DISCUSSION PAPERS
Institute of Economics
University of Copenhagen

Papers by Participants in
The Birgit Grodal Symposium, Copenhagen, September 2002

03-12

The Evolution of Conventions under Incomplete Information

Mogens Jensen, Birgitte Sloth and Hans Jørgen Whitta-Jacobsen

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

Birgit Grodal Symposium

Topics in Mathematical Economics

The participants in a September 2002 Workshop on *Topics in Mathematical Economics* in honor of Birgit Grodal decided to have a series of papers appear on Birgit Grodal's 60'th birthday, June 24, 2003.

The Institute of Economics suggested that the papers became Discussion Papers from the Institute.

The editor of *Economic Theory* offered to consider the papers for a special Festschrift issue of the journal with Karl Vind as Guest Editor.

This paper is one of the many papers sent to the Discussion Paper series.

Most of these papers will later also be published in a special issue of *Economic Theory*.

Tillykke Birgit

Troels Østergaard Sørensen	Karl Vind
Head of Institute	Guest Editor

The Evolution of Conventions under Incomplete Information

Mogens Jensen, Birgitte Sloth, and Hans Jørgen Whitta-Jacobsen,

Institute of Economics
University of Copenhagen
Studiestraede 6
DK-1455 Copenhagen K
DENMARK

This version: March 2003

Abstract

We formulate an evolutionary learning process with trembles for static games of incomplete information. For many games, if the amount of trembling is small, play will be in accordance with the games' (strict) Bayesian equilibria most of the time supporting the notion of Bayesian equilibrium. Often the process will select a specific equilibrium. We study an extension to incomplete information of the prototype conflict known as "Chicken" and find that the equilibrium selection by evolutionary learning may well be in favor of inefficient Bayesian equilibria where some types of players fail to coordinate.

JEL Classification: C72

1 Introduction

This paper suggests an evolutionary learning process in the spirit of Young ([7]) for static games of incomplete information and demonstrates in general how such a process may give justification for the notion of Bayesian equilibrium and may give selection among multiple Bayesian equilibria. For a specific game of economic interest, an extension to incomplete information of the prototype strategic conflict known as the "Chicken" game, the selection is characterized. Although the game is special, a general insight emerges: equilibrium selection by evolutionary learning cannot in general be expected to be in accordance with the more efficient equilibria where the different types of the players coordinate well.

There are mainly two motivations for extending the models of evolutionary learning to incomplete information. First, the foundations of Bayesian equilibrium are at least as shaky as those of Nash equilibrium. Any doubt one may have concerning the feature that players use best replies against each other is as relevant for Bayesian equilibrium in games of incomplete information as it is for Nash equilibrium in games of complete information. For Bayesian equilibrium one may further doubt if the idea that players plan for types they are actually not, is an adequate formalization of how players cope with uncertainty in games. It is therefore of interest if one can give a justification of Bayesian equilibrium for games of incomplete information, like the one given of Nash equilibrium for games of complete information by, e.g., [7]. Second, going from complete to incomplete information in games often adds a dimension of equilibrium multiplicity - in particular it often implies the existence of inefficient equilibria - such that equilibrium selection may be even more relevant in games of incomplete information. Since the evolutionary models with small trembles suggested by, e.g., [5] and [7], have proved to be strong devices for equilibrium selection, it is a natural idea to generalize such processes to games of incomplete information.

To define an evolutionary learning process connected to a game of incomplete information, one has to give a physical meaning to the types and priors that are part of the description of the game. We assume that there are two large pools of players containing players to take the "row" and the "column" position in the game respectively, and that each pool is subdivided in types. The existing types are as in the underlying game: if a player of a specific type from one pool plays against a player of a specific type from the other, the two will receive payoffs depending on their types and actions exactly as in the underlying game. Play goes on in subsequent rounds. In each round, one player is picked at random from each pool and these two play the game once, and their actions are observed by everybody, just as assumed by Young [7]. The two crucial modelling choices that have to be faced in an extension of Young's approach to games of incomplete information concern the degrees of knowledge of the priors and of observability of types. We consider assumptions at the most "well-informed" extreme: We assume: (1) that all players know the true probability distribution by which opponents are picked, and (2) that after each round of play the types of the picked players become known by everybody. We motivate these choices below.

For each type of each player there is a record of the actions a player of that type took in a certain number of earlier rounds where this type was picked for play. After a round of play the records on the two types who played are updated; the oldest observation is deleted and the new one is inserted. The records on other types are unchanged. When a player is about to play, he intends to play a best reply to the expectation on the opponent created from samples from the current records on the opponent. This defines the basic learning process. The perturbed learning process is defined from the basic one by adding a tremble: with a small probability ε a picked player does not (necessarily) play a best response, but takes an arbitrary action. The interest is in the perturbed process for ε small.

The considered game and the process may well be such that when the probability of trembling is small play will be in accordance with the games's strict Bayesian equilibria most of the time. For such games one has thus obtained the kind of support for the notion of Bayesian equilibrium looked for. Further, when a game has several strict Bayesian equilibria, it may well be that play will be in accordance with a specific one most of the time. A selection among the Bayesian equilibria is thus obtained. We study such selection in the Chicken game of incomplete information.

The game is an extension to incomplete information of the kind of coordination game typically studied in the earlier contributions on evolutionary learning in games of complete information and illustrates how incomplete information may add an economically important dimension of equilibrium multiplicity. Under complete information the game has multiple strict Nash equilibria, but these have similar efficiency properties: both pure Nash equilibria involve one player taking the tough action and the other taking the cautious one, and equilibrium selection is only a matter of who takes which action, i.e., who gets the (main part of the) surplus. Under incomplete information the game still has two equilibria where (independently of types), one side takes the tough action, and the other side takes the cautious one, but in addition there is an equilibrium where sometimes (for some combinations of types), both players take the tough action, and a waste occurs. The inefficiency possible in equilibrium is thus due to incomplete information. If there is considerable uncertainty about payoffs, equilibrium selection by evolutionary learning in the incomplete information Chicken turns out to be in favor of the less efficient equilibrium, where sometimes tough is played agains tough.

The present paper is closely related to contributions such as [1], [5], and [7], that introduced the approach of evolution or evolutionary learning with trembles for static games of complete information. Some papers have studied equilibrium selection by evolutionary models for dynamic games with both incomplete information and sequential moves, most notably for signalling games, [2], [4], and [6]. The present paper is (as far as we know) the first to study an evolutionary process in the same spirit for static games of incomplete information.

Of the modelling choices (1) and (2) above, the assumption of observability of types is the more fundamental and controversial one (if types can be observed then close to correct priors can be derived from long records). In favor of (2) speaks the following: First, if the player pools are large, each individual only plays rarely, so no individual has incentive to hide his type after having played. Second, for the formulation of a theory that is general enough to include games for which the type of the opponent can be inferred from knowledge of the actions and one's own type and payoff, it has to be assumed that types become known after each round of play. Third, to preserve from Young's work the essential structural equivalence between equilibria and particular states (so-called conventions), records have to list actions type-wise, since a Bayesian equilibrium reports an action for each type of each player. Records of past play can only be type-wise if types become known. Finally, assuming that only the opponent's actions can be observed, and consequently that records only list past actions, will imply a certain "drift" phenomenon, which has no counterpart in Young's work: Occasionally a certain type of a player will, by accident, be picked in a high concentration over some rounds, and for that reason a specific action may be observed much. This will shift the expectation concerning the player towards the much observed action, which may change the best response (of some types) of the other player, a shift that would not have occurred if it had been known that the high concentration of the particular action was only caused by a particular type coming out many times. In this way, even if play were currently in accordance with an equilibrium, behavior would be able to drift out of equilibrium (without trembles), whereas in Young's work, and for the process we formulate, if play is currently locked at a convention (corresponding to an equilibrium), play cannot drift away, but only change by trembling. So, a straight extension of Young's work to incomplete information implies observability of types. Of course, an alternative process where only actions could be observed would also be of interest.[1]

In Section 2 we give the definitions of games of incomplete information and Bayesian equilibrium. Section 3 defines the evolutionary learning process in two steps, the basic and the perturbed process, respectively. Further we state some general results about convergence and stochastically stable states. In Section 4 we analyze the incomplete information version of the Chicken game and characterize the long run behavior supported by evolutionary learning. Section 5 concludes. Proofs are given in Appendix.

2 Games of Incomplete Information and Bayesian Equilibrium

We describe a finite static two player game of incomplete information as follows. The Row player, Player 1, has finite action set R, and the Column player, Player 2, has finite action set C. Player 1 is of one type α out of the finitely many in A, while Player 2 is of one type β out of the finitely many in B.[2] Each player

[1] In Jacobsen et. al. (2001) we have studied such a process for the specific case of signaling games.

[2] For everything below we could, more generally, have chosen a formulation where the set of available strategies possibly depended on type, $R(\alpha)$ and $C(\beta)$ etc. This could be of some relevance in the context of evolutionary learning with trembles, see Section 3 below, but would imply a heavier notation.

knows his own type, but not the opponent's. Player 1's belief concerning Player 2's type is given by a probability measure b over B. Likewise, Player 2's belief concerning Player 1's type is given by a probability measure a over A. If players 1 and 2 of types α and β choose actions $r \in R$ and $c \in C$ respectively, they obtain von Neumann-Morgenstern payoffs $u(r, c, \alpha, \beta)$ and $v(r, c, \alpha, \beta)$.

A probabilistic expectation, or conjecture, of Player 1, concerning Player 2's choice, is a collection $q = (q_\beta)$ of probability measures over C, one for each of Player 2's possible types. Likewise, an expectation of player 2 is $p = (p_\alpha)$, where each p_α is a probability measure over R. The expected payoff of a Player 1 of type α, who holds a conjecture q, from choosing the specific action r, is,

$$U_\alpha(r, q) = \sum_{\beta \in B} b(\beta) \sum_{c \in C} u(r, c, \alpha, \beta) q_\beta(c),$$

and r is a best reply if it maximizes $U_\alpha(r', q)$ over $r' \in R$. Let the set of pure best replies be $BR_\alpha(q)$. Define the expected payoff of Player 2, $V_\beta(c, p)$, similarly and let the set of best replies for a Player 2 of type β holding a conjecture p be $BR_\beta(p)$. The sets $BR_\alpha(q)$ and $BR_\beta(p)$ are always non-empty.

A Bayesian equilibrium (in pure actions) is a pair $r(\cdot), c(\cdot)$, where $r(\alpha) \in R$ for all $\alpha \in A$, and $c(\beta) \in C$ for all $\beta \in B$, such that if one for each α defines the probability measure p_α by $p_\alpha(r(\alpha)) = 1$, and for each β defines q_β by $q_\beta(c(\beta)) = 1$, then $r(\alpha) \in BR_\alpha(q)$ for all $\alpha \in A$, and $c(\beta) \in BR_\beta(p)$ for all $\beta \in B$. A Bayesian equilibrium is strict if for all the α's and β's, $BR_\alpha(q)$ and $BR_\beta(p)$ are singletons. We only consider games with at least one strict Bayesian equilibrium.

3 Evolutionary Learning in Games of Incomplete Information

A Bayesian equilibrium is a Nash equilibrium of the extended game where the players' pure strategies are mappings $r : A \longrightarrow R$, and $c : B \longrightarrow C$, and payoffs associated to such strategies are the expected values of payoffs from the original incomplete information game, where expectations are taken with respect to types. However, the evolutionary learning process we define in this section is not just an application of the process of Young to this extended game. Applying Young's process directly to the extended game would imply an assumption that in each round the full strategy of the opponent, and not just the action taken by the relevant type of opponent, is observed, and this is not meaningful. Hence, the process we define is not a special case of Young's. On the other hand, Young's process is the special case of the process defined here where the game under consideration has only one type for each player.

The presence of incomplete information necessitates some modelling choices which do not have counterparts in the work of Young. We face these choices in a general formulation.

3.1 The Basic Learning Process

Envisage that there are two large (disjoint) pools of player 1s and 2s, both pools partitioned according to types. Players from the two pools play a game between them in subsequent rounds. In each round one player 1 is picked at random from pool 1, such that type α has probability $a(\alpha)$, and one player 2 is picked randomly from pool 2 with probability $b(\beta)$ of type β. The picked player 1 chooses an action r from R, while the picked player 2 chooses an action c from C. They receive payoffs according to their choices r and c, and their types α and β, as given by the underlying game, that is, $u(r, c, \alpha, \beta)$ to player 1, and $v(r, c, \alpha, \beta)$ to player 2.

As discussed in the introduction above, we make two important informational assumptions. **(1)** Players know the true probability measure, a or b, by which their opponent is picked. **(2)** After a round of play the actions r and c chosen in the round, as well as the true types α and β of the players who took them, become known to everybody.

The individuals keep records of past play. For each type α or β, a record h_α or h_β reports which actions were taken the last m_α or m_β times a player of that type played, $h_\alpha \in R^{m_\alpha}$ and $h_\beta \in C^{m_\beta}$. A state h is a complete description of the records, $h = ((h_\alpha), (h_\beta)) \in \prod_{\alpha \in A} R^{m_\alpha} \times \prod_{\beta \in B} C^{m_\beta}$. The state space is thus finite. After a round of play where a type α of player 1 chose r, against a type β of player 2 who chose c, only h_α and h_β are updated, and this is done by deleting in each of them the oldest observation and inserting as the newest observation r and c respectively. Given a state h, a state h' is a successor to h, if it is possible to go from h to h' in one step according to this procedure by picking α and β, and r and c appropriately.

When a player 1 has been picked, he first samples from the records on player 2;

for each β he takes a sample Q_β from h_β, where the sample size $k_\beta = \#Q_\beta$ fulfils $k_\beta \leq m_\beta$. The sampling goes on according to a random procedure, which is such that all observations in h_β have positive probability of ending in Q_β. Let Q be the collection of samples, $Q = (Q_\beta)$. A picked player 2 samples P_α from h_α, where again the sample size is $k_\alpha \leq m_\alpha$, and the set of samples is $P = (P_\alpha)$. Samples P_α and Q_β are converted into probability measures p_α and q_β over R and C respectively the obvious way: $p_\alpha(r)$ is the number of times r appears in the sample P_α divided by k_α, etc. It will cause no confusion to identify the samples P and Q with the so derived collections of probability measures p and q.

According to the basic learning process, a player 1 picked for play will take an action in $BR_\alpha(Q)$, if he is of type α. If $BR_\alpha(Q)$ has several elements, player 1 will pick one at random according to a full support probability measure on $BR_\alpha(Q)$.[3] Similarly a picked player 2 of type β will take an action in $BR_\beta(P)$.

Given the random procedures by which players are picked, sampling goes on, and ties are broken, there will for each pair of states h and h' be a specific probability of going from h to h' in one step. Call this transition probability $\pi^0(h, h')$. If h' is not a successor of h, then $\pi^0(h, h') = 0$. The matrix of all transition probabilities is Π^0, which defines a homogeneous Markov chain on H.

A convention is defined as a state h with two properties: First, it consists entirely of constant records, that is, for each α, the record h_α is a list of m_α identical actions $r(\alpha)$, and for each β, h_β is a list of m_β identical actions $c(\beta)$. Second, each recorded action is the unique best reply to the only samples that are possible from h: If one for each α lets P_α be the list of k_α times $r(\alpha)$, and for each β lets Q_β be the list of k_β times $c(\beta)$, then $BR_\alpha(Q) = \{r(\alpha)\}$ for each α, and $BR_\beta(P) = \{c(\beta)\}$ for each β.

A convention is an absorbing state for Π^0, that is, a state that one stays in for sure when transitions are governed by Π^0. Since we have assumed that in

[3] It could well be argued that only best responses, which are not weakly dominated, should be given positive probability. According to the "full" (below the "perturbed") learning process, unexpected actions will occur frequently due to trembling, and it is therefore dangerous to play dominated actions. With the alternative assumption it would not be exactly the strict equilibria that would correspond to the "resting points" (the conventions below) of the basic learning process, but rather all the "semi-strict" equilibria (where it is not required that that every best response is unique, but only that no type of no player has another best response that is undominated). We prefer, nevertheless, the formulation in the text to have full correspondence to Young's formulation on issues that are not related to information being compete or incomplete.

case of several best replies all have positive probability, it is also true that every absorbing state is a convention. Further, for any convention there is a strict Bayesian equilibrium defined in the obvious way, and every strict Bayesian equilibrium defines exactly one convention. Since we have assumed that there is a strict Bayesian equilibrium, there is at least one convention.

Assume that from any state, there is, according to Π^0, positive probability of reaching a convention in a finite number of steps. From a standard argument there is then probability one of finally reaching a convention irrespectively of initial state,[4] and we say that the basic learning process converges to a convention. One can formulate assumptions on the considered game and the details of the learning process that ensure convergence. The following is a straightforward extension of Proposition 1 in [7]. First define the best reply graph: Each node s is a combination $(r(\alpha), c(\beta))$ of actions, one for each type of each player. There is a directed edge from s to s' if and only if s' is different from s for exactly one type of one player, and for this type the action in s' is the unique best reply to the action combination of the opponent in s. The game is weakly acyclic if, from any node, there is a directed path (through directed edges) to some node out of which there is no edge (a sink). Every sink is a strict Bayesian equilibrium. For each node s, let $L(s)$ be the length of the shortest path to a sink. Let L be the maximum of $L(s)$ over all nodes. Finally, let k be the largest sample size, $k = \max_{\alpha,\beta}(k_\alpha, k_\beta)$, and let m be the shortest record length, $m = \min_{\alpha,\beta}(m_\alpha, m_b)$.

Proposition 0. *If the game is weakly acyclic and for each type γ: $k_\gamma \leq \frac{m_\gamma}{L+2}$, then the basic learning process converges to a convention.*

The proof is similar to the proof of Proposition 1 in [7], where it is demonstrated that in the one type model it is sufficient for convergence that the game is weakly acyclic and that for each player $k \leq \frac{m}{L+2}$. Young's proof is constructive, providing a finite sequence of samples that has positive probability and that implies that a convention is reached. In our model only one type of each player plays in each round

[4] If there is, starting from any state, at least probability π of reaching a convention in s steps, then there is at most probability $1-\pi$ of not reaching a convention in s steps, and then there is at most probability $(1-\pi)^t$ of not reaching a convention in st steps. Here $(1-\pi)^t$ goes to zero as t goes to infinity. So, independently of initial state it has probability zero to not eventually reach a convention, or probability one to reach one.

and a similar sequence of types and samples can be constructed. Since sampeling and updating is done for each type of each player separately, it is sufficient that the condition on k_γ is fulfilled typewise.

Sometimes it is possible to give a direct proof of convergence not referring to Proposition 0, as for the bilateral monopoly game studied in Section 5. Convergence implies that there are no other absorbing sets than the singleton sets of conventions. It then follows from a standard result on Markov chains that a stationary distribution for Π^0, a probability measure μ over H, such that $\mu\Pi^0 = \mu$, can only have $\mu(h) > 0$ for conventions h. The reverse, that any distribution with $\mu(h) > 0$ only for conventions h is stationary, is obvious.

3.2 The Perturbed Learning Process

The basic view in the literature on evolutionary learning is that a slightly perturbed process involving trembles is more realistic than the basic one described above. The basic process is modified by assuming that there is in each round a small probability ε (independent across rounds and players) that a picked player will not play a best reply to his samples, but take a random action according to a specific full support distribution (which is independent of ε) over his full strategy set. The remaining probability, $1-\varepsilon$, is assigned to the best replies as before. This means that in each round, any "suboptimal" action has a probability proportional to ε of being played. The described trembling can be interpreted as mistakes or experiments or "mutations". The transition probabilities of the modified process are called $\pi^\varepsilon(h, h')$, and the matrix of transition probabilities, the perturbed Markov chain, is called Π^ε.[5]

The process Π^ε is irreducible: for any pair of states h and h', there is according to Π^ε positive probability of going from h to h' in a finite number of steps; this is just a matter of picking types and actions (which now all have positive probability)

[5] It can be argued, in particular if trembles have the interpretation of experiments, that in the ε-eventuality of a tremble only all undominated (not all) actions should have positive probability. This could be captured by excluding from the beginning dominated actions from the strategy sets, which could well imply that the sets of available actions would depend on type, $R(\alpha)$ and $C(\beta)$, as already discussed. Here we have chosen the notationally simpler formulation in the text. We note, however, first that all of the general results reported in this section would be identical in a formulation with type dependent strategy sets as long as it is assumed that a tremble for a type α could go to any strategy in $R(\alpha)$, etc. Second, the equilibrium selection obtained for the specific game studied below does not depend on trembling to dominated actions.

appropriately. Further, Π^ε is aperiodic: this follows since Π^ε is irreducible and there is a state (namely a convention) h with $\pi^\varepsilon(h,h) > 0$. Finally, Π^ε is a regular perturbation of Π^0 in the sense of [7], i.e., Π^ε is irreducible and aperiodic, $\Pi^\varepsilon \longrightarrow \Pi^0$ as $\varepsilon \to 0$, and for any transition hh' for which $\pi^0(h,h') = 0$, there is a well defined order of the speed by which $\pi^\varepsilon(h,h')$ goes to zero with ε.

The resistance in a transition hh' is defined as this order: If $\pi^0(h,h') = 0$, and h' is not a successor of h, so even with trembles it is impossible to go from h to h', then also $\pi^\varepsilon(h,h') = 0$, and $\pi^\varepsilon(h,h')$ could be said to go to zero infinitely fast, so the resistance in hh' is infinite. If $\pi^0(h,h') = 0$, but one can go from h to h' if and only if one type of each player makes a tremble, so two trembles are necessary, then $\pi^\varepsilon(h,h')$ is some constant times ε^2, which goes to zero with a speed of order two, and the resistance in hh' is two. If $\pi^0(h,h') = 0$, but it takes just one tremble (by one type of one player) to go from h to h', then $\pi^\varepsilon(h,h')$ is a constant times ε, and the resistance is one. Finally, if $\pi^0(h,h') > 0$, the resistance is zero. To find the resistance in a transition from one state to another is just a matter of counting the trembles necessary to go from the first to the second.

It is standard from the theory of Markov chains that for each $\varepsilon > 0$, there is a unique stationary distribution μ^ε, $\mu^\varepsilon\Pi^\varepsilon = \mu^\varepsilon$, and this has the empirical content, that if one lets the process run for a long time according to the probabilities of Π^ε, then the relative frequencies by which the states are visited converge to the probabilities of μ^ε with probability one. Our interest is in Π^ε for small ε, and therefore in $\mu^0 \equiv \lim_{\varepsilon \longrightarrow 0} \mu^\varepsilon$. It follows by [7], Lemma 1, that this limit distribution exists, and that it is a stationary distribution for Π^0. The states in the support of μ^0 are called stochastically stable, and these are the states that will be observed frequently for small ε. Since μ^0 is stationary for Π^0, *if the basic learning process converges, then only the conventions of* Π^0 *can be stochastically stable.*

This has two important implications. In the long run observed play will most of the time be in accordance with the game's strict Bayesian equilibria. This supports the notion of Bayesian equilibrium, in so far as the game and the details of the process are such that the basic learning process is convergent. Further, it will often be the case that only one convention is, or only a few similar conventions are, stochastically stable, so only play according to, or very similar to, a specific of the game's strict Bayesian equilibria will be observed frequently. This provides a selection among the Bayesian equilibria.

To actually find the selected equilibria requires the use of a (quite complicated) theorem characterizing the stochastically stable states. The following is a result (Corollary to Proposition 2) from [7], specialized to the situation where all absorbing sets are singletons (conventions). Assume that Π^0 has several conventions $h_1, ..., h_T$. Above it was described how one defines the resistance of a direct transition between two states. For two conventions h and h', define the resistance in the (indirect) transition hh', as the minimal sum of resistances over all collections of direct transitions leading from h to h'. An h-tree is a collection of transitions $h_l h_n$ between conventions such that each convention other than h stands first in exactly one transition, and for any convention $h' \neq h$, there is a unique way to go from h' to h through the transitions of the collection. For each h-tree one can define the sum of resistances embodied in the transitions of the tree. The stochastic potential of h is the minimal of such total resistances over all h-trees.

Theorem 0. *Assume that the basic learning process converges to a convention. Then the stochastically stable states are exactly the conventions with minimal stochastic potential.*

4 Conventional Behavior in a Game of Chicken with Incomplete Information

Below to the left the complete information game "Chicken" is displayed, where $R = C = \{D, H\}$, and $\alpha > 0$ and $\beta > 0$.

	D	H
D	$0,0$	$0,\beta$
H	$\alpha,0$	$-1,-1$

	D	H
D	$0,0$	$1,\beta$
H	$\alpha,1$	$0,0$

This is a prototype strategic situation where each player has a cautious action, here D (Dove), and a tough one, here H (Hawk), and it is good to take the tough action against an opponent playing cautiously - the parameters α and β indicate how good - but bad to take it against an opponent who also plays toughly, while D against D is "neutral".[6] The game has exactly two strict Nash equilibria: One is (H, D), where Player 1 plays H, and Player 2 plays D; the other is (D, H).

[6] The assumption that a player using D receives the same payoff when the opponent plays D as when he plays H simplifies some formulas below, but is not essential for the basic result. The

For the risk dominance relation, [3] one transforms the game to the one to the right with only zero's on the off-equilibrium diagonal and the same best reply structure. The "Nash product" of an equilibrium is the product of the two players' payoffs in the transformed game, and the risk dominating equilibrium is the equilibrium with the largest Nash product, so if $\alpha > \beta$, then (H, D) risk dominates (D, H), and vice versa. [5] and [7] show that the risk dominance selection rule is supported by evolutionary learning. Young's process is as defined in Section 3 with only one type of each player. Assume that the record sizes for both positions in the game are m, and the sample sizes are both k. It should cause no confusion to let a vector of m times H be denoted also by H, etc. The two conventions are (H, D) and (D, H), corresponding to the two strict equilibria. Young shows that if $k < m/3$, then the basic learning process converges, and if further k is sufficiently large, then the convention corresponding to the risk dominating equilibrium is the only stochastically stable state.

We now consider a game of Chicken with incomplete information. The action sets are still $R = C = \{D, H\}$, and uncertainty concerns how good it is for a player to play H against D. Player 1 is of one of two types which, with a slight abuse of notation, are called $1/\alpha$ and α, where α is a number above one. Each type has probability 1/2. Similarly, player 2 is of type $1/\beta$ with probability 1/2, and of type β with probability 1/2, where $\beta > 1$. For the action combinations (D, D) and (H, H) payoffs are independent of types and as given in the complete information game, $(0, 0)$ and $(-1, -1)$ respectively. If the action combination is (H, D), then player 1's payoff is independent of the type of player 2, and it is $1/\alpha$ if player 1 is of type $1/\alpha$, and α if he is of type α, while player 2 gets 0 independently of types. Similarly, if the combination is (D, H), then player 2 gets $1/\beta$ if of type $1/\beta$, and β if of type β, independently of the type of player 1, while player 1 gets zero irrespective of types. For what follows it is convenient that the incomplete information game is (also) given by just two parameters α and β, and that these measure the degree of the players' uncertainty about payoffs.

The incomplete information game has three strict Bayesian equilibria. The first two are: $r(1/\alpha) = r(\alpha) = H, c(1/\beta) = c(\beta) = D$, and $r(1/\alpha) = r(\alpha) = D, c(1/\beta) =$

Chicken game formalizes, e.g., duopoly situations where the strategic conflict is really a battle over the roles as leader and follower.

$c(\beta) = H$. These are counterparts of the strict equilibria of the complete information game; on side plays H, and the other plays D irrespective of types. The players coordinate and clearly the equilibria are efficient ex ante and ex post. The third equilibrium is $r(1/\alpha) = c(1/\beta) = D, r(\alpha) = c(\beta) = H$, where the "low" types play D, and the "high" types play H. This is an equilibrium since against the strategy of player 2, player 1 of type $1/\alpha$ obtains zero in expected payoff from D and $\frac{1}{2}(\frac{1}{\alpha} - 1)$ from H, so D is best, while player 1 of type α obtains zero from D and $\frac{1}{2}(\alpha - 1)$ from H, so H is better, etc. There is a lack of coordination in it, since when both players are of low type or both are of high type, losses occur. The equilibrium is not inefficient ex ante, but it is ex post, and even ex ante it is dominated by a half-half convex combination of the other two equilibria. So, the third equilibrium is qualitatively different from the first two, in particular with respect to efficiency.

Consider the learning process defined in Section 3 for this particular game. Because of the symmetry of the game it is natural to assume $m_{1/\alpha} = m_a = m_{1/\beta} = m_\beta = m$, and $k_{1/\alpha} = k_\alpha = k_{1/\beta} = k_\beta = k$. The three conventions are, in obvious notation, (HH, DD) and (DD, HH) corresponding to the two efficient equilibria with coordination, and (DH, DH) corresponding to the equilibrium with lack of coordination. It is easy to check that the game is weakly acyclic and $L = 2$. Therefore we get directly from Proposition 0,

Proposition 1. *If $k < m/4$, then the basic learning process converges to a convention.*

So, when $k < m/4$ and the trembling probability is small, only behavior in accordance with the three conventions, or strict Bayesian equilibria, can be observed frequently. Theorem 1, which is proved in the Appendix, tells which will be observed,

Theorem 1. *Assume $k < m/4$. If k is sufficiently large, then if $(\alpha - 1)(\beta - 1) < 4$, the conventions (HH, DD) and (DD, HH) are the only stochastically stable states, while if $(\alpha - 1)(\beta - 1) > 4$, the convention (DH, DH) is the only stochastically stable state.*

Thus, if there is substantial difference between the payoffs of the different types (i.e, if α and β are so high, that $(\alpha - 1)(\beta - 1) > 4$), then the process selects the equilibrium (DH, DH), where players sometimes miscoordinate. This happens since high values of α and β imply that in the equilibria (HH, DD) and (DD, HH) it only takes a few trembles by the player using action H for the best response of the high type of the opponent (type α or β) to change, and similarly it only takes a few trembles of the player using action D for the best response of the low type of the opponent (type $1/\alpha$ or $1/\beta$) to change. This is in constrast to what happens in the equlibrium (DH, DH), where each type of each player "goes for his most preferred equilibrium", and it therefore takes more trembles to upset this action being a best response.

The general message is that under incomplete information equilibrium selection by evolutionary learning is no longer just a matter of which of the two efficient equilibria is supported, or who grasps the surplus. Rather, with a substantial amount of uncertainty about payoffs, the qualitatively different, less coordinated equilibrium, and only that, will be observed frequently. Under incomplete information, evolutionary learning may well select in favor of equilibria which are very different with respect to efficiency from the equilibria that are just counterparts of equilibria also occurring under complete information.[7]

5 Conclusions

The evolutionary learning process we have studied is meant to formalize the view that agents involved in strategic conflicts expect their opponents to act more or less as they usually do in identical or similar conflicts. Therefore they form conjectures on opponents from records of past play. They intend primarily to take actions that are best given such conjectures, but with small probability they take more or less arbitrary actions perhaps to test if they are right in the presumption that alternative actions give poorer results.

We have shown in general how this view may imply (i) support for the notion of Bayesian equilibrium, since the learning process may well generate play in accordance with the game's (strict) Bayesian equilibria most of the time, and

[7]It is just an artifact of the payoff specifications above that when $(\alpha-1)(\beta-1) < 4$, both of the conventions corresponding to a coordinated equilibrium are stocahstically stable. It is relatively easy to see that if one increases one of the player 1 payoffs a little bit above $1/\alpha$ or α, then only (HH, DD) is stochastically stable.

(ii) selection among multiple Bayesian equilibria, since the learning process may well generate play in accordance with one particular - or with several, but alike - Bayesian equilibria most of the time.

For a specific incomplete information "surplus division" game of economic interest we have shown that (i) and (ii) are indeed true and we have found the selected equilibrium. It turned out that equilibrium selection by evolutionary learning under incomplete information could well be in favor of an equilibrium with inferior efficiency properties.

Thus we have demonstrated how evolutionary learning can be used to give a foundation for Bayesian equilibrium in static games, and to obtain a selection among the Bayesian equilibria, and we have demonstated that - in contrast to the results for similar games of complete information - the selection is not always in favor of efficient equilibria.

A Proof of Theorem 1

The stochastically stable states are the conventions with minimal stochastic potential. We find the stochastic potentials for all three conventions by systematically computing the resistances in all transitions from one convention to another.

First consider the transition from (HH, DD) to (DH, DH). Obviously, the way to do this transition that requires the fewest trembles (the cheapest way) is either by player 1 type $1/\alpha$ trembling to D thus for possible sampling changing the best reply of player 2 type β to H, or by player 2 type β trembling to H thus changing the best reply of player 1 type $1/\alpha$ to D. Assume Player 1 of type $1/\alpha$ makes l trembles to D, that is, type $1/\alpha$ is picked in l consecutive rounds and plays D in all of them, where $l \leq k$. For a player 2 of type β, who samples all l of the Ds from the record on player 1 type $1/\alpha$, the best reply will have changed to H if $\frac{1}{2}\left[\frac{l}{k}\beta + \frac{k-l}{k}(-1)\right] + \frac{1}{2}(-1) \geq 0$, or if,

$$l \geq \frac{2}{\beta+1}k. \tag{1}$$

This number of trembles is indeed sufficient to get all the way from (HH, DD) to (DH, DH): according to the basic learning process, there is positive probability that in each of the k rounds following the l trembles, player 1 is of type α, so $h_{1/\alpha}$ is unchanged and contains the l times D, and player 2 is of type β and in all k rounds samples all l of the Ds from $h_{1/\alpha}$, and hence plays H in all k rounds. This will gradually insert k times H in h_β. The picked player 1s of type α will with positive probability according to Π^0 have played H in all k rounds (they may not have sampled any Hs from h_β). After this h_β contains k times H. Now it has, according to the basic learning process, positive probability that in each of the next m rounds player 1 is of type $1/\alpha$ and samples k times H from h_β, which gives him best reply D (since $\frac{1}{2}\frac{1}{\alpha} + \frac{1}{2}(-1) < 0$), while player 2 is of type β, and all the time samples at least l times D from $h_{1/\alpha}$, and hence plays H. Note that in these m rounds, when some of the old Ds go out of $h_{1/\alpha}$, new Ds are inserted, and when some of the old Hs go out of h_β, new Hs are inserted, which makes it possible that best replies for possible sampling continues to be D for type $1/\alpha$, and H for type β, during all m rounds. Then the convention (DH, DH) has been reached without further trembling.

Now assume Player 2s of type β make l trembles to H. If player 1 type $1/\alpha$ samples all l of the Hs his best reply will have changed to D if $\frac{1}{2}\frac{1}{\alpha}+\frac{1}{2}\left[\frac{l}{k}(-1)+\frac{k-l}{k}\frac{1}{\alpha}\right] \leq 0$, or if,

$$l \geq \frac{2}{\alpha+1}k. \tag{2}$$

After these l trembles it has, just as above, positive probability according to the basic process to reach (DH, DH) (first k rounds with types $1/\alpha$ and $1/\beta$, and then m rounds with types $1/\alpha$ and β).

So, the resistance in the transition from (HH, DD) to (DH, DH) is,[8]

$$\rho[(HH, DD) \rightarrow (DH, DH)] = \min\{\frac{2}{\beta+1}k, \frac{2}{\alpha+1}k\}.$$

By symmetry it follows that $\rho[(DD, HH) \rightarrow (DH, DH)] = \min\{\frac{2}{\alpha+1}k, \frac{2}{\beta+1}k\}$, which is the same.

Next we consider the transition from (DH, DH) to (HH, DD). It is again evident that this is most cheaply brought about either by trembling of type $1/\alpha$ of player 1, or of type β of player 2. Assume player 1s of type $1/\alpha$ tremble l times to H. The best reply of player 2 type β will for positive sampling have changed to D if $\frac{1}{2}\left[\frac{l}{k}(-1)+\frac{k-l}{k}\beta\right]+\frac{1}{2}(-1) \leq 0$, or if,

$$l \geq \frac{\beta-1}{\beta+1}k. \tag{3}$$

Again it can be shown that after these l trembles it has, according to the basic process, positive probability to go all the way to (HH, DD) (first k rounds with types α and β, and then m rounds with types $1/\alpha$ and β).

Now assume player 2s of type β tremble l times to D. The best reply of player 1 type $1/\alpha$ will for possible sampling have changed to H if $\frac{1}{2}\frac{1}{\alpha}+\frac{1}{2}\left[\frac{l}{k}\frac{1}{\alpha}+\frac{k-l}{k}(-1)\right] \geq 0$, or if

$$l \geq \frac{\alpha-1}{\alpha+1}k. \tag{4}$$

Again, after these l trembles it has, according to the basic process, positive probability to go all the way to (HH, DD). So,

$$\rho[(DH, DH) \rightarrow (HH, DD)] = \min\{\frac{\beta-1}{\beta+1}k, \frac{\alpha-1}{\alpha+1}k\}.$$

[8]Note that here, and in what follows, we are ignoring an integer problem: The resistance is really the smallest integer which is (weakly) above either $\frac{2}{\beta+1}k$ or $\frac{2}{\alpha+1}k$. It is because of the assumption of "k sufficiently large" that this will cause no error (if a resistance as defined here is smaller than another one, then it will also be smaller according to the correct definition if only k is large enough).

Due to symmetry, $\rho[(DH, DH) \to (DD, HH)] = \min\{\frac{\alpha-1}{\alpha+1}k, \frac{\beta-1}{\beta+1}k\}$ and the same.

Finally we consider the transition from (HH, DD) to (DD, HH) (and visa versa). It is evident that if this transition is to be initiated by trembles of player 1, then it is obtained most cheaply if it is type α who does it: trembling by type $1/\alpha$ and α are equally effective in making H a possible best reply for player 2, most easily so for type β (the payoffs of player 2 do not depend on the types of player 1, and these types have the same probability). However, when this has occurred and h_β has become full of Hs, then the only possible best reply of player 1 type $1/\alpha$ is D, so $h_{1/\alpha}$ is filled up with Ds at no further cost, so it can only work to make the transition cheaper that the Ds, that occur because of trembling, are to be found in h_α.

So, assume that from (HH, DD), player 1s of type α make l trembles to D. From (1) the best reply for player 2 type β will have changed to H for possible sampling if $l \geq \frac{2}{\beta+1}k$. After these l trembles it has, according to the basic process, positive probability that in each of the next $2m$ rounds the types are $\frac{1}{\alpha}$ and β, so h_α does not change, and the samples drawn from this record contain in each round all l of the Ds, so each player 2 of type β plays H. After the first m rounds $h_\beta = H$, which makes D the only possible best reply for player 1 type $1/\alpha$, so after the next m rounds also $h_{1/\alpha} = D$.

Now suppose a type $1/\beta$ of player 2 is drawn. With positive probability the sample drawn from h_α will contain all l of the Ds in it. In that case the best reply of type $1/\beta$ will be H if $\frac{1}{2}\frac{1}{\beta} + \frac{1}{2}\left[\frac{l}{k}\frac{1}{\beta} + \frac{k-l}{k}(-1)\right] > 0$. It is already required that $l \geq \frac{2}{\beta+1}k$. By inserting this number of trembles we find that it also suffices for the now considered best reply shift exactly if $-\beta^2 + 2\beta + 3 \geq 0$, that is, if $\beta \leq 3$.

Therefore first consider the case $\beta \leq 3$. It then has, according to the basic process, positive probability that in each of the next m rounds the types are $1/\alpha$ and $1/\beta$, and that, since already $h_\beta = H$, all the player 1s of type $1/\alpha$ play D, keeping $h_{1/\alpha}$ unchanged, and all the player 2s of type $1/\beta$ play H, so after these rounds $h_{1/\beta} = H$. Finally, with positive probability the types α and β are drawn in the next m rounds, and since now $h_{1/\beta} = h_\beta = H$, with positive probability all the player 1s of type α will play D, which can only work to keep H a best reply for player 2 type β, so all of these play H. The state will then be the convention (DD, HH). So, when $\beta \leq 3$, only the $\frac{2}{\beta+1}k$ trembles are required for transition.

Then consider the case $\beta > 3$. Further trembling is required. There are two possibilities: either type α of player 1 trembles further such that the best reply for type $1/\beta$ of player 2 *does* shift to H (for possible sampling), or type $1/\beta$ trembles such that the best reply for type α changes to D.

In the first case, given that already $h_{1/\alpha} = D$, the total number of Ds required in h_α to ensure that H is a possible best reply for type $1/\beta$ is l such that $\frac{1}{2}\frac{1}{\beta} + \frac{1}{2}\left[\frac{l}{k}\frac{1}{\beta} + \frac{k-l}{k}(-1)\right] \geq 0$, or

$$l \geq \frac{\beta - 1}{\beta + 1}k. \tag{5}$$

If this were the (larger) number of trembles first made by type α, then just as above we would with positive probability first over $2m$ rounds get $h_{1/\alpha} = D$ and $h_\beta = H$, and then, in the same way as for $\beta \leq 3$, over the next $2m$ also get $h_{1/\beta} = H$ and $h_\alpha = D$. So, the convention (DD, HH) would then have been reached with $\frac{\beta-1}{\beta+1}k$ trembles.

In the second case, the best reply changes to D for type α of player 1, if player 2 type $1/\beta$ makes l trembles to H, where $\frac{1}{2}\left[\frac{l}{k}(-1) + \frac{k-l}{k}\alpha\right] + \frac{1}{2}(-1) \leq 0$, or $l \geq \frac{\alpha-1}{\alpha+1}k$. By a similar argument as above one can verify that after these *further* trembles there is positive probability according to the basic process of reaching the convention (DD, HH). So, in the second case the total number of trembles required is,

$$\frac{2}{\beta + 1}k + \frac{\alpha - 1}{\alpha + 1}k. \tag{6}$$

We have now shown that if transition from (HH, DD) to (DD, HH) is to be initiated by trembling of player 1, then if $\beta \leq 3$, the number of trembles required is as given by (1), whereas if $\beta > 3$, it is the minimum of the expressions in (5) and (6).

If the transition is to be initiated by trembling of player 2, the trembling should be done by type $1/\beta$ to make the transition cheapest (since it is more easy to make type β play H by best reply). So, assume that player 2s of type $1/\beta$ make l trembles to H. Then from (2), the best reply for player 1 type $1/\alpha$ will (for possible sampling) have changed to D if $l \geq \frac{2}{\alpha+1}k$.

As above it now has positive probability according to Π^0 to reach a state with $h_{1/\alpha} = D$, $h_\alpha = H$, $h_\beta = H$, and in $h_{1/\beta}$ the l entries of H are still there (after $2m$ rounds with types $1/\alpha$ and β). Now (as with $\beta \leq 3$ above), if $\alpha \leq 3$ then for possible sampling, D will be a best reply also for type α of player 1, and (still as above) one can construct a sequence of positive probability events leading all the way to (DD, HH). If $\alpha > 3$, more trembling is required. (Still as above), this can *either* be further trembling to H by player 2 type $1/\beta$, which *does* make D a possible best reply for player 1 type α, the total number of trembles (including both initial and further) required for this being l such that $\frac{1}{2}\left[\frac{l}{k}(-1) + \frac{k-l}{k}\alpha\right] + \frac{1}{2}(-1) \leq 0$, or,

$$l \geq \frac{\alpha - 1}{\alpha + 1}k, \tag{7}$$

or it can be trembling to D by player 1 type α until H becomes a possible best reply for type $1/\beta$ of player 2. This requires l additional trembles where $\frac{1}{2}\frac{1}{\beta} + \frac{1}{2}\left[\frac{l}{k}\frac{1}{\beta} + \frac{k-l}{k}(-1)\right] \geq 0$, or $l \geq \frac{\beta-1}{\beta+1}k$. So, in the last case an overall of,

$$\frac{2}{\alpha + 1}k + \frac{\beta - 1}{\beta + 1}k \tag{8}$$

trembles are needed. In both cases there will, after the additional trembling, be transition all the way to (DD, HH) with positive probability according the basic process.

We have now shown that if transition from (HH, DD) to (DD, HH) is to be initiated by trembling of player 2, then if $\alpha \leq 3$, the number of trembles required is as given by (2), whereas if $\alpha > 3$, it is the minimum of the expressions in (7) and (8).

The resistance in the transition from (HH, DD) to (DD, HH) is the minimum over the numbers of trembles required for transition when trembling is first done by player 1, and when it is first done by player 2. So,

$$\rho[(HH, DD) \rightarrow (DD, HH)] = k \cdot \begin{cases} \min\{\frac{2}{\alpha+1}, \frac{2}{\beta+1}\} & \text{if } \alpha \leq 3, \beta \leq 3 \\ \min\{\frac{\alpha-1}{\alpha+1}, \frac{2}{\alpha+1} + \frac{\beta-1}{\beta+1}, \frac{2}{\beta+1}\} & \text{if } \alpha > 3, \beta \leq 3 \\ \min\{\frac{2}{\alpha+1}, \frac{2}{\beta+1} + \frac{\alpha-1}{\alpha+1}, \frac{\beta-1}{\beta+1}\} & \text{if } \alpha \leq 3, \beta > 3 \\ \min\{\frac{\alpha-1}{\alpha+1}, \frac{2}{\alpha+1} + \frac{\beta-1}{\beta+1}, \frac{2}{\beta+1} + \frac{\alpha-1}{\alpha+1}, \frac{\beta-1}{\beta+1}\} & \text{if } \alpha > 3, \beta > 3 \end{cases}$$

which is equivalent to,

$$\rho[(HH, DD) \rightarrow (DD, HH)] = k \cdot \begin{cases} \min\{\frac{2}{\alpha+1}, \frac{2}{\beta+1}\} & \text{if } \alpha \leq 3, \beta \leq 3 \\ \min\{\frac{\alpha-1}{\alpha+1}, \frac{2}{\beta+1}\} & \text{if } \alpha > 3, \beta \leq 3 \\ \min\{\frac{2}{\alpha+1}, \frac{\beta-1}{\beta+1}\} & \text{if } \alpha \leq 3, \beta > 3 \\ \min\{\frac{\alpha-1}{\alpha+1}, \frac{\beta-1}{\beta+1}\} & \text{if } \alpha > 3, \beta > 3 \end{cases}$$

From symmetry we get by permuting α and β in the expression above that $\rho[(DD, HH) \to (HH, DD)] = \rho[(HH, DD) \to (DD, HH)]$.

In the following we assume wlog. that $\alpha \geq \beta$ (again by symmetry the result is the same if $\beta \geq \alpha$). The expression above is then reduced to,

$$\begin{aligned}\rho[(HH, DD) &\to (DD, HH)] = \\ \rho[(DD, HH) &\to (HH, DD)] = \\ & k \cdot \begin{cases} \frac{2}{\alpha+1} & \text{if } \alpha \leq 3 \\ \min\{\frac{\alpha-1}{\alpha+1}, \frac{2}{\beta+1}\} & \text{if } \beta \leq 3 \leq \alpha \\ \frac{\beta-1}{\beta+1} & \text{if } \beta > 3 \end{cases} .\end{aligned}$$

Now note that $\rho[(DD, HH) \to (HH, DD)] \geq \rho[(DD, HH) \to (DH, DH)]$, which together with the symmetry property $\rho[(DH, DH) \to (DD, HH)] = \rho[(DH, DH) \to (HH, DD)]$ imply that their can be no (HH, DD)-tree with less total resistance than the one consisting of the transitions from (DD, HH) to (DH, DH) and from (DH, DH) to (HH, DD). The stochastic potential of (HH, DD) is then,

$$\begin{aligned}\gamma[(HH, DD)] &= \\ \rho[(DD, HH) &\to (DH, DH)] + \rho[(DH, DH) \to (HH, DD)].\end{aligned}$$

By symmetry, $\gamma[(DD, HH)] = \rho[(HH, DD) \to (DH, DH)] + \rho[(DH, DH) \to (DD, HH) = \gamma[(HH, DD)]$, where the last equality also uses $\rho[(HH, DD) \to (DH, DH)] = \rho[(DD, HH) \to (DH, DH)]$.

Again, since $\rho[(DD, HH) \to (HH, DD)] \geq \rho[(DD, HH) \to (DH, DH)]$ etc., there can be no cheaper way to go to (DH, DH), than to go from each of the other conventions separately, so

$$\begin{aligned}\gamma[(DH, DH)] &= \\ \rho[(HH, DD) &\to (DH, DH)] + \rho[(DD, HH) \to (DH, DH)].\end{aligned}$$

Then we simply arrive at the conclusion that both of (HH, DD) and (DD, HH), and only those conventions, are stochastically stable if and only if $\rho[(DH, DH) \to (HH, DD)] < \rho[(HH, DD) \to (DH, DH)]$, and (DH, DH), and only that, is stochastically stable if $\rho[(DH, DH) \to (HH, DD)] > \rho[(HH, DD) \to (DH, DH)]$ (where it is again used that k is sufficiently large).

Finally, from $\alpha \geq \beta$ it follows that $\rho[(DH, DH) \longrightarrow (HH, DD)] = \frac{\beta-1}{\beta+1}k$, and $\rho[(HH, DD) \rightarrow (DH, DH)] = \frac{2}{\alpha+1}k$. Since $\frac{2}{\alpha+1} < \frac{\beta-1}{\beta+1}$, if and only if $(\alpha-1)(\beta-1) > 4$ etc., the conclusion of Theorem 1 follows. $\square$

References

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

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

[3] Harsanyi, J.C., Selten R. *A General Theory of Equilibrium Selection in Games*, Cambridge: MIT Press (1988).

[4] Jacobsen, H.J., Jensen, M., Sloth, B.: Evolutionary Learning in Signalling Games, *Games and Economic Behavior*, **34**, 34-63 (2001)

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

[6] Nöldeke, G., Samuelson, L.: A Dynamic Model of Equilibrium Selection in Signalling Markets, *Journal of Economic Theory*, **73**, 118-156 (1997)

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