

The Evolution of Conventions under Incomplete Information

Hans Jørgen Jacobsen, Mogens Jensen, and Birgitte Sloth

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

First version: June 1997
This version: February 2000

Abstract

We formulate an evolutionary learning process in the spirit of Young (1993a) for games of incomplete information. The process involves trembles. For many games, if the amount of trembling is small, play will be in accordance with the games' (semi-strict) Bayesian equilibria most of the time. This supports the notion of Bayesian equilibrium. Further, often play will most of the time be in accordance with exactly one Bayesian equilibrium. This gives a selection among the Bayesian equilibria. For two specific games of economic interest we characterize this selection. The first is an extension to incomplete information of the prototype strategic conflict known as Chicken. The second is an incomplete information bilateral monopoly, which is also an extension to incomplete information of Nash's demand game, or a simple version of the so-called sealed bid double auction. For both games selection by evolutionary learning is in favor of Bayesian equilibria where some types of players fail to coordinate, such that the outcome is inefficient.

JEL Classification: C72

Keywords: Games of incomplete information, Bayesian equilibrium, evolution, learning, conventions.

1 Introduction

This paper suggests an evolutionary learning process in the spirit of Young (1993a) 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 two specific games of economic interest the selection is characterized. The first is an extension to incomplete information of the prototype strategic conflict known as the Chicken game. The second is an incomplete information bilateral monopoly; it can be viewed as an extension to incomplete information of Nash's demand game, and it is also equivalent to a so-called (sealed bid) double auction. For both games selection turns out to be in favor of Bayesian equilibria exhibiting certain inefficiencies. These can be attributed to the presence of incomplete information, since they are due to miscoordination between some types of players.

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 some games of incomplete information, like that given of Nash equilibrium for some games of complete information by, e.g., Young (1993a). Second, going from complete to incomplete information in games often adds a dimension of equilibrium multiplicity, in particular it often implies the existence of many 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., Young (1993a) and Kandori, Malait, and Rob (1993), 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 formulation 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. 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 chosen actions as well as the true types of the players who took them, are perfectly observed by everybody.

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 semi-strict Bayesian equilibria most of the time. For such games one has then obtained the kind of support for the notion of Bayesian equilibrium looked for. Further, also when a game has several Bayesian equilibria, it may well be that play will be in accordance with a specific one most of the time. This will then yield a selection among the Bayesian equilibria. We hope by our analyses of the Chicken and bilateral monopoly games to demonstrate the economic usefulness of this selection in situations with incomplete information.

Both games are extensions to incomplete information of the kind of coordination games typically studied in the earlier contributions on evolutionary learning in games, and both are illustrations of how incomplete information may add an eco-

onomically important dimension of equilibrium multiplicity. Each game has under complete information several strict Nash equilibria, but these have similar efficiency properties. In Chicken both pure Nash equilibria involve one player taking the tough action and the other taking the cautious one; equilibrium selection is only a matter of who takes which action, i.e., who gets the (main part of the) surplus. In Nash's demand game there is a strict Nash equilibrium for each possible strict division, but all equilibria are efficient since all surplus is exploited; equilibrium selection is only a matter of how large a share each player gets. Under incomplete information the games may well have equilibria with qualitatively different efficiency properties. In Chicken there are still the two equilibria where (independently of types), one side takes the tough, and the other side takes the cautious, action, but in addition there can be an equilibrium where sometimes (for some types), both players take the tough action, and a waist occurs. Our extension of Nash's demand game to incomplete information is one where the game is interpreted as a bilateral monopoly and the buyer and the seller are uncertain about each others' reservation prices. Such a game may well have both strict equilibria, where the bids of the different types are spread out in a way that ensures that any two types who can trade to mutual benefit do so, and strict equilibria where the bids are clustered in a way so that there are couples of types who could trade to mutual benefit, but do not do so; the potential gains from trade are not exhausted. The first type of equilibrium is efficient, the latter is inefficient. For both games the type of inefficiency possible in equilibrium is due to incomplete information.

A theory of equilibrium selection in the complete information games, for instance by evolutionary learning, can give insight into how surplus is shared. This is of economic interest. The issue if one should expect to observe equilibria with one or the other type of efficiency properties in the incomplete information games is also of economic interest; it may, for instance, be important for the recommendation of the double auction as a trade mechanism. We address the second issue by studying equilibrium selection by means of evolutionary learning in an incomplete information version of each game. In both a unique Bayesian equilibrium is selected. For the Chicken game with incomplete information this will, if there is considerable uncertainty about payoffs, be of the type where sometimes tough is played against tough. For the bilateral monopoly the selected equilibrium is one

with clustered bids where the potential gains from trade are poorly exploited. In both cases selection is in favor of an equilibrium exhibiting a certain kind of inefficiency that can be attributed to incomplete information. Equilibrium selection by evolutionary learning in incomplete information games may well be in favor of equilibria that have qualitatively different (efficiency) properties than the equilibria that also occur under complete information.

The present paper is closely related to contributions such as Young (1993a), Canning (1992a), and Kandori, Malaith, and Rob (1993), 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, Canning (1992b), Nöldeke and Samuelson (1997), Jacobsen, Jensen, and Sloth (1999). 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.

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 Sections 4 and 5 we analyze the incomplete information versions of the Chicken and the bilateral monopoly game respectively, and characterize the long run behaviors supported by evolutionary learning. Section 6 concludes. Proofs are given in Appendix. However, some of the proofs are very long in their complete versions and for these the Appendix sketches the main steps only, while the complete proofs are available from the authors upon request.

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 .¹ Each player 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 $BR_\alpha(q)$, and let $BUR_\alpha(q)$ be the subset of undominated best replies, that is, r is in $BUR_\alpha(q)$, if r is in $BR_\alpha(q)$ and there is no other action that (weakly) dominates r . Define the expected payoff of Player 2, $V_\beta(c, p)$, similarly and let the sets of best replies and best undominated replies for a Player 2 of type β holding a conjecture p be $BR_\beta(p)$ and $BUR_\beta(p)$ respectively. The sets $BUR_\alpha(q)$ and $BUR_\beta(p)$ are 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 semi-strict if for all the α 's and β 's, $BUR_\alpha(q)$ only contains $r(\alpha)$, and $BUR_\beta(p)$ only contains $c(\beta)$. We only consider games with at least one semi-strict Bayesian equilibrium.

¹For everything below we could, more generally, have chosen a formulation where the set of available strategies possibly depends 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.

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 $r : A \rightarrow R$, and $c : B \rightarrow C$, 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 (1993a) 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 (almost) 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 prefer to face these in a general formulation rather than for the specific games separately.

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.

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 being 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 $BUR_\alpha(Q)$, if he is of type α . If $BUR_\alpha(Q)$ has several elements, he will pick one at random according to a full support probability measure on $BUR_\alpha(Q)$. Similarly a picked player 2 of type β will take an action in $BUR_\beta(P)$.

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 $\pi^0(h, h')$. Of course, 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: **(1)** 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 $c(\beta)$. **(2)** Each recorded action is the unique undominated best reply to the samples that are possible from h : If one lets P_α be the list of k_α times $r(\alpha)$, and lets Q_β be the list of k_β times $c(\beta)$; then $BUR_\alpha(Q) = \{r(\alpha)\}$ for each α , and $BUR_\beta(P) = \{c(\beta)\}$ 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 case of several undominated best replies all have positive probability, it is also true that every absorbing state is a convention. Further, for any convention there is a semi-strict Bayesian equilibrium defined in the obvious way, and every semi-strict Bayesian equilibrium defines exactly one convention. Since we have assumed that there is a semi-strict Bayesian equilibrium, Π^0 has at least one convention and absorbing state.

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,² 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 a result in Young (1993a). 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 an undominated best reply to the strategy 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 semi-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.

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

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.

²If 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.

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.

The definition of the basic learning process involves some modelling choices, which are not just straightforward extensions of formulations in Young (1993a). One has nothing in particular to do with incomplete information: We assume that players only consider undominated best replies, not all best replies, as candidates for actions to take. Consequently they put positive probability on all undominated, not on all, best replies. Unexpected actions will occur frequently due to trembling, and it is therefore dangerous to play dominated actions. Since there are always undominated best replies, it is most natural to assume that players prefer these. As a consequence the equilibria that correspond to the absorbing states of the basic learning process are all the semi-strict, and not only all the strict as in Young (1993a).

The other two are: **(1)** individuals know the true measures a and b , and **(2)** they perfectly observe types (as well as actions). These are linked to incomplete information and can be said to be at the most well-informed extreme. A consequence of **(1)** is that the probabilities of types are not, as are the action choices of types, subject of learning. In favor of **(1)** it can be argued that information about a population's distribution over types is often more available than information about how types choose actions. For instance, it may be officially statistically recorded how people are distributed according to income, consumption, occupation, etc., or how firms are distributed according to plant size, financial or ownership structure, etc., features which may be relevant for the (not directly observable) "type" of a player. For **(2)** note that if the player pools are large, each individual only plays rarely, so no individual has incentive to hide his type after having played. Further, for some games of incomplete information there is enough "structure" in the way payoffs depend on types to ensure that one can read the opponent's type from one's own payoff. Therefore, if payoffs can be observed, so can types. A purely analytical motivation for **(1)** and **(2)** is that it seems natural in a first investigation of the evolutionary approach in static games of incomplete information to consider assumptions at the most well-informed extreme.

3.2 The Perturbed Learning Process

The basic view taken in the literature on evolutionary learning is that a slightly perturbed process involving trembles is more realistic than the basic one described above. We modify the process by assuming that there is in each round for each player independently a small probability ε 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 undominated 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 .

There is here a modelling choice which is particularly, but not only, relevant for games of incomplete information. An alternative specification could be that in the ε -eventuality of a tremble a random action is picked from a full support distribution only over the subset of undominated actions. This could be particularly relevant if trembles are interpreted as experiments. Such a restriction on trembling could easily 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 where all types of a player have the same strategy set and trembles may go to all alternative strategies in that set. We note, however, first that all of the qualitative general results reported in this section would be identical in the 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 games studied below does not depend on trembling to dominated actions. For the Chicken game this follows since there are no dominated actions. For the bilateral monopoly game there are many dominated actions since it is always dominated for a buyer, e.g., to bid at or above his reservation price. We demonstrate, however, that the selected equilibrium is the same if trembling can go to all bids as when trembling is restricted to go to bids on the "right" side of the reservation price. The selection results thus hold both for interpretations

of trembles that would naturally allow these to go to dominated actions, and for interpretations that would naturally exclude them from doing so.

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) appropriately. Further, Π^ε is aperiodic: this follows since Π^ε is irreducible and there is a convention h with $\pi^\varepsilon(h, h) > 0$. Finally, Π^ε is a regular perturbation of Π^0 in the sense of Young(1993a), i.e., Π^ε is irreducible and aperiodic, $\Pi^\varepsilon \rightarrow \Pi^0$ as $\varepsilon \rightarrow 0$, 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 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 $\mu^0 \equiv \lim_{\varepsilon \rightarrow 0} \mu^\varepsilon$. It follows by Young (1993a), 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 semi-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 close conventions are, stochastically stable, so only play according, or close, to a specific of the game's semi-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 from Young (1993a), specialized to the situation where all absorbing sets are singleton (i.e., conventions). Assume that Π^0 has several conventions h_1, \dots, 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 $h_i h_n$ 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 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".³ 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) .

For the risk dominance relation, Harsanyi and Selten (1988), 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. Young (1993a) and Kandori, Mailath, and Rob (1993) 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 also to let a vector of m times H be denoted by H , etc. The two conventions are (H, D) and (D, H) , corresponding to the two semi-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 only 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, that is, $(0, 0)$ 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

³The 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 Chicken game formalizes, e.g., duopoly situations where the strategic conflict is really a battle over the roles as leader and follower.

$1/\beta$, 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 $r(1/\alpha) = r(\alpha) = D, c(1/\beta) = c(\beta) = H$. These are counterparts of the strict equilibria of the complete information game; one 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, which are as the equilibria under complete information, 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 (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, will be observed frequently. Theorem 1 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.

The main message is that under incomplete information equilibrium selection by evolutionary learning is no longer only 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, it is the qualitatively different, less coordinated equilibrium, and only that, that will be observed frequently. Under incomplete information, evolutionary learning may well select in favor of equilibria with efficiency properties very different from those of the equilibria, which are counterparts of equilibria also occurring under complete information.⁴

5 Conventional Behavior in Bilateral Monopoly

Bargaining in a bilateral monopoly with uncertainty about reservation prices can be modelled as a game of incomplete information, as first suggested by Chatterjee and Samuelson (1983). There are two players: player 1 is a buyer characterized by a reservation price $\alpha \in A$, and player 2 is a seller characterized by a reservation price (or cost) $\beta \in B$, where A and B are the sets of possible reservation prices. The buyer offers a price $r \in R$, and the seller asks for a price $c \in C$, and it is assumed that $R = C = M$, where M is the set of allowed bids. Trade occurs if and only if $r \geq c$, at the trading price $p = \frac{r+c}{2}$.⁵ The payoff for the buyer is zero if no trade occurs, and $\alpha - p$ otherwise, and the payoff for the seller is zero if no trade occurs, and $p - \beta$ otherwise. We will refer to α and β as the types of the buyer and the seller respectively. Each player knows his own type, but not the opponent's. The belief concerning the latter is given by some probability distribution over A or B , which is independent of the player's own type.

This game has gained much attention in the literature. It captures some important aspects of bargaining under incomplete information. In fact it is an extension

⁴It 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 stochastically stable. It is relatively easy to see, e.g., that if one from such a situation increases one of the player 1 payoffs a little bit above $1/\alpha$ or α , then only $(H, H; D, D)$ is stochastically stable.

⁵The set of allowed trading prices is thus the set of all possible midpoints of allowed bids.

to incomplete information of the bargaining game known as Nash's demand game; a simple version of the latter appears as the complete information special case, $A = \{1\}$, $B = \{0\}$. Furthermore, it is a possible formalization of what goes on whenever a buyer meets a seller, and the two are not completely sure about the other's eagerness to sell or buy. Finally, it is a simple case of a "sealed bid double auction", a much studied trade mechanism, see Satterthwaite and Williams (1989).

In the special case of Nash's demand game, there is a strict Nash equilibrium for each bid $x \in M$ strictly between zero and one, where both players bid x . All these equilibria are efficient and only differ with respect to the division of the surplus. Consider for this game the learning process defined in Section 3, just with only one type for each player, and assume that record sizes on both players are m , and sample sizes are both k , and that $k \leq m/2$. There is a convention for each of the strict Nash equilibria, where each convention consists of two records, one for each player, with only the bid x of the equilibrium in them. It is an implication of the result in Young (1993b), that for large enough k , the convention with $x = 1/2$, is the only stochastically stable state (assuming M symmetric with enough bids in it, and $1/2 \in M$).

Under incomplete information also equilibria with inferior efficiency properties may exist. An often studied version of the game is where all of A , B , and M are the unit interval, and α and β are independently uniformly distributed. This game has infinitely many Bayesian equilibria as demonstrated by Leininger, Linhart, and Radner (1989).

Some are stepwise. In these the players cluster their bids; A and B are partitioned into the same number of intervals, and all types in some interval of reservation prices submit the same bid. The bids used by the two sides are the same. The simplest case is a one-step equilibrium: Let x be any number (strictly) between zero and one. The strategies where the buyer bids x if $\alpha \geq x$, and zero otherwise, and the seller bids x if $\beta \leq x$ and one otherwise, form a Bayesian equilibrium. Trade only occurs if β is below x , and α is above x , and the equilibrium is clearly inefficient: On either or both sides of x , there will be couples of types α, β who could trade to mutual benefit, $\alpha > \beta$, but do not do so since they are on the same side of x . There are also stepwise equilibria with two, three, etc. steps, and they are also inefficient (and not even second-best, see below), but as long as they do not involve

too many steps, they have the attractive property that they are relatively simple: the way bids are clustered reminds of a simple rule of thumb.

Other equilibria are "smooth": the strategies are smooth functions of type (but can be modified to non-smooth strategies in areas where no trade occurs). The strategies must satisfy a system of first order differential equations. A particular smooth equilibrium is of special interest: Myerson and Satterthwaite (1983) show that among all Bayesian equilibria the strategies given by $r(\alpha) = \frac{2}{3}\alpha + \frac{1}{12}$, and $c(\beta) = \frac{2}{3}\beta + \frac{1}{4}$ maximize the total expected gains from trade. If $\alpha < \frac{1}{4}$ or $\beta > \frac{3}{4}$, there is no trade, so the equilibrium is not fully efficient,⁶ only second-best. It could be held against this equilibrium, however, that it is complicated, since it entails a very "fine tuned" dependence of bid on type.

Incomplete information adds a dimension of multiplicity where equilibria differ also with respect to basic efficiency properties. The issue, if one should normally expect to see the more or the less efficient equilibria in actual play of this game, is important, e.g., for the recommendation of the double auction as a trade mechanism. The basic tension, or equilibrium selection problem, seems to stand between efficiency and simplicity.

We will investigate what equilibrium selection by evolutionary learning has to say on this tension. We have to consider a finite version of the game, and to be able to characterize the stochastically stable states, a simplified one. We consider the specification where $A = B = \{0, \frac{1}{3}, \frac{2}{3}, 1\}$, and type is independently uniformly distributed; each of the four types of a player has probability $\frac{1}{4}$. Further, $M = \{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}$, where $n = 10^N$ for some natural number N .⁷ This is a simplest possible version of the game still representing a "true" conflict between efficiency and simplicity in the sense that the game has semi-strict equilibria of two kinds:

⁶The feature that for $\alpha < \frac{1}{4}$, the buyer bids above his reservation price, is thus not so peculiar; there is never trade if $\alpha < \frac{1}{4}$. The (non-smooth) strategies with $r(\alpha) = 0$ for α below $\frac{1}{4}$, and $c(\beta) = 1$ for β above $\frac{3}{4}$, and otherwise as above, also form an equilibrium and give the same outcome.

⁷The assumption, $n = 10^N$, is technical and not essential for our results. It implies that $\frac{1}{3}$ and $\frac{2}{3}$ are not allowed bids, whereas $\frac{1}{2}$ is. Further it implies that if $y < \frac{2}{3}$ is an allowed bid, then there is another allowed bid $x < \frac{1}{3}$, such that $2x \geq y$. These properties are convenient since they contribute to establish that the only absorbing sets for the basic learning process are the singleton sets corresponding to semi-strict equilibria, or (which is the same), that the basic learning process converges to a convention. In the absence of $n = 10^N$, there could also be some non-singleton absorbing sets, but the presence of these would not affect the stochastically stable states and hence not the selection result we obtain.

3-step equilibria which are efficient, and 2-step equilibria which are inefficient, but simpler; see below.⁸

For the buyer it is dominated to bid anything at or above the reservation price α , except if of type $\alpha = 0$, in which case the only undominated bid is $r = 0$. On the other hand, any allowed bid $x < \alpha$ is undominated; it is the unique best reply for the conjecture that the seller bids x irrespective of type. Likewise, for a seller of type β , the undominated actions are all bids in M strictly above β , except for $\beta = 1$, where the only undominated bid is $c = 1$. There are many Bayesian equilibria in pure strategies using only undominated actions. The strategies $r(\alpha) = 0$ for all α , and $c(\beta) = 1$ for all β , form a "zero-step" equilibrium which involves no trade; it is highly inefficient. Next, for any allowed bid $x \neq 0, 1$, the strategies $r(\alpha) = x$ if $\alpha > x$, and $r(\alpha) = 0$ otherwise, and $c(\beta) = x$ if $\beta < x$, and $c(\beta) = 1$ otherwise, form a "one-step" equilibrium which involves some trade, but is still inefficient. A one-step equilibrium is illustrated in Figure 1(a). Zero- and one-step equilibria are not semi-strict: there will always be either a type $\alpha > 0$ of buyer who bids zero, or a seller $\beta < 1$, who bids one, which is best reply, but not the only undominated best reply, since, e.g., the smallest allowed bid above zero is also undominated best reply for such a buyer. Therefore none of these equilibria can correspond to conventions.

There are, however, semi-strict two-step equilibria. Let x and z be two allowed bids with $x < \frac{1}{3}$, and $z > \frac{2}{3}$. Consider the strategies $r(\cdot)$ and $c(\cdot)$ for the buyer and seller respectively,

$$r(0) = 0, r\left(\frac{1}{3}\right) = r\left(\frac{2}{3}\right) = c(0) = x, r(1) = c\left(\frac{1}{3}\right) = c\left(\frac{2}{3}\right) = z, c(1) = 1, \quad (1)$$

as illustrated in Figure 1(b). It is immediate that for $\alpha \in \{0, \frac{1}{3}, \frac{2}{3}\}$, the action $r(\alpha) = x$ is the unique undominated best reply to $c(\cdot)$. For $r(1) = z$ to be also the unique undominated best reply, a deviation by type $\alpha = 1$ down to x must yield less expected payoff, that is, one must have $\frac{1}{2}(1 - z) + \frac{1}{4}(1 - \frac{x+z}{2}) > \frac{1}{4}(1 - x)$, or $x > 5z - 4$. Likewise, $c(\beta)$ is obviously best reply for $\beta \in \{\frac{1}{3}, \frac{2}{3}, 1\}$, but for $\beta = 0$ to play a unique undominated best reply one must have $\frac{1}{2}x + \frac{1}{4}\frac{x+z}{2} > \frac{1}{4}z$, or $5x > z$.⁹ So, under the further restrictions $x > 5z - 4$, and $5x > z$, also implying $x > 0$ and

⁸With only the types $0, \frac{1}{2}$, and 1 , and the game otherwise the same, there would still be both efficient 2-step and inefficient 1-step equilibria, but only the first could be semi-strict.

⁹Evidently, the latter condition must appear from the first by symmetry: in $x > 5z - 4$, change x to $1 - z$, and z to $1 - x$, and $5x > z$ results.

$z < 1$, the strategies form a semi-strict Bayesian equilibrium, and such strategies always exist: just set $x = \frac{3}{10}$ and $z = \frac{7}{10}$. The only reason these equilibria are not fully strict is the presence of the "redundant" types $\alpha = 0$, and $\beta = 1$; other types play unique best replies. The two-step equilibria are not efficient, because if a buyer of type $\alpha = \frac{2}{3}$ happens to meet a seller of type $\beta = \frac{1}{3}$, the two will not trade according to equilibrium behavior, although they could do so to mutual benefit at any price between their reservation prices.¹⁰ There are also some equilibria involving trade at two different prices which are not semi-strict. For instance, there exist pairs of allowed bids fulfilling $0 < x < \frac{1}{3} < y < \frac{2}{3}$, and strategies $r(0) = 0$, $r(\frac{1}{3}) = x$, $r(\frac{2}{3}) = r(1) = y$, and $c(0) = x$, $c(\frac{1}{3}) = y$, $c(\frac{2}{3}) = c(1) = 1$, that form an equilibrium, but here the seller of type $\beta = \frac{2}{3}$ does not have a unique undominated best reply.

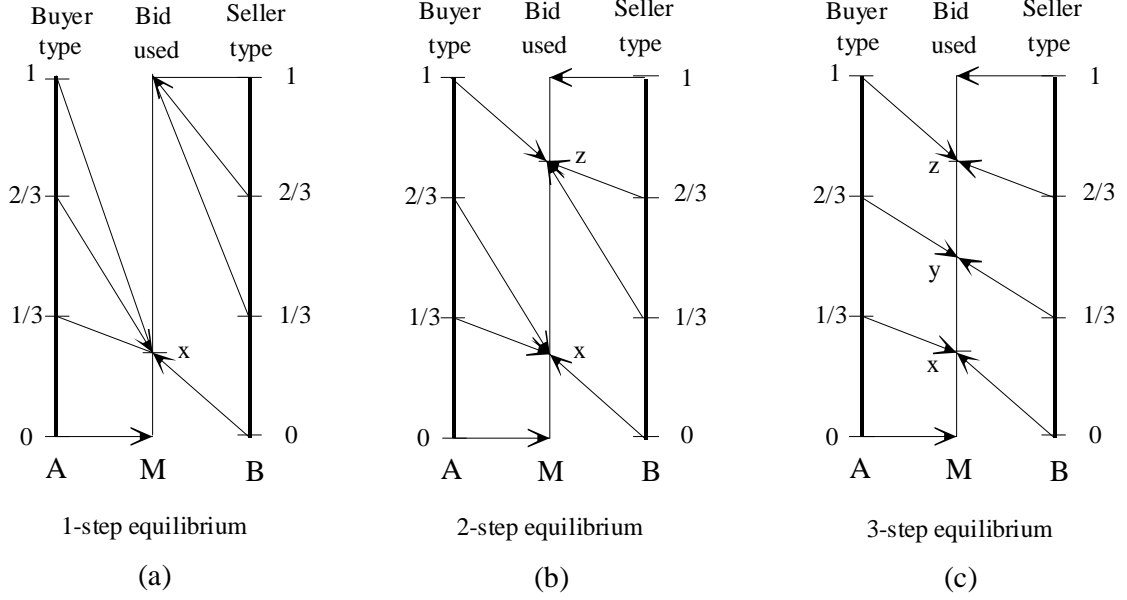
Finally, there are semi-strict three-step equilibria. Let three allowed bids fulfill $x < \frac{1}{3} < y < \frac{2}{3} < z$, and let the strategies be,

$$r(0) = 0, r(\frac{1}{3}) = c(0) = x, r(\frac{2}{3}) = c(\frac{1}{3}) = y, r(1) = c(\frac{2}{3}) = z, c(1) = 1, \quad (2)$$

as illustrated in Figure 1(c). Under some restrictions these form a semi-strict equilibrium. Obviously, the buyer types $\alpha \in \{0, \frac{1}{3}\}$ play best replies. For $\alpha = \frac{2}{3}$, it must not pay to deviate from y down to x : $\frac{1}{4}(\frac{2}{3} - y) + \frac{1}{4}(\frac{2}{3} - \frac{x+y}{2}) > \frac{1}{4}(\frac{2}{3} - x)$, or $4 + 3x > 9y$. Similarly one finds that the condition for z being strictly better than y for type $\alpha = 1$, is $y > 2z - 1$. One does not have to check whether a deviation by $\alpha = 1$ all the way down to x could pay, due to an obvious monotonicity property: higher types can only have higher best replies. So, if deviation down to x payed for $\alpha = 1$, so it would have done $\alpha = \frac{2}{3}$. By symmetry (in the above conditions change x to $1 - z$, and z to $1 - x$, and y to $1 - y$), the condition for a seller of type $\beta = 0$ to strictly prefer x to y is $2x > y$, and for $\beta = \frac{1}{3}$ to strictly prefer y to

¹⁰The presence of an unexploited opportunity of trade is not a full proof of inefficiency. One has to demonstrate the existence of better strategies. Change the equilibrium strategies so that $r(\frac{2}{3}) = c(\frac{1}{3}) = y$, where $\frac{1}{3} < y < \frac{2}{3}$, the strategies being otherwise the same. At the new strategies, buyers of types different from $\alpha = \frac{2}{3}$ obviously get higher or the same expected payoff. Type $\alpha = \frac{2}{3}$ gets higher payoff if $\frac{1}{4}(\frac{2}{3} - y) + \frac{1}{4}(\frac{2}{3} - \frac{x+y}{2}) > \frac{1}{4}(\frac{2}{3} - x)$, or $9y < 4 + 3x$. Similarly, only for the seller type $\beta = \frac{1}{3}$ could there be doubt about payoff, and this type gets higher expected payoff if $\frac{1}{4}(\frac{y+z}{2} - \frac{1}{3}) + \frac{1}{4}(y - \frac{1}{3}) > \frac{1}{4}(z - \frac{1}{3})$, or $9y > 2 + 3z$. Under the equilibrium restrictions on x and z , one has $4 + 3x > 2 + 3z$, so one can find a number y , such that $\frac{2+3z}{9} < y < \frac{4+3x}{9}$ (also implying $\frac{4}{9} < y < \frac{5}{9}$), and in fact the set M will always contain an allowed bid with this property. The new strategies are at least as good as the equilibrium strategies for all types and strictly better for some types.

Figure 1: Types of equilibria in the bilateral monopoly



z is $9y > 3z + 2$. Again, due to monotonicity in best reply as a function of type (and since seller types $\beta \in \{\frac{2}{3}, 1\}$ obviously play best replies), one does not have to check further. The obtained restrictions imply that $x > 0$ and $z < 1$, and strategies obeying all restrictions exist, e.g., for $x = \frac{3}{10}, y = \frac{1}{2}, z = \frac{7}{10}$. These strategies form a semi-strict Bayesian equilibrium which is also fully strict except for the redundant types. It goes without proof that the three-step equilibria are efficient. We have established,

Proposition 2. *There are exactly the following semi-strict Bayesian equilibria of the considered game: Two-step equilibria given by the strategies in (1) (as illustrated in Figure 1(b)), and the restrictions that x and z are both in M , that $0 < x < \frac{1}{3}, \frac{2}{3} < z < 1$, and,*

$$5x > z \text{ and } x > 5z - 4, \quad (3)$$

and three-step equilibria given by the strategies in (2) (as illustrated in Figure 1(c)), and the restrictions that x, y , and z are all in M , that $0 < x < \frac{1}{3} < y < \frac{2}{3} < z < 1$,

and,

$$4 + 3x > 9y \text{ and } y > 2z - 1 \text{ and } 2x > y \text{ and } 9y > 3z + 2. \quad (4)$$

All the two-step equilibria are inefficient, whereas all the three-step equilibria are efficient.

In our simple representation of the bilateral monopoly with incomplete information the distinction between equilibria that are second-best and equilibria that are not second-best, takes the sharp form of a distinction between fully efficient and inefficient equilibria, which is convenient. Essentially this is because we have only few types.

For the evolutionary learning process defined in Section 3 it seems natural to specify $m_\alpha = m_\beta = m$, and $k_\alpha = k_\beta = k$ for all α, β , where $k \leq m$, because of the symmetry of the game. The usual assumption of "not too much sampling", like $k < m/3$ in Theorem 1, contributes to avoid cyclical behavior of the basic learning process Π^0 , and hence to establish convergence to a convention. We do not need such a restriction here, since the presence of the redundant types $\alpha = 0$ and $\beta = 1$, essentially does the same and eliminates cyclical behavior of Π^0 .¹¹

Our interest is in the perturbed learning process Π^ε , given by an $\varepsilon > 0$, for ε small. Since we use the general definition of Π^ε to the particular game studied here, a consequence is, as discussed in Section 3.2, that in the ε -eventuality of a tremble, all bids, including the dominated, have positive probability. This may be problematic for some interpretations of the trembling. In the complete proof of Theorem 1 we therefore show that under the alternative assumption, that trembling can only be to bids on the "right" side of a player's reservation price, the theorem still holds.

A state h now consists of four records h_α and four records h_β , each being a list

¹¹It is open for discussion, if the redundant types should be present in an appropriate formalization of the bilateral monopoly with uncertainty about opponents' reservation prices. We have included them here to preserve the full symmetry from the game with a continuum of types. However, a buyer of type $\alpha = 0$, for instance, does not really want to buy from any present type of seller. The issue is whether it could happen in a "realistic" description of the trade situation, that in an encounter a seller does not really meet an opponent, here represented by meeting type $\alpha = 0$, and therefore "bids out in the open air", so to say. We think that in some situations that could happen, and in others not. Therefore, in the complete proof of Proposition 3 below stating convergence to a convention we demonstrate that also without the redundant types, one would still have convergence to a convention provided $k < \frac{m}{2}$. It follows that the equilibrium selection described in Theorem 2 would, under these alternative circumstances, also be the same.

of the bids from M observed the last m times the type in question were picked for actual play of the game. Each semi-strict Bayesian equilibrium $r(\cdot), c(\cdot)$ defines the convention with constant records $h_\alpha = (r(\alpha), \dots, r(\alpha))$ for each α , and $h_\beta = (c(\beta), \dots, c(\beta))$ for each β , and there are no other conventions than the so defined; there is a one-to-one correspondence between conventions and semi-strict equilibria. This is just a repetition of general properties described in Section 3. A state can also correspond to an equilibrium that is not semi-strict, if it has constant records only listing actions in accordance with the equilibrium, but such a state is not a convention.

Proposition 3. *The basic learning process converges to a convention.*

Let h^* be the convention that corresponds to the particular two-step equilibrium given by (1) with $x = \frac{1}{4}$, and $z = \frac{3}{4}$; from Proposition 1, this is indeed a semi-strict equilibrium as long as $N \geq 2$.

Theorem 2. *The convention h^* is stochastically stable. If n and k are sufficiently large, then h^* is the only stochastically stable state.*

In words Theorem 2 says: If conventional behavior evolves gradually as described by the evolutionary learning process with small trembling probability ε , and there is not too little sampling (or too few allowed bids), then the only behavior that will be observed frequently in the long run is behavior in accordance with the inefficient two-step equilibrium given by $x = \frac{1}{4}$ and $z = \frac{3}{4}$.

That conventional behavior is in accordance with a simple, inefficient equilibrium with clustered bids has only been demonstrated for the version of the game with four types, three of which are non-redundant. Obviously a result for a version with T non-redundant uniformly distributed types would be more general and therefore to prefer, but also extremely difficult to derive.¹² From the present Theorem 2, however, one can safely conclude that if the uncertainty of the trading situation is properly formalized by only "a few" types, and observed behavior is "conventional", then there is no reason to expect the outcome of the considered trading

¹²The game we consider is quite complicated as it is. It contains the kind of equilibrium multiplicity in bids known from the Nash demand game, for which Young (1993b) characterizes conventional behavior in an elegant, but also lengthy proof. In addition it contains the multiplicity in bid clustering originating from incomplete information.

mechanism to be efficient (or second-best). We think this is of relevance. Note in particular, that the simple game considered here is not a special case of the game with a continuum of types uniformly distributed over the unit interval. Despite its simplicity it may well approximate the uncertainty in actual trade situations better than the formalization with a continuum of types.

Theorem 2 is a statement on the limit behavior that will be observed as $\varepsilon \rightarrow 0$, if n and k are sufficiently large. One may rightfully wonder, if there is only a reasonable degree of precision in the prediction pointed to by Theorem 1 for extreme values of ε , n , and k . We therefore performed a simulation with moderate parameter values which rejects this suspicion.¹³ The game studied in the simulation is as above, only with $M = \{0, \frac{1}{16}, \dots, \frac{15}{16}, 1\}$. This is the second-smallest symmetric set that includes the bids $\frac{1}{4}$ and $\frac{3}{4}$. For the learning process we used $k = m = 5$, and $\varepsilon = 1\%$. The iterations were started in the convention corresponding to the three-step equilibrium $x = \frac{1}{4}, y = \frac{1}{2}, z = \frac{3}{4}$. From the proof of Theorem 2, this is, so to say, the closest three-step competitor to h^* . Control simulations revealed, however, that one gets almost the same figures if one starts the iterations in h^* itself. The iterations went on until several jumps from conventions corresponding to two-step equilibria to conventions corresponding to three-step equilibria, and vice versa, had occurred. The program counts the number of steps where the state is exactly equal to h^* . However, this underestimates how often the process is really at the two-step equilibrium. Occasionally trembles occur which will in the subsequent rounds pass through the records without making these essentially different from h^* . The program therefore also counts the number of rounds in which the state is such that the best replies of all types are the same as in state h^* . Finally the program counts the number of rounds in which best replies are as in some three-step equilibrium. The results are that in 71% of the time the state is exactly h^* , and in 97% of the time play generated by the state is (best replies are) as in h^* . In 0.37% of the time the state will generate play according to a three-step equilibrium. The conclusion is that also for moderate values of ε , n , and k , play will most of the time correspond to the two-step equilibrium with $x = \frac{1}{4}$ and $z = \frac{3}{4}$.

¹³The GAUSS-programme used for these simulations is available from the authors upon request.

6 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 and 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 (semi-strict) Bayesian equilibria most of the time, and *(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 two specific incomplete information games of economic interest we have shown that *(i)* and *(ii)* are indeed true and we have found the selected equilibria. It turned out that equilibrium selection by evolutionary learning under incomplete information could well be in favor of equilibria that are inefficient as compared to the equilibria that are counterparts of equilibria occurring under complete information.

By our analyses of the specific games we hope to have convinced our readers that evolutionary learning is an important instrument both for support of Bayesian equilibrium, and for selection among Bayesian equilibria. One should, however, also be aware of the limitations of this instrument. Considering a game and specifications of the learning process that make *(i)* and *(ii)* hold, we think that the main limitations are: Whatever has been obtained in terms of *(i)* and *(ii)* above, has only been obtained for situations of recurrent play, and has only been obtained for situations where the types and their probabilities have a physical meaning in terms of subdivisions of player populations.

For the first of these limitations we note that this is just as the support of Nash equilibrium, and selection among multiple Nash equilibria, obtained in work such as Young (1993a) and Kandori, Malait, and Rob (1993); it only pertains to situations of recurrent play.

For the second we note that if this is already the way one thinks about types, etc., there is no real limitation here. However, often the types are meant to describe a purely subjective uncertainty. There is really only one type of each player, and the different types of a player represent the uncertainty in the head of the opponents. In cases where this is the appropriate interpretation of types our analysis cannot be used to motivate Bayesian equilibria or selection among multiple such equilibria. In that case, however, it is perhaps not at all clear why one should expect to observe behavior in accordance with Bayesian equilibrium.

A Proofs

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 D s 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,

$$l \geq \frac{2}{\beta+1}k. \quad (5)$$

This number of mistakes 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 D s 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 H s 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 D s go out of $h_{1/\alpha}$, new D s are inserted, and when some of the old H s go out of h_β , new H s 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 any further trembling.

Now assume Player 2 type β makes l trembles to H . If player 1 type $1/\alpha$ samples all l of the H s, 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,

$$l \geq \frac{2}{\alpha+1}k. \quad (6)$$

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 $1/\beta$ entering, and then in m rounds with types $1/\alpha$ and β).

So, the resistance in the transition from (HH, DD) to (DH, DH) is,¹⁴

$$\rho_{HH,DD \rightarrow DH,DH} = \min\left\{\frac{2}{\beta+1}k, \frac{2}{\alpha+1}k\right\}.$$

By symmetry it follows that $\rho_{DD,HH \rightarrow DH,DH} = \min\left\{\frac{2}{\alpha+1}k, \frac{2}{\beta+1}k\right\}$, 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 1 type $1/\alpha$ trembles 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,

$$l \geq \frac{\beta-1}{\beta+1}k. \quad (7)$$

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 entering, and then m rounds with types $1/\alpha$ and β).

Now assume player 2 type β trembles 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

$$l \geq \frac{\alpha-1}{\alpha+1}k. \quad (8)$$

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

$$\rho_{DH,DH \rightarrow HH,DD} = \min\left\{\frac{\beta-1}{\beta+1}k, \frac{\alpha-1}{\alpha+1}k\right\}.$$

¹⁴Note 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, $\rho_{DH, DH \rightarrow DD, HH} = \min\{\frac{\alpha-1}{\alpha+1}k, \frac{\beta-1}{\beta+1}k\}$.

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 H s, then the only possible best reply of player 1 type $1/\alpha$ is D , so $h_{1/\alpha}$ is filled up with D s at no further cost, so it can only work to make the transition cheaper than the D s, that occur because of trembling, are to be found in h_α .

So, assume that from (HH, DD) , player 1 type α makes l trembles to D . From (5) 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 D s, 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 D s 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 2s of 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 D s 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

$$l \geq \frac{\beta - 1}{\beta + 1}k. \quad (9)$$

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,

$$\frac{2}{\beta + 1}k + \frac{\alpha - 1}{\alpha + 1}k. \quad (10)$$

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 (5), whereas if $\beta > 3$, it is the minimum of the expressions in (9) and (10).

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 2 type $1/\beta$ makes l trembles to H . Then from (6), 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,

$$l \geq \frac{\alpha - 1}{\alpha + 1}k, \quad (11)$$

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,

$$\frac{2}{\alpha + 1}k + \frac{\beta - 1}{\beta + 1}k \quad (12)$$

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 (6), whereas if $\alpha > 3$, it is the minimum of the expressions in (11) and (12).

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,

$$\rho_{HH,DD \rightarrow DD,HH} = k \cdot \begin{cases} \min\left\{\frac{2}{\alpha+1}, \frac{2}{\beta+1}\right\} & \text{if } \alpha \leq 3, \beta \leq 3 \\ \min\left\{\frac{\alpha-1}{\alpha+1}, \frac{2}{\alpha+1} + \frac{\beta-1}{\beta+1}, \frac{2}{\beta+1}\right\} & \text{if } \alpha > 3, \beta \leq 3 \\ \min\left\{\frac{2}{\alpha+1}, \frac{2}{\beta+1} + \frac{\alpha-1}{\alpha+1}, \frac{\beta-1}{\beta+1}\right\} & \text{if } \alpha \leq 3, \beta > 3 \\ \min\left\{\frac{\alpha-1}{\alpha+1}, \frac{2}{\alpha+1} + \frac{\beta-1}{\beta+1}, \frac{2}{\beta+1} + \frac{\alpha-1}{\alpha+1}, \frac{\beta-1}{\beta+1}\right\} & \text{if } \alpha > 3, \beta > 3 \end{cases}$$

which is equivalent to,

$$\rho_{HH,DD \rightarrow DD,HH} = k \cdot \begin{cases} \min\left\{\frac{2}{\alpha+1}, \frac{2}{\beta+1}\right\} & \text{if } \alpha \leq 3, \beta \leq 3 \\ \min\left\{\frac{\alpha-1}{\alpha+1}, \frac{2}{\beta+1}\right\} & \text{if } \alpha > 3, \beta \leq 3 \\ \min\left\{\frac{2}{\alpha+1}, \frac{\beta-1}{\beta+1}\right\} & \text{if } \alpha \leq 3, \beta > 3 \\ \min\left\{\frac{\alpha-1}{\alpha+1}, \frac{\beta-1}{\beta+1}\right\} & \text{if } \alpha > 3, \beta > 3 \end{cases}$$

From symmetry we get by permuting α and β in the expression above that $\rho_{DD,HH \rightarrow HH,DD} = \rho_{HH,DD \rightarrow DD,HH}$.

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,

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

Now note that $\rho_{DD,HH \rightarrow HH,DD} \geq \rho_{DD,HH \rightarrow DH,DH}$, which together with the symmetry property $\rho_{DH,DH \rightarrow DD,HH} = \rho_{DH,DH \rightarrow HH,DD}$ implies that there 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,

$$\gamma(HH, DD) = \rho_{DD,HH \rightarrow DH,DH} + \rho_{DH,DH \rightarrow HH,DD}.$$

By symmetry, $\gamma(DD, HH) = \rho_{HH,DD \rightarrow DH,DH} + \rho_{DH,DH \rightarrow DD,HH} = \gamma(HH, DD)$, where the last equality also uses $\rho_{HH,DD \rightarrow DH,DH} = \rho_{DD,HH \rightarrow DH,DH}$.

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

$$\gamma(DH, DH) = \rho_{HH,DD \rightarrow DH,DH} + \rho_{DD,HH \rightarrow DH,DH}.$$

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 $\rho_{DH,DH \rightarrow HH,DD} < \rho_{HH,DD \rightarrow DH,DH}$, and (DH, DH) , and only that convention, is stochastically stable if $\rho_{DH,DH \rightarrow HH,DD} > \rho_{HH,DD \rightarrow DH,DH}$ (where it is again used that k is sufficiently large).

Finally, from $\alpha \geq \beta$ it follows that $\rho_{DH,DH \rightarrow HH,DD} = \frac{\beta-1}{\beta+1}k$, 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

Sketch of Proof of Proposition 3. (A complete proof is available from the authors upon request).

We need to prove that from any state it has, according to the basic learning process Π^0 , positive probability to reach a convention in a final number of rounds. The proof of that involves the following steps:

1. From any state it has, according to Π^0 , positive probability to reach in a final number of rounds, a state in which all records are constant (meaning that each record on each type is a list of m times one and the same action), and where in particular the record on buyer type $\alpha = 0$ only contains the bid zero, and the record on seller type $\beta = 1$ only contains the bid one. This is true since from any state it has positive probability that in all of the next m rounds a buyer of type $\alpha = 0$, and a seller of type $\beta = 1$, are picked for play, and they bid 0 and 1 respectively (these are always the unique undominated best replies). After that, the state will have constant records on these two types. Now it has positive probability that in all of the following m rounds, the buyers are again of type $\alpha = 0$ and bid 0, and the sellers are of type $\beta = 0$, and sample in all m rounds exactly the same from the buyer records (note that this is possible since the buyer records do not change during the m rounds, since buyer $\alpha = 0$ bidding zero appears all the time), and hence play the same best reply in all m rounds. Then a state is reached in which the record on seller type $\beta = 0$ is also constant. Continuing in this way the process reaches, in at most $7m$ rounds, a state in which all records are constant.

2. From any state with the property described in 1, it has, according to Π^0 , positive probability to reach in a final number of rounds a state, also with this property, in which any non-zero bid appearing in a buyer record also appears in a seller record, and any bid different from one appearing in a seller record also appears in a buyer record. This holds because, obviously, for any buyer type, the set of best replies to samples drawn from constant seller records either contains the bid zero (if the buyer type does not want trade), or, if it does not contain zero, only contains actions which appear in the seller records. So, if in each of the next m rounds the seller is of type $\beta = 1$ and bids 1, and the buyer is of some particular type α , there is positive probability that in all m rounds, the buyer bids zero or some other bid contained in the seller records. By continuing with the other types of buyers, and then with the sellers, and perhaps repeating for some rounds, one will obtain that all bids appearing in buyer records are either zero or also appearing in seller records, and all bids appearing in seller records are either one or also appearing in buyer records. It is possible that the state corresponds to a zero-step equilibrium, so all seller

records only contain the bid one, and all buyer records only contain zero. It can also correspond to a one-, two-, or three-step equilibrium. Finally, it can be a state not corresponding to an equilibrium.

3. From any state with the property described in 2, and not corresponding to an equilibrium, it has, according to Π^0 , positive probability to reach in a final number of rounds a state corresponding to an equilibrium. Either the state already corresponds to an equilibrium, or there is some type who does not play a best reply to the unique sampling possible on the opponent. It has positive probability that in all of the next m rounds exactly this type is picked together with the redundant type of opponent, and will in all m rounds play one and the same best reply to the unique sampling possible. Then a state is reached which is an equilibrium or in which there is some other type not playing a best reply. Continuing in this way bids can disappear from records, but they cannot reappear, so with positive probability the process will eventually reach a state corresponding to a zero-, one-, two-, or three-step equilibrium.
4. From any state corresponding to an equilibrium, it has, according to Π^0 , positive probability to reach in a final number of rounds a state corresponding to a semi-strict equilibrium, that is, to reach a convention. This is shown through a series of lemmas:

Lemma 1 *From the state corresponding to the zero-step equilibrium where all buyer types bid zero, and all seller types bid one, there is positive probability of reaching a state corresponding to a one-step equilibrium.*

Lemma 2 *From a state corresponding to a one-step equilibrium with a bid below $\frac{1}{3}$ or a bid above $\frac{2}{3}$, there is positive probability of either reaching a state corresponding to a one-step equilibrium with a bid between $\frac{1}{3}$ and $\frac{2}{3}$, or to reach a state corresponding to a two-step equilibrium.*

Lemma 3 *From a state corresponding to a one-step equilibrium with a bid between $\frac{1}{3}$ and $\frac{2}{3}$, there is positive probability of reaching a state corresponding to a two-step equilibrium.*

Lemma 4 *From a state corresponding to a two-step equilibrium that is not semi-strict, there is positive probability of reaching either a state corresponding to a two-step equilibrium with a bid below $\frac{1}{3}$ and a bid above $\frac{2}{3}$, or a state corresponding to a three-step equilibrium.*

Lemma 5 *From a state corresponding to a two-step equilibrium with a bid below $\frac{1}{3}$ and a bid above $\frac{2}{3}$, there is positive probability of reaching a state (convention) corresponding to a semi-strict two-step equilibrium.*

Lemma 6 *From a state corresponding to a three-step equilibrium that is not semi-strict, there is positive probability of reaching a state (convention) corresponding to a semi-strict three-step equilibrium.*

Each of these lemmas 1-6 is proved by identifying a sequence of events which has positive probability according to Π^0 , that is, by identifying a specific sequence of types entering and playing specific best replies to possible samples, which leads to the desired outcome. We now show how this is done for For Lemma 1; the proofs of the other lemmas are similar in structure, but admittedly more complicated.

For Lemma 1, start in the state where all buyer records only contain zero, and all seller records only contain one. With positive probability (according to Π^0), in each of the next m rounds the seller is of type $\frac{2}{3}$ and bids one, and the buyer is of type 1 and bids $\frac{1}{2}$ (anything is a best reply for this type of buyer given the records on the seller). In each of the following m rounds with positive probability the buyer is of type $\frac{1}{3}$ and bids zero, and the seller is of type 0 and bids $\frac{1}{2}$ (the unique best reply), and in each of the then following m rounds with positive probability the buyers are all of type $\frac{1}{3}$ and bid 0, and the sellers are all of type $\frac{1}{3}$ and bid $\frac{1}{2}$ (the unique best reply). In the following m rounds with positive probability the sellers are of type $\frac{2}{3}$ and bid 1, and the buyers are of type $\frac{2}{3}$ and bid $\frac{1}{2}$ (the unique best reply), which leads the process to the state that corresponds to the one-step equilibrium with $y = \frac{1}{2}$.

Note that in this proof of Lemma 1, the redundant types were not used. In fact, all of the Lemmas 1-6 can be proved without letting the redundant types enter. These are only used to prove steps 1 - 3 above. In the complete proof, we show how 1 - 3 can also be proved without "sending in the redundant types", under a usual

assumption of not too much sampling, $k < \frac{m}{2}$. Basically this implies that it has positive probability that in the next k rounds only buyers of a specific type enter and all draw the same sample and play the same best reply, even if seller records do change during the k rounds. One can use this to arrange records "piece by piece".
 \square

Sketch of Proof of Theorem 2. (A complete proof is available from the authors upon request).

We need to identify the conventions with minimal stochastic potential. To do this we find for each convention h the smallest number of trembles required to "escape" the convention, that is, to reach from it another state from which it has positive probability, according to the basic process Π^0 , to reach a convention different from h . We call this smallest number of trembles the *escape resistance* from h . In Section 3 we defined the resistance in any transition hh' between two conventions. The escape resistance from h is just the minimal resistance over all transitions hh' . If the resistance in the transition hh' is the escape resistance we say that hh' is an *escape direction* from h .

We find escape resistances and escape directions by systematically going through all conventions, and for each convention going through all the bids it uses, for each such bid calculating how many trembles are (minimally) required to change the best reply of either a seller type or a buyer type using the bid to a higher or lower bid, and checking that from the state reached by this trembling it has positive probability, according to Π^0 , to go all the way to a new convention. By minimizing over all the bids used in a convention one finds the escape resistance of it. The procedure is exactly as the procedure we applied in the proof of Theorem 1 above. For conventions corresponding to two-step equilibria using bids $x < \frac{1}{3}$ and $z > \frac{2}{3}$ we find by this procedure,

Lemma 7 *On the minimal amount of trembling required to go to another convention by first shifting best replies away from x : (i) The number of trembles required to go from a convention corresponding to a two-step equilibrium to a new convention by changing some best reply away from x , is maximized for conventions with x as close as possible to $\frac{1}{4} \frac{n-1}{n}$. (ii) From conventions with x above $\frac{1}{4} \frac{n-1}{n}$, the minimal number of trembles is required when moving to another convention corresponding to*

a two-step which either uses bids $x - \frac{1}{n}$ and z , or (if x is sufficiently close to $\frac{1}{3}$) uses bids $\frac{x}{2}$ and z (if there is such a convention), or just uses one of the bids $\frac{x}{2}$ or z . From conventions with x below $\frac{1}{4} \frac{n-1}{n}$, the minimal number of trembles is required when moving to another convention corresponding to a two-step equilibrium which either uses bids $x + \frac{1}{n}$ and z , or (if $5x$ is sufficiently close to z) just uses the bid z .

Lemma 8 *About the minimal amount of trembling required to go to another convention by first shifting best replies away from z : (i) The number of trembles required to go from a convention corresponding to a two-step equilibrium to a new convention by changing some best reply away from z , is maximized for conventions with z as close as possible to $\frac{1}{4} \frac{3n-1}{n}$. (ii) From convention with z below $\frac{1}{4} \frac{3n-1}{n}$, the minimal number of trembles is required when moving to another convention corresponding to a two-step equilibrium which either uses bids x and $z + \frac{1}{n}$, or (if z is sufficiently close to $\frac{2}{3}$) uses bids x and $\frac{z}{2} + \frac{1}{2}$ (if there is such a convention), or just uses one of the bids x or $\frac{z}{2} + \frac{1}{2}$. From conventions with z above $\frac{1}{4} \frac{3n-1}{n}$, the minimal number of errors is required when moving to another convention corresponding to a two-step equilibrium which either uses bids x and $z - \frac{1}{n}$, or (if $5z - 4$ is sufficiently close to x) just uses the bid x .*

Thus, within the set of conventions corresponding to two-step equilibria, the convention with the highest escape resistance is h^* , whenever n is large enough. Further it can be concluded from the two lemmas, that for all the other conventions in the set, escape directions are towards h^* , see Figure 2. In fact, *on the set of conventions corresponding to two-step equilibria*, one can now construct an h^* -subtree that only involves transitions that are escape directions. Since h^* itself has maximal escape resistance, this subtree has the lowest possible total resistance among all subtrees on the set, see again Figure 2.

We then go through a similar procedure for the set of conventions corresponding to three-step equilibria. Unfortunately this case is a bit more complicated even beyond the facts that a three-step equilibrium involves three bids, and the set of possible equilibrium bids can therefore not be drawn in two dimensions. Keeping the bid z above $\frac{2}{3}$ fixed, the possible equilibrium combinations of a bid x below $\frac{1}{3}$, and a bid y between $\frac{1}{3}$ and $\frac{2}{3}$, as given by the restrictions (4) in Proposition 1, can

Figure 2: Escape directions for two-step equilibria

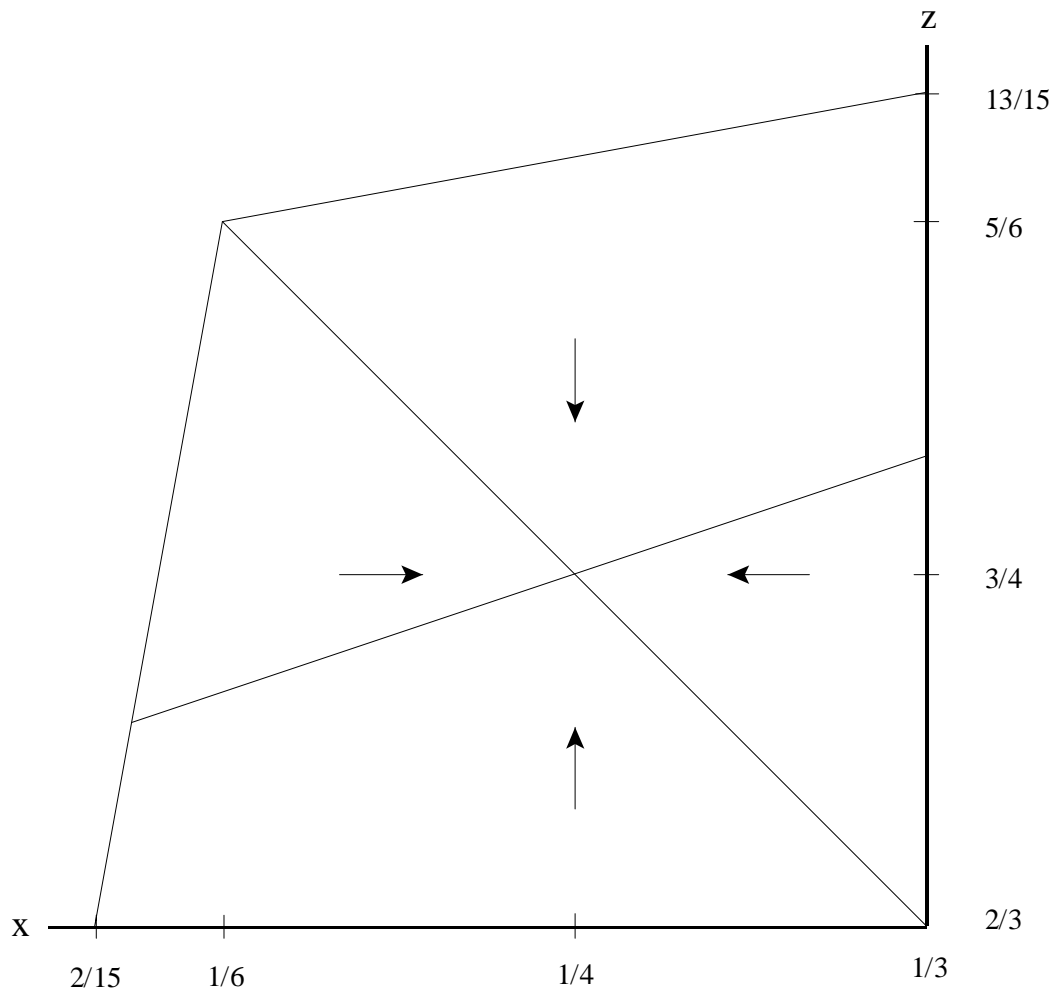
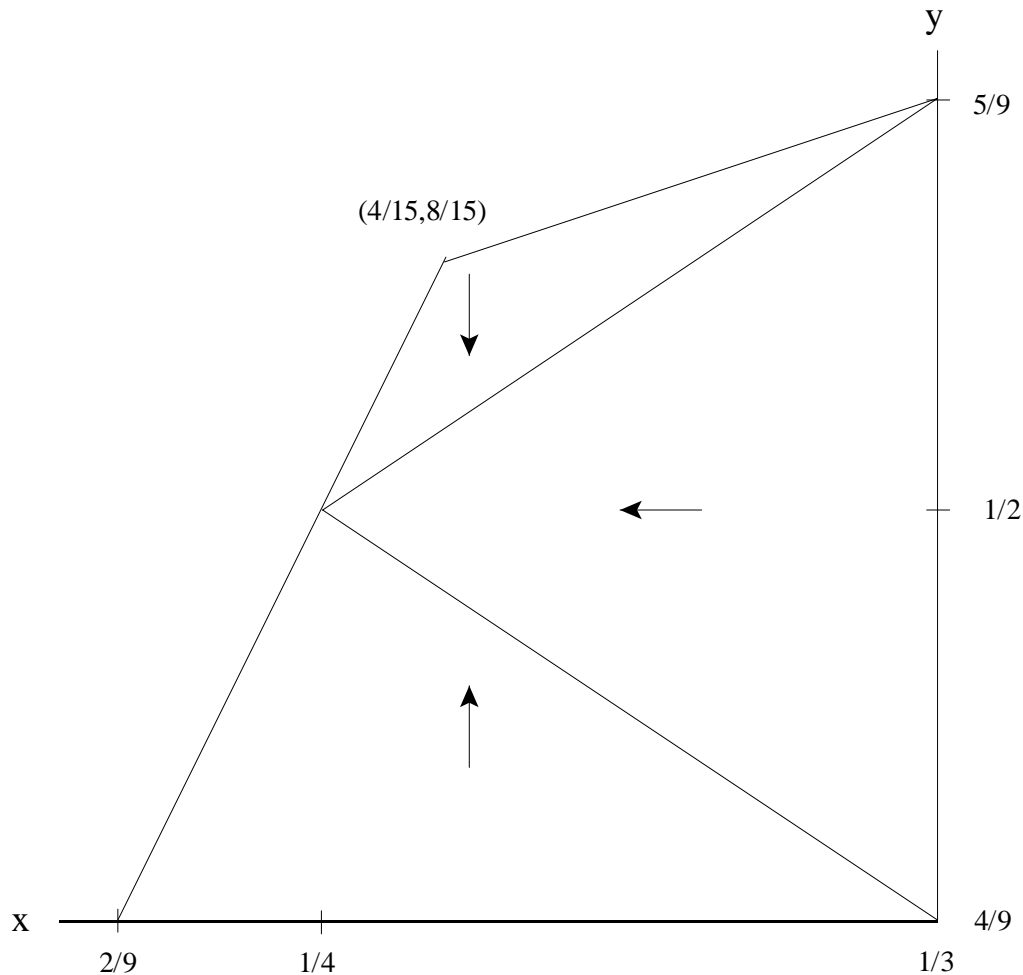


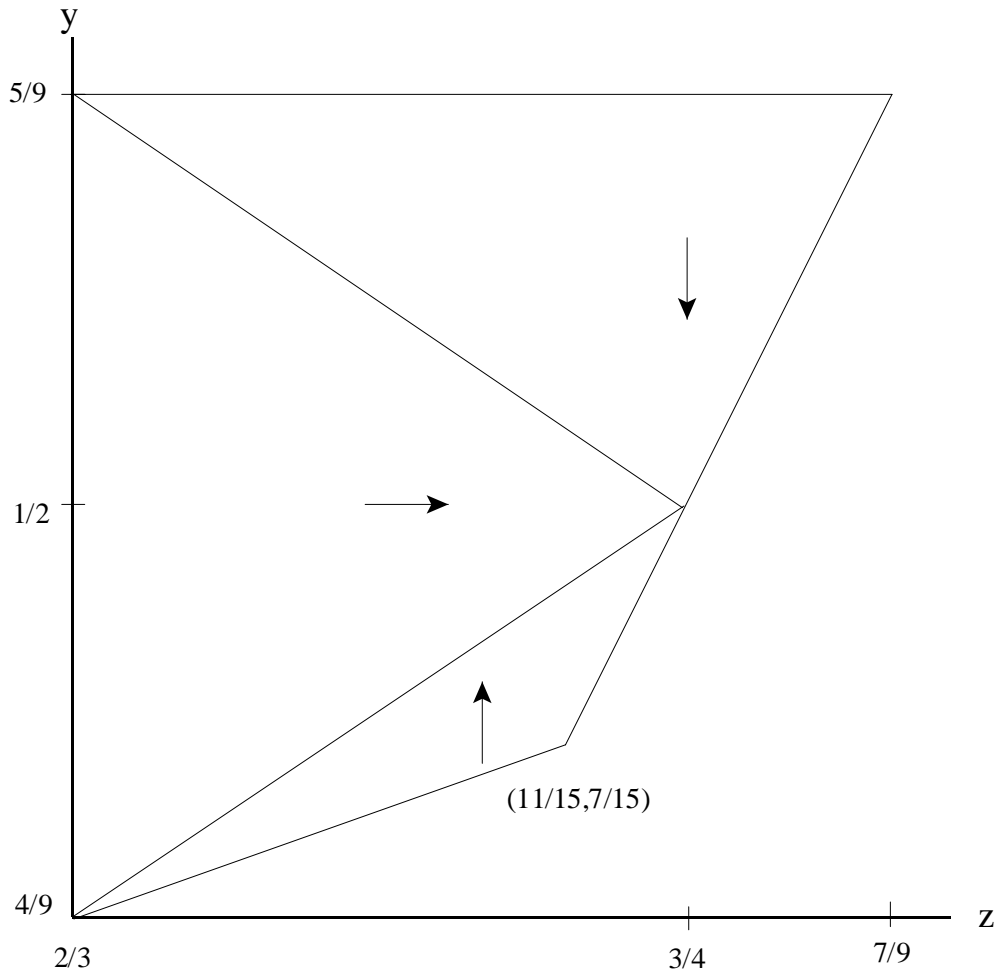
Figure 3: Escape directions in the (x, y) -dimension for three-step equilibria



be illustrated as in Figure 3 (not all of these combinations will form an equilibrium with any $z > \frac{2}{3}$).

Now systematic calculations similar to the ones performed for two-step equilibria reveal that the number of trembles required for going from one convention to another by first changing best replies away from x or y (or both) is largest close to the point $x = \frac{1}{4}, y = \frac{1}{2}$. To get away from the other conventions by as few trembles as possible one has to move in the directions indicated in Figure 3. Close to the boundary there may be several directions which require the same (low) number of errors. A similar analysis applies for getting away from a convention by first changing the value of y

Figure 4: Escape directions in the (y, z) -dimension for three-step equilibria



or z , see Figure 4.

Combining the above elements it is possible to see, that the maximal escape resistance for conventions corresponding to three-step equilibria is attained in the conventions with $(x, y, z) = (\frac{1}{4} + \frac{2}{n}, \frac{1}{2}, \frac{3}{4} - \frac{2}{n})$ and $(\frac{1}{4} + \frac{2}{n}, \frac{1}{2} + \frac{1}{n}, \frac{3}{4} - \frac{1}{n})$ and $(\frac{1}{4} + \frac{2}{n}, \frac{1}{2} + \frac{1}{n}, \frac{3}{4} - \frac{2}{n})$ and $(\frac{1}{4} + \frac{2}{n}, \frac{1}{2} - \frac{1}{n}, \frac{3}{4} - \frac{2}{n})$ and $(\frac{1}{4} + \frac{1}{n}, \frac{1}{2} - \frac{1}{n}, \frac{3}{4} - \frac{2}{n})$. Further, it is possible to partition the set of conventions corresponding to three-step equilibria into five different subsets each containing one of the conventions with maximal escape resistance, such that on each subset of the partition there is a subtree that

only involves transitions which are escape directions and the subtree leads exactly to the convention in the subset with maximal escape resistance.

We now have a partition of the full set of conventions into six classes (one with conventions corresponding to two-step equilibria, and five with conventions corresponding to three-step equilibria), such that in each class there is a convention with maximal escape resistance, and such that there is on the class an escape direction subtree leading to this convention. The six conventions with (local) maximal escape resistance we call the *candidates*.

The partition obtained is useful because of the following general result,

Lemma 9 *Any stochastically stable state is a candidate.*

This is true since if we have a full tree for some convention h that is not a candidate, then it is possible to construct a tree for the candidate in the class to which h belongs with lower total resistance. Just use the transitions of the h -tree for all conventions outside the class to which h belongs, and the transitions of an escape direction subtree for the class to which h belongs.

Now we use the subtrees (one on each class) to construct an h^* -tree with very low resistance. This tree consists of escape directions for all conventions except for the following: the two conventions corresponding to the three-step equilibria with bids $(x, y, z) = (\frac{1}{4} + \frac{1}{n}, \frac{1}{2} - \frac{1}{n}, \frac{3}{4} - \frac{1}{n})$ and $(\frac{1}{4} + \frac{2}{n}, \frac{1}{2}, \frac{3}{4} - \frac{1}{n})$, and the one candidate convention with bids $(x, y, z) = (\frac{1}{4} + \frac{2}{n}, \frac{1}{2} + \frac{1}{n}, \frac{3}{4} - \frac{1}{n})$. It can be calculated that for n sufficiently large the resistance in this tree is less than $(\frac{1}{7} + 0.001)k$ above the sum of all the escape resistances except those for h^* and for one of the five three-step candidates. Let K be this sum of escape resistances except for h^* and one three-step candidate.

Finally we show that for n sufficiently large any tree to a three-step candidate must have total resistance exceeding at least $0.15k + K$, which is larger than $(\frac{1}{7} + 0.001)k + K$, and the theorem follows. This is done by considering, for a given three-step candidate h , the different possibilities of constructing a path that connects h^* to h , and then construct an h -tree, where all the other transitions in the tree are escape transitions. It turns out, that it is impossible to construct an h -tree with total resistance less than $0.15k + K$.

If n or k is smaller than "sufficiently large", the h^* -tree constructed still has lowest possible total resistance, but there may be trees to other conventions with

this resistance, so in this case there need not be a unique stochastically stable state.

■

References

- [1] Canning, D. (1992a), Average Behavior in Learning Models, *Journal of Economic Theory*, **57**, 442-72.
- [2] Canning, D. (1992b), Language Conventions in Common Interest Signaling Games, *Discussion Paper # 607*, Columbia University.
- [3] Chatterjee, K. and W. Samuelson (1983), Bargaining under Incomplete Information, *Operations Research*, **31**, 835-851.
- [4] Harsanyi, J.C. and R. Selten (1988), *A General Theory of Equilibrium Selection in Games*, Cambridge: MIT Press.
- [5] Jacobsen, H.J., M. Jensen, and B. Sloth (1999), Evolutionary Learning in Signalling Games, *Discussion Paper 99-01*, Institute of Economics, University of Copenhagen. Forthcoming in *Games and Economic Behavior*.
- [6] Kandori, M., G.J. Mailath, and R. Rob (1993), Learning, Mutation, and Long Run Equilibria in Games, *Econometrica*, **61**, 29-56.
- [7] Leininger, W., P. Linhart, and R. Radner (1989), Equilibria of The Sealed-bid Mechanism for Bargaining with Incomplete Information, *Journal of Economic Theory*, **48**, 63-106.
- [8] Myerson, R.B. and M.A. Satterthwaite (1983), Efficient Mechanisms for Bilateral Trading, *Journal of Economic Theory*, **29**, 265-281.
- [9] Nöldeke, G. and L. Samuelson (1997), A Dynamic Model of Equilibrium Selection in Signalling Markets, *Journal of Economic Theory*, **73**, 118-156.
- [10] Satterthwaite, M.A. and S.R. Williams (1989), Bilateral Trade with the Sealed Bid k-Double Auction: Existence and Efficiency, *Journal of Economic Theory*, **48**, 107-133.

- [11] Young, H. Peyton (1993a), The Evolution of Conventions, *Econometrica*, **61**, 57-84.
- [12] Young, H. Peyton (1993b), An Evolutionary Model of Bargaining, *Journal of Economic Theory*, **59**, 145-168.