ can be constructed from ω by this procedure of updating by appropriate choices of signal, type, and rational action.

3.1. The Unperturbed Process, P^0

When a player is to choose an action, he draws a sample from the current record, forms beliefs of the opponent's action or type by computing the relative frequencies from the sample, and then chooses a (myopic) best response. Precisely, when a sender is to choose a signal, he picks k elements from *each* record g^x, and uses the relative frequencies of actions in the sample as his conjecture of what action the receiver will take after x. Let the so derived conjecture concerning signal x be denoted by σ_x, where $\sigma_x(a)$ is the relative frequency of a in the sample from g^x. The picked sender then plays the signal that yields the highest expected payoff, that is, if he is of type t, he chooses a signal x that maximizes $\sum_{a \in A} u(t, x, a) \sigma_x(a)$. The receiver, after observing signal x, samples k elements from the record h^x, and uses the relative frequency of H in the sample, call it μ_x, as belief of the opponent's type. He then chooses an action yielding maximal expected payoff given this belief, that is, an action in $BR(\mu_x, x)$. We assume that k is strictly less than $m/2$, and that k is such that for no samples ties occur.[4]

This defines a Markov chain on Ω, where the transition probabilities are given as follows. For any state $\bar{\omega}$ that is not a successor of ω, the transition probability $p^0(\omega, \bar{\omega})$ is zero. Let

---

[3]This assumption simplifies the analysis, but is not essential to our results as will be explained in Appendix B.

[4]By A2 ties occur only at knifeedge beliefs, so such a k always exists. An assumption of "not too much sampling", here $k < m/2$, is standard in the literature, see Young (1993a) and (1993b), and serves to avoid cyclical behavior of the unperturbed process which could lead to the existence of "strange" absorbing sets. It is exactly in the proof of Proposition 4.3 below, which characterizes the set of absorbing sets, that $k < m/2$ is used.

$\bar{\omega}$ be the successor of ω that appears by updating with respect to signal x', and has t' as the last element in $\bar{h}^{x'}$, and a' as the last element in $\bar{g}^{x'}$. The transition probability $p^0(\omega, \bar{\omega})$ equals the probability of drawing type t' multiplied by the probability that the sender draws samples from $(g^x)_{x \in X}$ making x' the best response multiplied by the probability that the receiver draws a sample from $h^{x'}$ making a' the best response (this may be zero). The transition probabilities are collected in the transition matrix P^0.

A set $O \subseteq \Omega$ is absorbing if it is minimal with respect to the property,

$$\text{if } \omega \in O \text{ and } p^0(\omega, \bar{\omega}) > 0, \text{ then } \bar{\omega} \in O. \tag{3.1}$$

A set is absorbing if for every state in the set the probability of transition to a state in the set is one, and for every two states in the set there is positive probability of transition between the two states in a finite number of rounds. If the singleton $\{\omega\}$ is absorbing, we say that ω is an absorbing state.

A probability distribution η on Ω is an invariant distribution for P^0, if $\eta P^0 = \eta$. So, if the process starts in some state drawn randomly according to η, then the (ex ante) probabilities of the different states in the following round are also given by η.

There is always at least one absorbing set, since the whole state space Ω fulfills (3.1), but may not be minimal. The following results are well known from the theory of Markov chains, see, e.g., Billingsley (1995),

1. *There is probability one of ending up in an absorbing set in finite time.*

2. *For each absorbing set O, there exists a unique invariant distribution with support O, and for each invariant distribution every state in the support belongs to an absorbing set.*

3.2. The Perturbed Process, P^ε

As is standard in the literature, we will study a Markov chain P^ε, which is a slight perturbation of P^0. The idea is that, as compared to P^0, a process in which players may occasionally tremble, is more realistic. We simply assume that the description defining P^0 is modified as follows: whenever a player is to move, with probability $\varepsilon > 0$, the player chooses a random signal or a random rational action - let's say drawn from the uniform distribution over possible signals or rational actions[5] - while with probability $1 - \varepsilon$ the

---

[5] The assumption can be relaxed. All that is required is that the probability of each signal for the sender and each rational action for the receiver is strictly positive, and that the probabilities are independent of the state, see Bergin and Lipman(1996).

player behaves according to the unperturbed process. The trembles can be interpreted as mistakes, experiments, or mutations. This leads to a perturbed Markov chain on Ω with transition probabilities $p^\varepsilon(\omega, \bar{\omega})$ collected in the matrix P^ε. The new process is irreducible: for any two states, there is positive probability of transition from one to the other in a finite number of rounds. This holds since there is strictly positive probability of transition between two arbitrary states ω and $\bar{\omega}$ in at most mn steps; just let the types of sender be drawn appropriately, and let the players choose the relevant signals/actions by trembles or as best response. Irreducibility implies that the long run behavior of the process in terms of the relative frequencies over time of the different states does not depend on the initial state, since it is a general result that,

3. *If a Markov chain is irreducible, then it has a unique invariant distribution, and independently of the initial state, (with probability one) the relative frequencies by which the different states are visited over time converge to the probabilities of the unique invariant distribution as time goes to infinity.*

Thus to predict the long run behavior of the process in terms of relative frequencies we just need to identify the invariant distribution. The process P^ε is also aperiodic: for any two states ω_1 and ω_2, there exists $T < \infty$ such that transition has positive probability in $T, T + 1, T + 2$ etc. rounds. This can be seen as above, since transition obviously is also possible in $mn + 1$ steps. For irreducible and aperiodic (that is, ergodic) processes the long run behavior in terms of probabilities of the states also does not depend on the initial distribution, since it is a general result that,

4. *If a Markov chain is irreducible and aperiodic, then independently of the initial distribution, the (ex ante) probabilities of the different states after T rounds converge to the probabilities of the unique invariant distribution, as T goes to infinity.*

We are interested in identifying the invariant distribution of P^ε (for ε small), since this characterizes the long run behavior both in terms of relative frequencies of the states over time and in terms of the long run, ex ante probabilities of the states.

The process P^ε is a so-called regular perturbation of P^0: P^ε is irreducible and aperiodic, and for each transition $(\omega, \bar{\omega})$ the probability $p^\varepsilon(\omega, \bar{\omega})$ tends to $p^0(\omega, \bar{\omega})$ as ε goes to zero, and there is a well-defined speed by which this happens, see Young (1993a) for a formal definition. It follows from Young (1993a), Theorem 4, that,

5. *As ε goes to zero, the associated sequence (η^ε) of the unique invariant distributions of P^ε, converges to a specific distribution η^0, and this distribution is invariant for P^0, that*

is, $\eta^0 P^0 = \eta^0$.

Since we are interested in the long run behavior of P^ε for ε small, our concern will be with η^0. It follows from the above that (for small ε) *the states in the support of η^0, called the stochastically stable states, are the states, and the only states, that will be observed frequently in the long run* if players learn, play, and tremble as described, and ϵ is small. It also follows that only states in absorbing sets of P^0 can be stochastically stable.

In the following section we characterize the stochastically stable states. Our main result is that if the signalling game has separating SE, then only states that correspond to the Riley equilibrium will be stochastically stable, implying that only play as in the Riley equilibrium will be observed frequently in the long run when ε is small.

Before demonstrating this we want to point to the major forces driving our results and discuss important aspects of our evolutionary learning process. This is done most illustratively by comparing to an alternative evolutionary approach.

3.3. Drift in in-Equilibrium Beliefs

The evolutionary learning process defined above is a development of the approach of Young (1993a) to signalling games. A similar development of the evolutionary approach of KMR, as suggested by Nöldeke and Samuelson (1997), would go as follows. It would be assumed that all players in each player pool play in each round. Random matching would decide which senders would meet which receivers. The fraction π of senders would be of type H. The players' expectations concerning a given signal, that is, the receivers' probability assessment of type and the senders' probability assessment of action in response, would in a given round be determined by the empirical frequencies of types and actions over the entire sender and receiver populations respectively for the signal in question in the immediately preceding round. Play from earlier rounds would be "forgotten". The difference is that according to the KMR approach expectations depend on observation of the populations' behaviors in one, namely the last, round, whereas according to Young's approach expectations are derived from observation of the behavior of a string of pairs back in time. Although the two approaches lead to very similar processes for simultaneous move games of complete information, their developments to signalling games are essentially different, and the differences are due both to sequential moves and to incomplete information.

First, consider a KMR type of process and assume that in a certain round a specific signal, x say, was not played by any sender. In the next round no player can then have an

observation based expectation concerning the signal x. This has important consequences. It must be assumed exogenously where the expectations given x (who sent it, what will the response be) go. Nöldeke and Samuelson (1997) assume that these expectations stay unchanged from the last time there was an observation based expectation. After a round where x was not played, the beliefs concerning x will be extremely sensitive to trembles. If exactly one sender of type L, say, by experiment or mistake plays x, then in the next round the belief of all receivers will be that x has to come from a type L. We describe this as the process being very sensitive to "drift in out-of-equilibrium beliefs", where the term *out-of-equilibrium* is used to indicate that the described change of belief concerns a signal that is currently *not* played with positive probability (according to the unperturbed process). The relative strength of drift in out-of-equilibrium beliefs is important for the equilibrium selection obtained by Nöldeke and Samuelson (1997).

For the Young type of process, on the other hand, it is an important feature that players always have observations, that is, records g^x and h^x, concerning any possible signal x. If some signal has not been played for a long time, the observations concerning this will be old, but it never happens that the "memory" concerning a signal is emptied, and therefore the players will always base their conjectures about opponent's type or play on observations, some of which may be old. It seems to us more reasonable to assume that players base their expectations on observations, even if old, than on something else, and this is one reason why we find a development of the Young approach to signalling games interesting (note here that when players tremble, all signals will indeed occasionally be observed). Also for a Young process, beliefs concerning signals not presently played according to the unperturbed process are sensitive to trembles, and this drift in out-of-equilibrium beliefs is indeed one of the driving forces for our results, but in our model those beliefs are not extremely sensitive to trembles; one experiment or mistake will not completely change the players' beliefs.

A second important difference origins from incomplete information. In a KMR process, since all senders would play in each round, the receivers could simply observe how many were type H; the true prior would be known to the receivers. For the Young process, as we have defined it, the types of those senders who were actually picked for play in a finite number of former rounds are observed, but this does not completely reveal the prior. The receivers keep track of which types sent signal x the last m times it was sent, and use (a sample from) this record to form a probabilistic expectation, a posterior (or belief), of type the next time an x is observed. Estimation of the prior is not even attempted.[6]

---

[6] Alternatively, the receivers could be assumed to estimate from a given set of past observations both

We find the assumption that the prior is neither known, nor precisely estimated, and hence the procedure of direct estimation of posteriors (beliefs), in good accordance with the boundedly rational, learning type of behavior that is anyway assumed.[7] The procedure leads to a phenomenon that can descriptively be called "drift in in-equilibrium beliefs", and this is the other major force driving our results.

To explain it assume that currently the records are such that (for any sampling) both types of sender will find it optimal to play a specific signal, x' say. To be precise assume that $g^{x'}$ contains many, relatively "high" actions, while for other signals x, the records g^x contain many bad responses close to $\underline{a}(x)$. Assume further, that the receivers' records contain relatively many H's in $h^{x'}$, so that the receivers give relatively good responses to x', and many L's in other records h^x, so other signals are given bad responses. For the sake of the argument assume that this corresponds to a pooling equilibrium: there is indeed a SE, where both types of sender choose x'. This is supported by beliefs $\mu(x') = \pi$, and $\mu(x) = 0$ for $x \neq x'$. The prior π is large enough to ensure that the best action after x', call it a', is so good for the sender, that x' followed by a', is better for both types of sender than any alternative x followed by $\underline{a}(x)$. Now, as the game is played recurrently it will happen with positive probability according to the unperturbed process, that is, without any trembles, that many senders of type L enter in a row (or just in a sufficiently strong concentration). They will all play the signal x'. This will make the receivers' record $h^{x'}$ contain many L's, so beliefs constructed from samples from $h^{x'}$ will be twisted towards higher probability on L. This will eventually change the best response after signal x' towards $\underline{a}(x')$, that is, to an action that is not so good for the senders, and as the record g^x is now being filled up with bad actions, it may well happen that for one of the types of sender, the best signal changes. In this way, without any trembling at all, the pooling situation that one started from is upset, and records may be brought to a place from which it may require no or only a little trembling to go further all the way to a separating situation. We call this drift in *in-equilibrium* beliefs because the described change in beliefs concerns a signal that *is* currently played with positive probability (according to the unperturbed process).

the prior and the (mixed) strategies of types L and H respectively, and derive posteriors for each signal x by applying Bayes' rule. Generally, for a given set of observations, however, this can never improve the estimates of posteriors. Of course, if the receivers knew the true prior or estimated it based on additional observations or other information, the alternative procedure might be better.

[7]Note that sequential moves play a role here. Since the receivers do not have to move before a certain signal has been observed, it seems natural to record types, and to estimate directly the posterior probability of type (that is, the belief), for each signal separately. In a simultaneous move game of incomplete information it may be more reasonable to assume that type is estimated from a larger set of past observations than are strategies, see Jacobsen, Jensen, and Sloth (1998).

An initial separating situation corresponding to a separating equilibrium cannot be similarly upset. Assume that there is a separating SE using signals x_L^* and x_H, and consider records which are currently such that (for any sampling), x_L^* is best for type L, and x_H is best for type H, and x_L^* is responded to by $\underline{a}(x_L^*)$, because $h^{x_L^*}$ contains many L's, and x_H is responded to by $\overline{a}(x_H)$, because h^{x_H} contains many H's. Again, it has positive probability (without trembling) that a lot of type L's (or H's) enter in a concentration, and this will insert many L's in $h^{x_L^*}$ (or H's in h^{x_H}), but $h^{x_L^*}$ already contained many L's (and h^{x_h} many H's), and therefore no players will change behavior; the separating situation is not upset.

This gives the intuition why pooling equilibria are more fragile than separating according to our evolutionary learning process, and hence why we obtain support for (a particular) separating equilibrium. Solely by a force associated to the unperturbed process, drift in in-equilibrium beliefs, most pooling situations are upset, but separating are not; separating equilibria are self-confirming, while pooling are not. To reach the full conclusion that only behavior in accordance with a particular separating equilibrium is observed in the long run *also* requires the force of trembles, and the induced drift in out-of-equilibrium beliefs, but a key is to understand that most pooling situations are upset solely by "first order" events (with positive probability even in the absence of trembles), while separating can only be upset by "second order" events (only by trembles). A KMR process as in Nöldeke and Samuelson (1997) is different in this respect. If all senders in the population in one round play a specific signal x', then the belief of type in the next round, given x' is observed is simply π leading to a', and thus a pooling SE can be self-confirming. This is the main reason for a main difference in results between this paper and Nöldeke and Samuelson's: that we obtain stronger support for separating equilibrium than they do.

4. Long Run Behavior

As explained above, a state can be stochastically stable only if it belongs to an absorbing set for P^0. It is therefore of interest to identify these absorbing sets.

4.1. Absorbing sets

Let $x(L) = x_L^*$, $x(H) = x_H$ be part of a separating SE. Consider a state ω that reports that signal x_L^* was always sent by type L, and always responded to by $\underline{a}(x_L^*)$, and x_H was always sent by type H, and responded to by $\bar{a}(x_H)$. If the records g^x for other signals are

such that for all possible samples, x_L^* is best response for sender type L, and x_H is best response for type H, then ω is an absorbing state. If a sender of type L is drawn in the next round, the signal x_L^* will be played and it will be responded to by $\underline{a}(x_L^*)$ leaving all records unchanged, and similarly if type H is drawn, the signal x_H will be played leaving all records unchanged. This proves,

Lemma 4.1. *Assume that $X^S \neq \emptyset$, and let $x_H \in X^S$. If the state $\omega = (g, h)$ fulfils, that*

$$g^{x_L^*} = (\underline{a}(x_L^*), \ldots, \underline{a}(x_L^*)), \ h^{x_L^*} = (L, \ldots, L), \text{ and} \tag{4.1}$$

$$g^{x_H} = (\bar{a}(x_H), \ldots, \bar{a}(x_H)), \ h^x = (H, \ldots, H),$$

and for any sampling, x_L^* is best response for a sender of type L, and x_H is best response for a sender of type H, then ω is absorbing.

If there exists a separating sequential equilibrium, $x(L) = x_L^*$, $x(H) = x_H$, then there are indeed absorbing states fulfilling the assumptions of the lemma; just let the records for signals $x \neq x_H$ be $g^x = (\underline{a}(x), \ldots, \underline{a}(x))$. In such a state behavior will, according to the unperturbed process, be as in the separating equilibrium all the time, and we say that the state corresponds to the equilibrium in question.

There is also another type of absorbing sets, which may correspond to pooling equilibria. Suppose a state ω has been reached for which all the records g^x for signals other than x_H^* contain so many "bad" actions close to $\underline{a}(x)$, that even if the record $g^{x_H^*}$ consisted entirely of $\underline{a}(x_H^*)$s, the best signal for senders of type H would be x_H^* for any possible sampling. So, senders of type H are currently locked at signal x_H^* according to the unperturbed process. Assume further that for signals x different from x_H^* and x_L^*, ω is such that x is not for any sampling a best signal for type L, and that if $x_L^* \neq x_H^*$ then record $h^{x_L^*}$ contains only Ls, and $g^{x_L^*}$ contains only $\underline{a}(x_L^*)$s. Then according to the unperturbed process all following states must also have these properties, such that there is an absorbing set consisting of states with exactly these properties, because all records g^x, h^x for signals x other than x_H^* must be unchanged in all future rounds. If a type H enters he plays x_H^*, which leaves all records g^x, h^x for x different from x_H^* unaffected, and cannot change the property that x_H^* is best for type H (it can only affect $g^{x_H^*}$ in a way that makes samples from it better than only $\underline{a}(x_H^*)$s). If a type L enters, either he plays x_L^* different from x_H^*, and gets response $\underline{a}(x_L^*)$, which only confirms the records concerning x_L^* as they are, or he plays x_H^*, and obtains some response, which may affect the records concerning x_H^*, but not in a way that can change the property that x_H^* is best for type H (since this was before best even for the

worst possible sampling concerning x_H^*, and records g^x for other signals have not changed). So, in all future rounds x_H^* is best for type H, and either x_L^* or x_H^* is best for type L, and hence records for other signals than x_H^* can never change. This almost proves the lemma below, but a formal proof is given in Appendix B,

Lemma 4.2. *Assume $x_H^* \notin X^S$.[8] For each possible choice of $(\bar{g}^x, \bar{h}^x)_{x \neq x_H^*}$ which fulfills,*

(1) If $x_L^* \neq x_H^*$, then $g^{x_L^*} = (\underline{a}(x_L^*), \ldots, \underline{a}(x_L^*))$, and $h^{x_L^*} = (L, \ldots, L)$, and

(2) For any sample for sender, which for $x \neq x_H^*$ is sampled from $\bar{g}^x$ and for x_H^* is $(\underline{a}(x_H^*), \ldots, \underline{a}(x_H^*))$, the best response for type t is x_t^* for $t = L, H$,

there is exactly one absorbing set such that all states in the set have $g^x = \bar{g}^x$, and $h^x = \bar{h}^x$ for $x \neq x_H^*$.

There is always at least one absorbing set as described in this lemma; just let g^x and h^x for all signals x different from x_H^* contain only $\underline{a}(x)$'s and L's respectively; then the lemma's assumptions are fulfilled. We call an absorbing set as in Lemma 4.2, an x_H^*-absorbing set.

Now, consider the case $x_L^* = x_H^*$. Then there is indeed a pooling SE with $x(L) = x(H) = x_H^*$, supported by beliefs $\mu(x) = 0$ for $x \neq x_H^*$. Every x_H^*-absorbing set will correspond to this pooling equilibrium in the sense that in any state of the set, both types of sender will choose x_H^* (the response to x_H^*, however, will not necessarily always be in $BR(\pi, x_H^*)$, as in the equilibrium; due to drift in in-equilibrium beliefs it will from time to time go close to $\underline{a}(x_H^*)$ or to $\overline{a}(x_H^*)$).

Consider next the case $x_L^* \neq x_H^*$. There may or may not (depending on π) be pooling sequential equilibria where $x(L) = x(H) = x'$, and if there are, there may or may not be one with $x' = x_H^*$. Under all circumstances, however, one will according to the unperturbed process circulate around in an x_H^*-absorbing set by drift in in-equilibrium beliefs as follows. If in the current state both types of sender have x_H^* as best signal (for any sampling), then with positive probability a sufficiently strong concentration of senders of type L will enter to ensure that $g^{x_H^*}$ is finally filled so much up with $\underline{a}(x_H^*)$s, that the best signal of type L changes to x_L^*. When one has reached such a state, on the other hand, one will be brought back to a state of the first kind when sufficiently many type Hs have arrived, and they do not have to come in a concentration (every time a type H arrives and plays x_H^*, an H will

---

[8]We only assume that x_H^* is not a separating signal because, if it is, the absorbing sets described in this lemma are the same as the absorbing states described in the former.

be inserted in $h^{x_H^*}$, and eventually x_H^* will be responded to by $\overline{a}(x_H^*)$, filling $g^{x_H^*}$ up with $\overline{a}(x_H^*)$, finally changing the best signal of type L back to x_H^*).

So, in an x_H^*-absorbing set senders of type H will play x_H^* all the time, while senders of type L will shift between x_L^* and x_H^*. However, if there is a SE with pooling on x_H^*, requiring among other things a sufficiently large π, then intuition suggests that if the sample size k is sufficiently large, every x_H^*-absorbing set will correspond to this equilibrium in the sense that both types of sender will most of the time play x_H^*. As explained above, to get out of a state where this is the case requires that a number of type L's arrive in a row (or concentration), and this number is large when k is large, and the event therefore has relatively low probability ($1 - \pi$ is small). To get back to a "pooling state" one just has to wait for enough type H's, and these come with the relatively high probability π. Therefore, intuitively, if one has reached an x_H^*-absorbing set and there is a pooling equilibrium on x_H^*, then both types of sender will most of the time play x_H^* as in the equilibrium if k is large enough. However, the same intuition suggests that in order to come close to the equilibrium k may have to be extremely large.[9]

If there is no pooling equilibrium on x_H^*, then an x_H^*-absorbing set cannot be said to correspond to a pooling equilibrium. In this case behavior in an x_H^*-absorbing set more resembles play in the kind of equilibrium that was above called semi-separating.

Proposition 4.3. *There are no other absorbing sets than those described in Lemma 4.1 and Lemma 4.2.*

A first conclusion is that a pooling equilibrium with pooling on a signal different from x_H^* cannot have an absorbing set that corresponds to it. To conclude further with respect to absorbing sets: If $X^S = \emptyset$, implying $x_H^* \notin X^S$, then there are only x_H^*-absorbing sets. In these, senders of type H always play x_H^*, while senders of type L shift between x_L^* and x_H^*, so a specific pattern of behavior is connected to such absorbing sets, and since all stochastically stable states are in absorbing sets one can conclude without explicitly studying the perturbed process that this behavioral pattern will prevail in the long run. If $X^S \neq \emptyset$, then there are for each separating equilibrium absorbing singleton sets corresponding to that equilibrium, and if $x_H^* \in X^S$, these are the only absorbing sets, while if $x_H^* \notin X^S$, there

---

[9]Jacobsen and Sloth (1998) study an example, a $2 \times 2 \times 2$ entry deterrence game and an evolutionary learning process with $k = m$. For the example they state and prove formally the result that as k goes to infinity, behavior in the x_H^*-absorbing set converges to the behavior of the x_H^*-pooling equilibrium. Calculations reveal that, even for the simple example, a very large k may be needed to get reasonably close to the equilibrium behavior.

are also x_H^*-absorbing sets. So, when there are separating equilibria it requires explicit investigation of the perturbed process to find which patterns of behavior will prevail in the long run.

4.2. Stochastically stable states

These patterns are given by the stochastically stable states which can be found by using the characterization of Young (1993a), based on Freidlin and Wenzell (1984)'s characterization of the invariant distribution of a Markov chain.

Let Θ be the set of absorbing sets defined by P^0. For each pair of absorbing states (O_1, O_2), the resistance in the transition from O_1 to O_2, is the minimal number of probability ε-trembles (mistakes/experiments/mutations) required for going from a state in O_1 to a state in O_2. For each absorbing set $O \in \Theta$, an O-tree is a directed graph on Θ fulfilling, that there is a unique path of directed edges (transitions) leading to O from every other state $O' \in \Theta$, and from O there is no outgoing edge.

Obviously, for each O there can be many O-trees. For each O-tree, calculate the total resistance for the tree, i.e., the sum of resistances along its edges. The stochastic potential of O is the minimal total resistance among all O-trees. From Young (1993a) follows that,

6. *The stochastically stable states are exactly the states which are in the absorbing sets with minimal stochastic potential.*

If the signalling game under consideration has separating SE, a Riley state is an absorbing state that corresponds to the Riley equilibrium, and Θ_r is the set of the absorbing (singleton) sets each consisting of a Riley state.

Lemma 4.4. *If $X^S \neq \emptyset$, then for each absorbing set O, that is not a Riley state, there exists a Riley state ω, and a path from O to ω, such that the resistance is one in each transition along this path.*

A formal proof is given in Appendix B, but the intuition is simple. Consider an absorbing state ω as described in Lemma 4.1, corresponding to a separating equilibrium, $x(L) = x_L^*$, $x(H) = x_H$, different from the Riley one, $x_H \neq r$. For simplicity assume that ω is the particular state that for all signals $x \neq x_H$, including r, has the records g^x consisting entirely of $\underline{a}(x)$s, and the records h^x consisting entirely of Ls. Consider one tremble by a sender of type H, sending r instead of x_H. This will insert an H as last element in h^r, and

not give rise to any other changes of records. The new state ω' will also be absorbing and correspond to the same equilibrium as did ω. So, one has moved from one absorbing set to another by one tremble. From ω', consider again one tremble of a sender of type H, playing r. This will insert a second H in h^r, and (most probably) define a new absorbing state ω'', still corresponding to the same separating equilibrium. After a number of trembles there will be so many H's in h^r, that the receivers will (with positive probability) begin to respond with $\overline{a}(r)$ to the r's played by tremble. Further trembles by senders of type H playing r will then begin to insert $\overline{a}(r)$'s in g^r, again leading, tremble by tremble, to new absorbing states corresponding to the same old separating equilibrium. Eventually, however, there will be so many $\overline{a}(r)$'s in g^r, that for some samples from it, r will be the best signal for senders of type H, and one will now go to a Riley state without further trembles. So, one has gone all the way to a Riley state from absorbing set to absorbing set by only one tremble at a time. If one starts from an x_H^*-absorbing set, a similar argument applies, but here drift in in-equilibrium beliefs is important. Without any trembles the process will with positive probability reach states where the records concerning x_H^* contain many Ls and $\underline{a}(x_H^*)$s, and from such states the transition over absorbing sets by one tremble at a time all the way to a Riley state is similar.

Lemma 4.5. *If $X^S \neq \emptyset$, then for k sufficiently large, for each Riley state ω, and each absorbing set O, that does not consist of a Riley state, there is resistance at least two of going from $\{\omega\}$ to O.*

This is proved by checking that after any single tremble from a Riley state, the process will with probability one go back to a Riley state (the original state or another one). A main intuition is that one cannot go away from a Riley state and all the way to an absorbing set that is not a Riley state, in the same way (described above) as one can go away from an absorbing state corresponding to a non-Riley separating equilibrium and all the way to a Riley state, that is, jumping from absorbing set to absorbing set by one tremble at a time. Assume that senders of type H by mistake begin to play some other separating signal x instead of r. This will indeed by one mistake at a time lead to new Riley states and eventually insert $\overline{a}(x)$s into g^x. However, even when g^x consists entirely of $\overline{a}(x)$s, it will still, by definition of r, be best response for senders of type H to play r for any sampling. At the end of the chain of absorbing sets, one is still at a Riley state. Another important intuition is that one mistake by a receiver giving the bad response $\underline{a}(r)$ to signal r, cannot change the best response for senders of type H away from r, as long as k is sufficiently large, but note that often a moderate sized k will suffice.

Lemma 4.6. *Assume that the set Θ of absorbing sets can be partioned into two subsets Θ_1 and Θ_2 such that,*

(1) from any absorbing set $O \in \Theta_2$, there is a path ending in some $O' \in \Theta_1$, such that all transitions along the path have resistance one, and

(2) for any pair of absorbing sets $O_1 \in \Theta_1$, $O_2 \in \Theta_2$ the transition from O_1 to O_2 has resistance at least two.

then all stochastically stable states are elements in absorbing sets in Θ_1.

Lemma 4.4 - Lemma 4.6 imply that if the game has separating SE, then only states corresponding to the Riley equilibrium can be stochastically stable, since Lemma 4.4 and Lemma 4.5 ensure that Lemma 4.6 can be applied with $\Theta_1 = \Theta_r$, and $\Theta_2 = \Theta \backslash \Theta_r$. Since we already know which states are stochastically stable if the game has no separating equilibria, this proves,

Theorem 4.7. *Consider a signalling game fulfilling assumptions A1 and A2.*

(1) If the game has separating SE, and k is sufficiently large, then only behavior in accordance with the Riley equilibrium will be observed frequently in the long run.

(2) If the game does not have a separating SE, then the only behavior that will be observed frequently in the long run is that senders of type H play x_H^* all the time, while senders of type L shift between x_L^* and x_H^*.

5. Final Remarks

The main conclusion from our analysis is that for monotone signalling games with separating equilibria only behavior in accordance with the Riley equilibrium will be observed frequently in the long run if players learn and tremble as assumed in the definition of our learning process.

It is of interest to compare our selection to that obtained by refinement concepts based on imposing restrictions on out-of-equilibrium beliefs. For games that do have separating equilibria our selection is stronger. For instance, our selection in favor of the Riley equilibrium also holds in the situation well-known from economically relevant (entry deterrence) games, where there is a pooling equilibrium which is better for both types of

sender than the Riley equilibrium, and where the Intuitive Criterion, although it does select in favor of the Riley equilibrium among the separating equilibria, does not eliminate this "good" pooling equilibrium. Cho and Sobel (1990) show that for signalling games that are monotone in our sense and fulfill an additional assumption on the costliness of signalling (basically that signals can be ordered such that for given actions of receivers the high type prefers "higher" signals than the low type), the divinity criterion D1 of Cho and Kreps (1986) selects the Riley equilibrium whenever there are separating equilibria. Since the economically relevant games we know of do fulfill Cho and Sobel's "cost of signalling" assumption, their selection and ours are identical in such games (but obtained by completely different methods). However, one can construct examples of two type signalling games that are monotone and have both pooling and separating sequential equilibria for which criterion D1 does not eliminate all pooling equilibria and even examples where only pooling equilibria survive D1, see Appendix A. In these examples our learning process still selects unambiguously in favor of the Riley equilibrium. For pooling equilibria, whenever our learning process points to one (in a monotone signalling game) it has to involve pooling on the signal x_H^*, whereas, whenever criterion D1 selects a pooling equilibrium (in a monotone game also fulfilling the "cost of signalling" assumption) it has to involve pooling on the highest of all signals, see Cho and Sobel (1990).

Since for monotone signalling games criterion D1 gives the same selection as generally stronger devices such as Never Weak Best Response all the way up to Strategic Stability, Kohlberg and Mertens (1986), it follows that the equilibrium selection obtained by evolutionary learning is stronger than, and only partly in accordance with, that obtained by known rationality based selection devices

The motivation of our monotonicity assumption is mainly that it encompasses most signalling games studied for economic reasons in the literature. It is, nevertheless, of interest to investigate if monotonicity is fundamental for the selection being so strongly in favor of separating equilibria.

Indeed it is. Figure 5.1 displays a $2 \times 2 \times 2$ signalling game in extensive form.[10] Note that the game is not monotone: If the sender is of type L, he prefers action up in response to signal $left$, and action $down$ in response to signal $right$, which is ruled out by A1. Also, type L dislikes the response up after signal $right$, but up is the response that the receiver will give if he believes that the sender is type H. In fact type L dislikes this response even more than type H dislikes the response $down$, in the sense that there are samples after $left$ such that type L would play $left$ given the sample if he believes the response after $right$ to
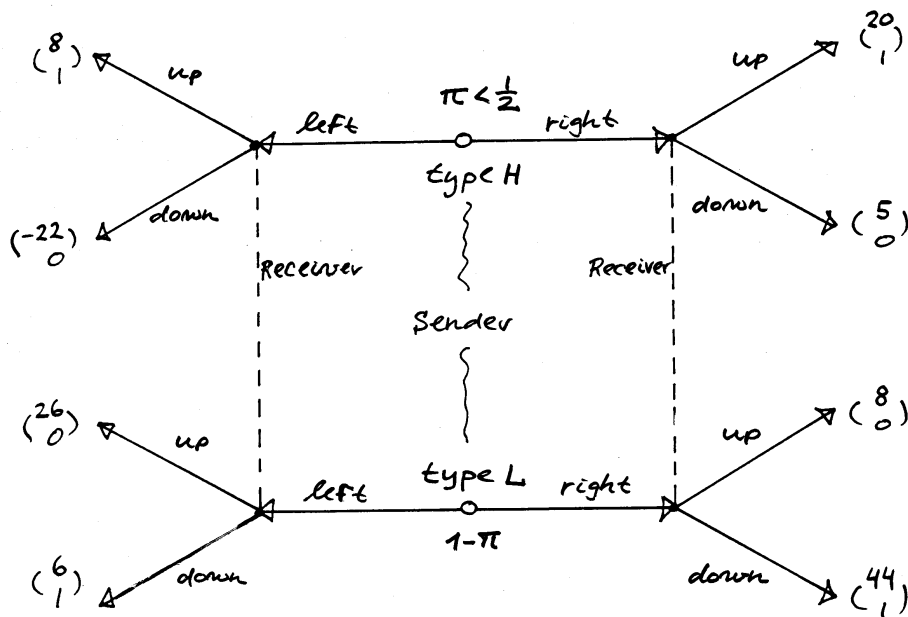
---

[10]This example is originally from Jensen (1996).

Figure 5.1:

be *up*, but type H would play *right* given the sample even if he believes that the response after *right* is *down*.

The game has two sequential equilibria in pure strategies, a separating one where the sender goes *left* if of type H (and obtains action *up* in response), and *right* if of type L (and obtains *down* in response), and a pooling one where the sender goes *right* independently of type (and both signals are responded to by *down*).

Jensen (1996) shows that if k is sufficiently large, there are stochastically stable states that correspond to the pooling equilibrium, but none that correspond to the separating equilibrium. There are absorbing sets for $\mathbf{P}^0$ corresponding to both the pooling equilibrium and the separating (Riley) equilibrium. However, when trembles are considered the separating equilibrium turns out to be more fragile than the pooling. Mutations may change the off-path beliefs in states corresponding to the pooling equilibrium such that drift in in-equilibrium beliefs (i.e., by the $\mathbf{P}^0$-process) may lead to a state where some type of sender has *left* as a best response. This happens first by changing the best response for type L, leading to a state from which - after a lot of type L senders have entered and played *left* - receivers' belief after *left* changes in the direction of L and hence the best response after *left* changes to *down*, discouraging both types of senders from playing *left* and thereby

23

receiver

		a_1	a_2	a_3
sender	L	6,3	8,2	9,0
	H	3,0	5,2	6,3

x_1

	a_1	a_2	a_3
L	1,3	3,2	5,0
H	2,0	4,2	7,3

x_2

	a_1	a_2	a_3
L	2,3	4,2	7,0
H	1,0	8,2	9,3

x_3

Figure A.1: The probability of sender being type H is $\frac{3}{4}$.

restoring the pooling equilibrium.

A. Example

Consider the signalling game with payoffs as in Figure A.1

The game has pooling equilibria where both types send signal x_3, and separating equilibria where type L sends x_1 and type H sends x_2.

In pooling equilibrium strategies and beliefs $(x(\cdot), a(\cdot), \mu(\cdot))$ fulfill:
Sender: $x(H) = x(L) = x_3$.
Beliefs: $\mu(x_1) \in [0, \frac{1}{3}], \mu(x_2) \in [0, 1]$, and $\mu(x_3) = \frac{3}{4}$.
Receiver: $a(x_1) = a_1, a(x_2) \in \{a_1, a_2, a_3\}$, and $a(x_3) = a_3$.

The separating equilibria are strategies and beliefs $(x(\cdot), a(\cdot), \mu(\cdot))$ fulfilling:
Sender: $x(H) = x_2$, and $x(L) = x_1$.
Beliefs: $\mu(x_1) = 0, \mu(x_2) = 1$, and $\mu(x_3) \in [0, \frac{1}{3}]$.
Receiver: $a(x_1) = a_1, a(x_2) = a_3$, and $a(x_3) = a_1$.

Since the game satisfies assumptions A1 and A2, Theorem 4.7 gives that all stochastically stable states of the game correspond to the unique separating equilibrium outcome.

Consider some equilibrium of the game, and let $u^*(t)$ be sender type t's payoff in the equilibrium. For each unused signal x and each type t define the following sets: $P(t|x) = \{a \in A | u^*(t) < u(t, x, a)\}$. Thus $P(t|x)$ is the set of (pure) actions from receiver that gives sender type t a higher payoff than the equilibrium payoff given sender sends signal x.

The equilibrium satisfy the D1 criterion of Cho and Kreps (1986) and Cho and Sobel (1990) if for all unused signals x and all types t, t', with $t \neq t'$, the following is satisfied:

$$\text{If } P(t|x) \subset P(t'|x) \text{ then } \mu_t(x) = 0.$$

First consider the separating equilibria. The only unused signal is x_3. It follows that $P(L|x_3) = \{a_3\}$, and $P(H|x_3) = \{a_2, a_3\}$. So for the equilibrium to satisfy D1 we must have $\mu(x_3) = 1$. But this is in contradiction to the equilibrium constraints. Thus, every separating equilibrium violates D1.

Then look at the pooling equilibria. Both x_1 and x_2 are unused. It follows that $P(H|x_1) = P(H|x_2) = P(L|x_2) = \emptyset$, and $P(L|x_1) = \{a_2, a_3\}$. So, D1 gives the restriction $\mu(x_1) = 0$. This is not in contradiction to the pooling equilibria, so we conclude that only pooling equilibria survive the D1 criterion.

If the payoff for sender in cell (H, x_3, a_2) is changed to 6 (to a number between 1 and 7), then the separating equilibria will survive D1. In this case the set $P(H|x_3) = \{a_3\}$, and therefore there is no restriction on beliefs after signal x_3. In this case both the pooling and the separating equilibria fulfill D1.

B. Proofs

Proof of Lemma 4.2

Let $\omega = (g, h)$ be a state where $g^x = \bar{g}^x$ and $h^x = \bar{h}^x$ for $x \neq x_H^*$, and (1) and (2) in Lemma 4.2 are fulfilled.

If the sender is of type H in the succeeding round, he will play x_H^*. This is best reply since, from (2), even if the worst possible sample from $g^{x_H^*}$ is drawn, the best response is x_H^*, and thus this is also true for any other sample (by A2). So after this round all records for signals different from x_H^* remain unchanged.

If the sender is of type L in the succeeding round, he will play x_L^* or x_H^*, since, from (2), x_L^* is a better response than any other signal different from x_H^*. If x_L^* is the best response, and $x_L^* \neq x_H^*$, then receiver will respond by $\underline{a}(x_L^*)$, since, from (1) any sample from $h^{x_L^*}$ will tell him that x_L^* came from a type L. Hence no records will change. If x_H^* is the best response only the records for that signal can change.

Thus, if a state fulfills the assumptions of Lemma 4.2, all its succeeding states fulfill the

same assumptions. We conclude that at least one absorbing set exists where for all states $g^x = \bar{g}^x$ and $h^x = \bar{h}^x$ for $x \neq x_H^*$.

To see that only one absorbing set exists for each possible choice of $\bar{g}^x, \bar{h}^x$, note that for any state $\omega = (g, h)$, where $g^x = \bar{g}^x$ and $h^x = \bar{h}^x$ for $x \neq x_H^*$, and (1) and (2) in Lemma 4.2 are fulfilled, there is positive probability of reaching the state where $g^x = \bar{g}^x$ and $h^x = \bar{h}^x$ for $x \neq x_H^*$, and $g^{x_H^*} = (\bar{a}(x_H^*), \ldots, \bar{a}(x_H^*))$ and $h^{x_H^*} = (H, \ldots H)$, namely by drawing type H in $2m$ rounds in a row. ∎

Proof of Proposition 4.3

For any state $\omega = (g, h)$, and any signal x, let $\tilde{u}_t(x|\omega)$ be the expected payoff to a sender type t from sending signal x, when the sample from g^x is the best possible, that is, consists of the k highest actions listed in g^x.

Lemma B.1. *For each $\omega \in \Omega$, let ω_T be a state that can be reached if the sender is of type L in all of the next T rounds, and in all of those rounds for every signal picks the best possible sample. There exists $T < \infty$, such that for every state $\omega \in \Omega$,*

(a) $\omega_T = (g_T, h_T)$ fulfills

$$g_T^{x_L^*} = (\underline{a}(x_L^*), \ldots, \underline{a}(x_L^*)), \text{ and } h_T^{x_L^*} = (L, \ldots, L), \tag{B.1}$$

and x_L^* is the best response for a sender of type L for all possible samples from g_T.

(b) $\tilde{u}_t(x|\omega) \geq \tilde{u}_t(x|\omega_T)$ for all $x \in X$, $t = H, L$, with strict inequality for all signals $x \neq x_L^*$ that were played during the T rounds.

Proof: Start in an arbitrary state ω. Let type L be drawn in each of the next $T = 2mn$ rounds, and assume that in each of these rounds the sender picks the best possible sample for each signal. If the sender is type L and has chosen the same signal x in $2m$ rounds (without any sender of type H choosing it), it must be the case that $g^x = (\underline{a}(x), \ldots, \underline{a}(x))$, and $h^x = (L, \ldots, L)$. Notice that if for some signal x, $g^x = (\underline{a}(x), \ldots, \underline{a}(x))$, signal x can be best reply for a sender type L only if $x = x_L^*$. Thus signals $x \neq x_L^*$ can be played in at most $2m$ rounds during the T rounds, and after T rounds the best response of the best possible sample (and hence for all samples) for type L will be x_L^*, and (a) will be fulfilled.

For any x that was not sent in any of the T rounds, it must be the case that the record g^x is unchanged, and then the best expected payoff is also unchanged, $\tilde{u}_t(x|\omega) = \tilde{u}_t(x|\omega_T)$.

Consider some signal $x \neq x_L^*$ that *was* sent in some round, and thus optimal (given best samples) in that round, say round τ. If the response to x in round τ were such that the best expected payoff from sending signal x increased, then x is also best in round $\tau + 1$ (given best samples), because the records for all other signals are unchanged. Signal x will remain being best (given best samples in every round) at least until the best expected payoff from sending x is strictly less than it was in round τ. This proves (b) for $x \neq x_L^*$. For x_L^* note that $\tilde{u}_t(x_L^*|\omega_T)$ is the lowest possible expected payoff from sending x_L^*, since $g_T^{x_L^*}$ contains only $\underline{a}(x_L^*)$. Hence $\tilde{u}_t(x_L^*|\cdot)$ cannot have increased from ω to ω_T. ∎

Lemma B.2. *For each $\omega \in \Omega$, let ω_Q be any state that can be reached if the sender is of type H in all of the next Q rounds, and in all of those rounds for every signal picks the best possible sample. There exists a signal x' in X, and a $Q < \infty$, such that for every state $\omega \in \Omega$,*

(a) $\omega_Q = (g_Q, h_Q)$ fulfills,

$$g_Q^{x'} = (\bar{a}(x'), \dots, \bar{a}(x')), \text{ and } h_Q^{x'} = (H, \dots, H), \tag{B.2}$$

and x' is the best response for the sender of type H for all possible samples from g_Q.

(b) $\tilde{u}_t(x|\omega) \geq \tilde{u}_t(x|\omega_Q)$ for all $x \neq x'$, $t = H, L$, with strict inequality for signals $x \neq x'$ that were played during the Q rounds.

Proof: As the proof of Lemma B.1. ∎

Note, that after the Q rounds of play, x' is best response given any sample for type H.

To conclude from Lemma B.1 and Lemma B.2, in any absorbing set of the process there must be at least one state with property (B.1) and one with property (B.2).

Proof of proposition 4.3 Let the process be in an arbitrary state $\omega \in \Omega$. Suppose in the following that the sender always picks the best possible sample for each signal. This happens with strictly positive probability.

1. Let T be as in Lemma B.1. There is a strictly positive probability, that type L is drawn in the next T rounds, and that the receiver samples all possible Hs in the first k rounds, and samples all possible Ls in the remaining rounds.

   Lemma B.1 implies that the process reaches a state, ω_T, with property (B.1), where $\tilde{u}_t(x|\omega) \geq \tilde{u}_t(x|\omega_T)$ for all signals with strict inequality for all signals different from x_L^* used during the T rounds.

2. Let Q be as in Lemma B.2. With strictly positive probability type H is drawn in the next Q rounds, and the receiver samples all possible Ls in each of these rounds.

   Lemma B.2 implies that the process then reaches a state, ω_{T+Q}, with property B.2 for some $x' \in X$, where $\tilde{u}_t(x|\omega_T) \geq \tilde{u}_t(x|\omega_{T+Q})$ with strict inequality for all used signals different from x'.

Consider the state $\omega_{T+Q} = (g_{T+Q}, h_{T+Q})$. There are two possible cases, (i) Signal x_L^* was played by type H in some round in phase 2, and (ii) Signal x_L^* was never played by type H in phase 2.

(i) Signal x_L^* was played by type H in some round in phase 2. In the first round where signal x_L^* was played by type H, the record on receivers' action after signal x_L^* was $(\underline{a}(x_L^*), ..., \underline{a}(x_L^*))$ by phase 1 and Lemma B.1. This implies that it must be the case that $x_L^* = x_H^*$, and since type H picks the best possible sample after each signal in each round it must be the case that x_L^* is the best response for all succeeding type Hs as well such that $x' = x_L^*$. In state ω_{T+Q} it must also be the case that x_L^* is the best signal for type L for any possible sample, since x_L^* was a best response for all possible samples for type L in state ω_T by phase 1 and Lemma B.1, and $\tilde{u}_L(x|\omega_T) \geq \tilde{u}_L(x|\omega_{T+Q})$ for all signals different from x_L^*, and the utility to type L of playing signal x_L^* is higher than in state ω_T, where it was at the lowest possible. Thus the process has reached an x_H^*-absorbing set.

(ii) Signal x_L^* was never played by type H in phase 2. It must then be the case that $x' \neq x_L^*$, and $g_{T+Q}^{x_L^*} = (\underline{a}(x_L^*), ..., \underline{a}(x_L^*))$, $h_{T+Q}^{x_L^*} = (L, ..., L)$, $g_{T+Q}^{x'} = (\bar{a}(x'), ..., \bar{a}(x'))$ and $h_{T+Q}^{x'} = (H, ..., H)$. Now there are two possible subcases, (a) $u(L, x_L^*, \underline{a}(x_L^*)) > u(L, x', \bar{a}(x'))$ and (b) $u(L, x_L^*, \underline{a}(x_L^*)) < u(L, x', \bar{a}(x'))$.

(a) In case $u(L, x_L^*, \underline{a}(x_L^*)) > u(L, x', \bar{a}(x'))$, state ω_{T+Q} corresponds to a separating equilibrium where the best response for all possible samples (including the best samples) of sender type L is x_L^*, and of Sender type H is x'. This state is absorbing by Lemma 4.1.

(b) In case $u(L, x_L^*, \underline{a}(x_L^*)) < u(L, x', \bar{a}(x'))$ the best response for all possible samples for both types of sender in state ω_{T+Q} is x'. Then with positive probability phase 1 and phase 2 are repeated once more leading to a state $\omega_{2(T+Q)}$.

As long as the sender in phase 1 samples only $\bar{a}(x_L')$'s after x_L', then it is best response for the sender type L to play x'. During the first k rounds the receiver samples only Hs in $h^{x'}$, and plays $\bar{a}(x_L')$. For the remaining of phase 1 the receiver samples only Ls in $h^{x'}$, and plays $\underline{a}(x')$. After the first m rounds, the sender type L samples some combination of $\bar{a}(x')$ and $\underline{a}(x')$ after x', and at some point the best response for a sender of type L

changes to x_L^*, which is played in the rest of phase 1. Thus after phase 1 a state $\omega_{T+Q+T} = (g_{T+Q+T}, h_{T+Q+T})$ is reached, where $g_{T+Q+T}^{x'} = (\bar{a}(x'), \ldots, \bar{a}(x'), \underline{a}(x'), \ldots, \underline{a}(x'))$, with at most $k-1$ elements $\bar{a}(x')$, and $h_{T+Q+T}^{x'} = (L, \ldots, L)$. Note that for a sender type L, x_L^* is a better response than x', if the record $g^{x'}$ remains unchanged or more $\underline{a}(x')$ is added to it.

If x_L^* is ever played by type H in phase 2, then as in case (i) this implies $x_L^* = x_H^*$ and the process has reached an absorbing set corresponding to Lemma 4.2.

Since the receiver draws all possible Ls in all other rounds than the first k rounds where H is drawn, it must be the case if x' is played for $k-1$ more rounds, then record after x' will be $g^{x'} = (\underline{a}(x'), \ldots, \underline{a}(x'))$, so if x' is being played by type H for more than $k-1$, then it must be case that $x' = x_H^*$, and the process has reached a state corresponding to an absorbing set as described in Lemma 4.2. If this is not the case, then by Lemma B.2 there is some signal $x'' \notin \{x_L^*, x'\}$, such that state $\omega_{2(T+Q)} = (g_{2(T+Q)}, h_{2(T+Q)})$ fulfills $g_{2(T+Q)}^{x''} = (\bar{a}(x''), \ldots, \bar{a}(x''))$ and $h_{2(T+Q)}^{x''} = (H, \ldots, H)$. As case (ii(a)) above if $u(L, x_L^*, \underline{a}(x_L^*)) > u(L, x'', \bar{a}(x''))$, then the process has now reached an absorbing state corresponding to a separating equilibrium. If $u(L, x_L^*, \underline{a}(x_L^*)) < u(L, x'', \bar{a}(x''))$, then the best response for both types of sender in state $\omega_{2(T+Q)}$ is now x''. Then with positive probability phase 1 and 2 are repeated once more leading to a state $\omega_{3(T+Q)}$. By repeating the above arguments, it is seen that if x' or x'' is ever played more than $k-1$ rounds, then the process has reached an absorbing set as described in Lemma 4.2. If not, then state $\omega_{3(T+Q)}$ either (case(i)) corresponds to pooling on signal $x_L^* = x_H^*$, or (case(ii(a))) to a separating equilibrium or (case(ii(b))) there is some signal $x''' \notin \{x_L^*, x', x''\}$, such that $\omega_{3(T+Q)} = (g_{2(T+Q)}, h_{2(T+Q)})$ fulfills $g_{3(T+Q)}^{x'''} = (\bar{a}(x'''), \ldots, \bar{a}(x'''))$ and $h_{3(T+Q)}^{x'''} = (H, \ldots, H)$, and x''' is best response for a sender of type H.

By repeating this argument at most p times it follows that $\omega_{p(T+Q)}$ must be in one of the absorbing sets corresponding either to Lemma 4.1 or to Lemma 4.2. Since there cannot be some $x^{(p)} \notin \{x_L^*, x', x'', \ldots, x^{(p-1)}\}$, so case (ii(b)) cannot occur after the p'th repetition of phases 1 and 2. ∎

Proof of Theorem 4.7

In this section we prove Lemma 4.4 - Lemma 4.6, and then Theorem 4.7 follows as explained in the text.

Proof of Lemma 4.4 Consider first an absorbing state $\omega = (g, h)$ corresponding to a separating SE where a sender type H's best response is x' different from r. Then consider

one tremble by a sender of type H sending signal r, and assume that the receiver picks the best possible sample after x'. If now the best response of the sender is still x', even if he picks the best possible sample after r, then the process has reached a new absorbing state corresponding to the separating equilibrium where type H plays x'.[11] If there are samples such that the best response of a sender type H is r, assume that as long as this is the case each sender of type H picks such a sample and plays r. Then either the process will reach a Riley state, or it will again reach a state corresponding to the separating equilibrium where type H plays x'.

If the process after the first tremble has reached a state corresponding to the separating equilibrium where type H plays x', then assume that one more tremble by type H playing r takes place. Again either the process will reach a new absorbing state corresponding to the separating equilibrium where type H plays x', or it will reach a Riley state. Continuing this, eventually the best response of the receiver after signal r becomes $\bar{a}(r)$, and also the sender's belief become $g_r = (\bar{a}(r), ..., \bar{a}(r))$, and the process has reached a Riley state.

So, there is a chain of absorbing states corresponding to the same separating equilibrium connected by just one tremble, and such that at the end of the chain only one tremble is required to reach a Riley state.

Then consider a state belonging to an absorbing set as in Lemma 4.2. By using phase 1 and step 2 of the proof of Proposition 4.3 above, one can verify that with strictly positive probability the process reaches a state where $g^{x^*_H} = (\underline{a}(x^*_H), \ldots, \underline{a}(x^*_H))$. Now use the above argument, considering one tremble by sender type H playing r, and thereby changing the record at signal r and thus leading to a new absorbing state corresponding to Lemma 4.2 or to an absorbing state corresponding to the Riley outcome. ∎

Proof of Lemma 4.5 Assume that the process is in an arbitrary Riley state ω. We need to show that the unperturbed process after only one tremble by either the sender or the receiver always leads back to a Riley state. Suppose first a sender type L made one tremble and chose a signal that is not best response, i.e. a signal $x \neq x^*_L$. Three cases must be considered:

(i) $x = r$. By assumption A2, this does not change the best responses for any player, any

---

[11]It is different from the orginal state, since either it was the case that ω had $h^r \neq (H, H, ..., H)$ and then one new observation of a type H changed h^r, or it was the case that ω had $h^r = (H, ..., H)$ and $g^r \neq (\bar{a}(r), ..., \bar{a}(r))$ and then the receiver sampled only Hs and played $\bar{a}(r)$, which changed g^r. It is not possible that ω had $g^r = (\bar{a}(r), ..., \bar{a}(r))$, since then x' would not have been best response for the sender of type H.

sample, given k is sufficiently large, such that the best response of a receiver to a sample consisting of $k-1$ observations of type H and 1 observation of type L, is the same as the best response to a sample consisting of only type H. So, if no more trembles occur, then in all future rounds the receiver will play $\bar{a}(r)$ after signal r, and with probability one the process reaches the same Riley state as before the tremble.

(ii) $x \neq r$, $x \in X^S$. In this case the tremble only changes beliefs after x, and this cannot change the best response of a sender type H, but it may change the best response of a sender type L to be x for some samples. If it does change the best response for a sender type L, then with probability one some senders of type L will play x, which will eventually lead to a new state in which the receiver takes some low action after x, such that it is no longer best response for type L to signal x.

(iii) $x \notin X^S$. In this case the best response may be changed for one or both types of sender after the tremble. By Proposition 4.3, if no more trembles occur, the process will with probability one reach an absorbing set as characterized by either Lemma 4.1 or Lemma 4.2. The records g^r and h^r are unchanged by the tremble, and will not be changed after the tremble since it is never best response for type L to play r. However, when $g^r = (\bar{a}(r), ..., \bar{a}(r))$ and r is a separating signal, condition (2) of Lemma 4.2 cannot be fulfilled, so the absorbing set can not be an x_H^*-absorbing set. Further, when $g^r = (\bar{a}(r), ..., \bar{a}(r))$ it will never be best response for a sender of type H to send a separating signal other than r. So, the process must reach a Riley state.

Next suppose the tremble was made by a sender of type H choosing a signal x different from r. Again three cases must be considered:

(iv) $x = x_L^*$. As in case (i) above this will not change any best responses provided k is sufficiently large, so the process will with probability one get back to the original Riley state.

(v) $x \neq r$, $x \in X^S$. This is similar to case (ii) above.

(vi) $x \notin X^S$. This is similar to case (iii) above.

Finally, consider trembles by the receiver. Those can occur only following the signals sent in equilibrium, x_L^* and r, since the sender does not tremble. As in cases (i) and (iv) above, if k is sufficiently large, then one such tremble is not sufficient to change the best response of any type of sender, so with probability one the process will get back to the original Riley state. ∎

Proof of Lemma 4.6 Assume there exists $O^* \in \Theta_2$ with minimal total resistance. Let G be the minimal O^* - tree. By assumption there exists a path, Δ_1, from O^* to O' with resistance 1 for each transition and where $O' \in \Theta_1$. Denote by R_1 the set of sets along Δ_1.

Furthermore there is a branch in G from O' to O^*. Let O'' be the last set in Θ_1 first time along this path there is transition from a state in Θ_1 to a state in Θ_2. Let R_2 be the sets along the path, Δ_2, from O' to O''.

We can now construct a O''-tree, for all sets in $(R_1 \cup R_2)^c$ just use the same transitions as in G. For sets in $R_1 \backslash O'$ use the path Δ_1 and for sets in $R_2 \backslash O''$ use the path Δ_2.

Thus for all sets different from O^* and O'' the resistance is less or equal as the one in G, and since the edge going out of O'' (in G) has resistance at least 2 and the edge going out of O^* has resistance 1 the total resistance for the O''-tree is strictly less than the total resistance for G, which is a contradiction and the result follows. ∎

The theorem also holds without the assumption that the receiver always takes a rational action (even if he trembles). Omitting the assumption many more absorbing sets will emerge. In each of those new sets at least one record will contain a non-rational action from the receiver, say $g^{x'}$. However one tremble at the time from the sender sending x' (without any trembles from the receiver) will lead the process to an absorbing set where $g^{x'}$ only consists of rational actions, and thus from any of these new absorbing sets, there exists a chain of one-trembles by the sender such that the process ends up in an absorbing set fulfilling the assumption. The theorem now follows from Lemma 4.4-4.6.

References

[1] Banks, J., and J. Sobel (1987), Equilibrium selection in signalling games, *Econometrica*, **55**, 647-662.

[2] Bergin, J., and B.L. Lipman (1996), Evolution with State-Dependent Mutations, *Econometrica*, **64**, 943-956.

[3] Billingsley, P. (1995), *Probability and Measure,* (3.ed.), Wiley.

[4] Cho, I.K. and D.M. Kreps (1987), Signalling games and stable equilibria, *Quarterly Journal of Economics,* **102**, 179-221.

[5] Cho, I.K. and J. Sobel (1990), Strategic Stability and Uniqueness in Signaling Games, *Journal of Economic Theory* **50**, 381-413.

[6] Canning, D. (1992a), Average Behavior in Learning Models, *Journal of Economic Theory* **57**, 442-72.

[7] Canning, D. (1992b), Learning, Language Conventions in Common Interest Signaling Games, Discussion Paper # 607, Columbia University.

[8] Freidlin, D. and A. Wenzell (1984): Random perturbations of dynamical systems. *New York: Springer Verlag.*

[9] Fudenberg, D. and J. Tirole (1991), *Game Theory,* MIT Press: Cambridge, Mass. and London.

[10] Hellwig, M. (1987), Some Recent Developments in the Theory of Competition in Markets with Adverse Selection, *European Economic Review*, **31**, 319-25

[11] Jacobsen, H.J., M. Jensen and B. Sloth (1998), On the Evolution of Conventions under Incomplete Information, mimeo, University of Copenhagen.

[12] Jacobsen, H.J. and B. Sloth (1998), Evolutionary learning in simple signalling games: equilibrium selection by drift in-equilibrium beliefs, mimeo, University of Copenhagen.

[13] Jensen, M. (1996), Udvælgelse af Nash-ligevægte via stokastisk learning og evolution, Master thesis, Institute of Mathematics, University of Copenhagen (in Danish).

[14] Kandori, M., G.J. Mailath and R. Rob (1993), Learning, Mutation, and Long Run Equilibria in Games, *Econometrica*, **61**, 29-56. (KMR)

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

[16] Milgrom, P. and J. Roberts (1982), Limit pricing and entry under incomplete information: An equilibrium analysis, *Econometrica*, **50**, 443-459.

[17] Nöldeke, G. and L. Samuelson (1997), A dynamic model of equilibrium selection in signalling markets, *Journal of Economic Theory*, **73**, 118-156.

[18] Young, H.P. (1993a), The Evolution of Conventions, *Econometrica*, **61**, 57-84.

[19] Young, H.P. (1993b), An Evolutionary Model of Bargaining, *Journal of Economic Theory* **59**, 145-68.