ACADEMIC PRESS

Games and Economic Behavior 40 (2002) 322–338

GAMES and Economic Behavior

www.academicpress.com

Note

A note on the one-deviation property in extensive form games

Andrés Perea

Departamento de Economía, Universidad Carlos III de Madrid, Calle Madrid 126, Getafe, Madrid, 28903 Spain

Received 7 December 2000

Abstract

In an extensive form game, an assessment is said to satisfy the *one-deviation property* if for all possible payoffs at the terminal nodes the following holds: if a player at each of his information sets cannot improve upon his expected payoff by deviating unilaterally at this information set only, he cannot do so by deviating at any arbitrary collection of information sets. Hendon et al. (1996. Games Econom. Behav. 12, 274–282) have shown that pre-consistency of assessments implies the one-deviation property. In this note, it is shown that an appropriate weakening of pre-consistency, termed *updating consistency*, is both a sufficient and necessary condition for the one-deviation property. The result is extended to the context of rationalizability. © 2002 Elsevier Science (USA). All rights reserved.

JEL classification: C72

Keywords: Extensive form games; Sequential rationality; One-deviation property; Rationalizability

1. Introduction

In dynamic one-person and multi-person decision making, the *one-deviation property* (also called *one-shot deviation principle*) reflects the phenomenon that a stream of "locally optimal" decisions constitutes a "globally optimal" decision stream. By "locally optimal," we mean that the decision for an individual at a particular stage maximizes his expected payoff, taking as given the decisions

E-mail address: perea@eco.uc3m.es.

PII: S0899-8256(02)00005-2

chosen at all other stages (including his own decisions at other stages). "Globally optimal" refers to the fact that the decision maker cannot improve upon his expected payoff by changing his decisions at any arbitrary subset of stages. The one-deviation property thus reflects a kind a time-consistency, stating that for optimal decision making it should be sufficient to check the optimality of each of the decisions on a one-by-one basis.

It is a well-known fact that the one-deviation property holds generally for the context of one-person decision making; a result which is known as the *optimality principle* for dynamic programming. If more than one decision maker is involved, the fact whether the one-deviation property holds or not depends crucially on the way decision makers (from now on called players) update their conjectures about the opponents' behavior as times passes by. It is the aim of this note to figure out which conditions on the players' updating behavior are necessary and sufficient in order for the one-deviation property to hold.

To this purpose, we focus on two different contexts which have both been important for the development of rationality concepts for extensive form games. In the first it is assumed that players, at each of their information sets, hold conjectures about the opponents' future behavior that coincide exactly with the "real" behavior of the opponents. The uncertainty of a player at an information set about the actual play of the game thus reduces to uncertainty about the past play, captured formally by the notion of *beliefs* at information sets. The conjecture about future play is completely determined by a fixed behavior strategy profile, which prescribes a randomization over actions at each information set. The above mentioned assumption implies that a player at an information set always believes that future play will be according to this strategy profile, also if the event of reaching this information set actually *contradicts* this strategy profile. This assumption is used in the backward induction concept for games with perfect information and most of the extensive form equilibrium refinements, such as subgame perfect equilibrium (Selten, 1965), sequential equilibrium (Kreps and Wilson, 1982), different versions of perfect Bayesian equilibrium and extensive form perfect equilibrium (Selten, 1975).

In this particular setting, the players' choices and conjectures about the play by opponents are represented by a so-called *assessment*: a combination of a behavior strategy profile and a system of beliefs at information sets. Consider an extensive form structure, that is, a combination of the game tree, the information sets, the actions and possibly chance moves, together specifying how the game is to be played. An extensive form structure is extended to an extensive form game by assigning a vector of payoffs to each of the terminal nodes. Formally, an assessment for a given extensive form structure is said to satisfy the *one-deviation property* if for every extensive form game having this extensive form structure the following holds: if a player at each of his information sets cannot improve upon his expected payoff by deviating unilaterally at this information set, while leaving his behavior at other information sets unchanged, he cannot do so by deviating at any arbitrary collection of information sets. The fact that this property should hold for *all extensive form games having this extensive form structure* implies that the one-deviation property puts restrictions on an assessment that solely depend on the extensive form structure, and not on the particular choice of payoffs at the terminal nodes.

In games with perfect information it is well-known that every strategy profile, together with the trivial beliefs at the singleton information sets, satisfies the one-deviation property. For games with imperfect information the *consistency* condition on assessments, which is part of Kreps and Wilson's definition of sequential equilibrium, turns out to be sufficient for the one-deviation property. Hendon et al. (1996) show that some weakening of consistency, termed *pre-consistency,* is enough to imply the one-deviation property. In Theorem 2.2 we prove that a further weakening, called *updating consistency*, is both sufficient and necessary for the one-deviation property to hold. Intuitively, updating consistency states that player i's conjecture at information set B about the opponents' behavior should be induced by his conjecture at information set A whenever B comes after A and the conjecture at A does not exclude reaching B. Important is that this condition should hold also if player i's *own strategy choice* at A prevents B from being reached.

The second context we focus on leaves more freedom to the players' conjectures about the opponents' behavior, since it is now no longer assumed that players hold correct conjectures about the opponents' future strategy choices. This more flexible setting corresponds to rationalizability concepts for extensive form games, such as *extensive form rationalizability* (Pearce, 1984; see also Battigalli, 1997), subgame perfect rationalizability (Bernheim, 1984) and weak extensive form rationalizability (Ben-Porath, 1997; Battigalli and Bonanno, 1999), among others. It also applies to "intermediate" models that place restrictions on the players' conjectures that are weaker than in the first context discussed above, but stronger than in rationalizability. For instance, the concept of self-confirming equilibrium (Fudenberg and Levine, 1993) requires the players' conjectures to coincide with the actual behavior on the equilibrium path, but allows them to differ from the actual behavior at unreached information sets. In Dekel et al. (1999, 2000), the concept of self-confirming equilibrium is refined to the case where conjectures about the opponents' behavior at unreached information sets should, in addition, be "rationalizable." Greenberg (1996) proposes a model in which the players' conjectures about the play are assumed to agree at some, but not necessarily all information sets, and defines a corresponding notion of stability. Within this context, players may thus have different conjectures about the play of the game at information sets for which no agreement is required.

As a primitive to model the players' conjectures about the opponents' behavior we use the notion of *updating systems* (cf. Battigalli, 1997), which specifies for each player and each information set controlled by this player a subjective randomization on the set of opponents' strategy profiles that are compatible with reaching this information set. In order to avoid the issue whether such randomizations should be correlated or uncorrelated, we restrict our attention to the case of two players. There should be no problem, however, in extending the result to games with more than two players, once it is decided which class of conjectures (correlated or uncorrelated) is to be used.

For a given extensive form structure, an updating system for a player is said to satisfy the one-deviation property if for all extensive form games having this extensive form structure and all strategies for this player the following holds: if at each of his information sets the player cannot improve upon his expected payoff by deviating at this information set only, given his conjecture about the opponent's behavior and given his decisions at other information sets, then he cannot improve by deviating at any arbitrary collection of information sets. We present a condition on updating systems, termed *updating consistency*, which is a weakening of the notion of *consistent* updating systems, used by Battigalli (1997). The intuition of updating consistency is the same as in the first context: if the player holds a certain conjecture at an information set A, then conjectures at future information sets should be derived from this by Bayesian updating, as long as reaching these information sets does not contradict the conjecture at A. What distinguishes it from consistent updating systems is that, unlike the latter, players are allowed to reshuffle conjectures at information sets as long as it does not affect the expected outcome conditional on reaching this information set. It thus leaves some more freedom than updating consistency. In Theorem 3.1 it is shown that updating consistency is both a necessary and sufficient condition for the one-deviation property.

The note is organized as follows. Section 2 deals with the context in which players are required to hold correct conjectures about the opponents' future behavior. It first provides some notation and definitions, and then presents the result which characterizes the assessments that satisfy the one-deviation property. Section 3 procedes identically for the context of updating systems.

2. One-deviation property for assessments

2.1. Notation in extensive form games

An *extensive form structure* $\mathcal{S}$ specifies a finite set of players, a finite game tree, a collection of information sets for each player, a set of actions at each information set and the probabilities of each of the chance moves. Let I be the set of players. For every $i \in I$, let H_i be the collection of information sets controlled by player i, and let H be the collection of all information sets in the game. For every $h \in H_i$ denote by $A(h)$ the set of actions available at h. We assume that $A(h)$ contains at least two actions for every h. Suppose that two actions available at different information sets are labelled differently, that is, $A(h) \cap A(h') = \emptyset$

if $h \neq h'$. It is assumed, moreover, that $\mathcal{S}$ satisfies *perfect recall* (Kuhn, 1953), which means that two different paths leading to the same player i information set h contain the same player i actions. Since actions at different information sets are, by assumption, different, perfect recall implies in particular that two paths leading to the same player i information set h pass through the same collection of "preceding" player i information sets. The set of terminal nodes is denoted by Z. An *extensive form game* is a pair $\Gamma = (\mathcal{S}, u)$ where $\mathcal{S}$ is an extensive form structure and u is the *payoff function* assigning to every terminal node $z \in Z$ a vector $u(z) = (u_i(z))_{i \in I} \in \mathbb{R}^I$ of payoffs.

2.2. Strategies and beliefs

A *behavior strategy* for player i is a vector $\sigma_i = (\sigma_{ih})_{h \in H_i}$ that assigns to every information set $h \in H_i$ some probability distribution σ_{ih} on $A(h)$. A vector $\sigma = (\sigma_i)_{i \in I}$ of behavior strategies is called a behavior strategy profile. A *belief system* is a vector $\beta = (\beta_h)_{h \in H}$ where β_h is a probability distribution on the set of nodes in h for all $h \in H$. A pair (σ, β) is called an *assessment*. Note that the set of assessments in a game depends only on the extensive form structure.

2.3. Sequential rationality

Let σ be a behavior strategy profile, x a node and $Z(x)$ the collection of terminal nodes that follow x. For every $z \in Z(x)$, let $\mathbb{P}_\sigma(z|x)$ be the probability that z is reached under σ, conditional on the event that the game has reached x. By $u_i(\sigma \mid x) = \sum_{z \in Z(x)} \mathbb{P}_\sigma(z|x) u_i(z)$ we denote the expected payoff for player i, conditional on x being reached. For a given assessment (σ, β) and an information set $h \in H_i$, let $u_i(\sigma \mid h, \beta_h) = \sum_{x \in h} \beta_h(x) u_i(\sigma|x)$ be the expected payoff for player i conditional on h being reached, given the beliefs β_h at h. The assessment (σ, β) is called *sequentially rational* if for every player i and every $h \in H_i$ it holds that $u_i(\sigma \mid h, \beta_h) = \max_{\sigma_i'} u_i((\sigma_i', \sigma_{-i}) \mid h, \beta_h)$. Here, (σ_i', σ_{-i}) is the behavior strategy profile in which player i plays σ_i' and the other players act according to σ. The assessment is called *locally sequentially rational* if for every player i and every $h \in H_i$ it holds that $u_i(\sigma \mid h, \beta_h) = \max_{\sigma_{ih}'} u_i((\sigma_{ih}', \sigma_{-h}) \mid h, \beta_h)$. Here, $(\sigma_{ih}', \sigma_{-h})$ is the behavior strategy profile in which player i plays the local strategy σ_{ih}' at information set h and σ is played at all other information sets (including the other player i information sets). The difference between sequential rationality and local sequential rationality is thus that the former takes into account all possible deviations by a player, whereas the latter concentrates on those deviations in which a player changes his behavior at only one information set.

2.4. One-deviation property

Let $\mathcal{S}$ be an extensive form structure and (σ, β) an assessment in $\mathcal{S}$. We say that (σ, β) satisfies the *one-deviation property* if for every payoff function u the following holds: (σ, β) is sequentially rational in the game $\Gamma = (\mathcal{S}, u)$ if and only if it is locally sequentially rational in Γ.

2.5. Updating consistency of assessments

In Hendon et al. (1996) it has been shown that the set of so-called *pre-consistent* assessments satisfies the one-deviation property. Their definition of pre-consistency consists of two parts. The first part, which we call *updating consistency*, states that a player should update his beliefs in some consistent manner to be specified below. The second part, called Bayesian consistency, is an equilibrium condition which assures that every player holds a correct conjecture about the opponents' past behavior at information sets reached with positive probability under σ. Since Bayesian consistency is not needed in their proof, it follows that the larger set of updating consistent assessments satisfies the one-deviation property as well.

Formally, an assessment (σ, β) is called *updating consistent* if for every player i, every two information sets $h^1, h^2 \in H_i$ where h^2 comes after h^1, and every behavior strategy σ_i^* for player i,

$$\beta_{h^2}(x) = \frac{\mathbb{P}_{(\sigma_i^*, \sigma_{-i})}(x \mid h^1, \beta_{h^1})}{\mathbb{P}_{(\sigma_i^*, \sigma_{-i})}(h^2 \mid h^1, \beta_h^1)}$$

for all $x \in h^2$, whenever $\mathbb{P}_{(\sigma_i^*, \sigma_{-i})}(h^2 \mid h^1, \beta_{h^1}) > 0$. Here, $\mathbb{P}_{(\sigma_i^*, \sigma_{-i})}(x \mid h^1, \beta_{h^1})$ is the probability that the node x is reached, conditional on h^1 being reached and given the beliefs β_{h^1} at h^1. By $\mathbb{P}_{(\sigma_i^*, \sigma_{-i})}(h^2 \mid h^1, \beta_h^1) = \sum_{y \in h^2} \mathbb{P}_{(\sigma_i^*, \sigma_{-i})}(y \mid h^1, \beta_{h^1})$ we denote the probability that h^2 is reached, conditional on h^1 being reached and given the beliefs at h^1. By perfect recall, every path from a node in h^1 to a node in h^2 contains the same player i actions. Consequently, the ratio in the definition of updating consistency does not depend on the particular choice of σ_i^*, as long as $\mathbb{P}_{(\sigma_i^*, \sigma_{-i})}(h^2 \mid h^1, \beta_{h^1}) > 0$.

The intuition behind updating consistency is the following. Consider two information sets h^1 and h^2 which are controlled by the same player i, and assume that h^2 comes after h^1. Player i's conjecture about the opponents' past behavior at h^1 is reflected by the beliefs β_{h^1}. If we assume that players hold correct conjectures about the opponents' future behavior, also at information sets which should actually have been avoided by σ, it follows that player i at h^1 believes that the opponents' future behavior is determined by σ_{-i}. Updating consistency states that player i's conjecture about the opponents' past behavior at h^2 should be induced by his conjectures about past and future behavior at h^1

whenever the event of reaching h^2 is compatible with his conjectures at h^1 (i.e., whenever there is some σ_i^* with $\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(h^2 \mid h^1, \beta_{h^1}) > 0$). Important is that this condition should also hold when $\mathbb{P}_\sigma(h^2 \mid h^1, \beta_{h^1}) = 0$. Hence, even if player i's own behavior after h^1 precludes the information set h^2 from being reached, his beliefs at h^2 should be induced by his conjecture about past and future behavior at h^1.

This property is satisfied in concepts such as sequential equilibrium and extensive form perfect equilibrium. The reason is that in both concepts, the players' beliefs are derived from taking a sequence of strictly positive behavior strategy profiles converging to the original one.[1] Along the sequence, it is clear that the beliefs of a player at two consecutive information sets are always in accordance with each other, since all information sets are reached with positive probability. As may be verified easily, this property remains valid in the limit, and hence every consistent assessment is updating consistent.

The following result is due to Hendon et al.

Theorem 2.1 (Hendon et al., 1996). *Let $\mathcal{S}$ be an extensive form structure. Then, every updating consistent assessment in $\mathcal{S}$ satisfies the one-deviation property.*

The theorem below shows that updating consistency is not only sufficient, but also necessary for the one-deviation property.

Theorem 2.2. *Let $\mathcal{S}$ be an extensive form structure. Then, an assessment (σ, β) in $\mathcal{S}$ satisfies the one-deviation property if and only if it is updating consistent.*

Proof. In view of Theorem 2.1, it suffices to show that every assessment which is not updating consistent fails to satisfy the one-deviation property. Let (σ, β) be an assessment in $\mathcal{S}$ which is not updating consistent. We show that there is a payoff vector u for the terminal nodes such that in the extensive form game $\Gamma = (\mathcal{S}, u)$ the assessment (σ, β) is locally sequentially rational but not sequentially rational.

Since (σ, β) is not updating consistent, there is some player i, two information sets $h^1, h^2 \in H_i$ where h^2 follows h^1, and some behavior strategy σ_i^* such that $\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(h^2 \mid h^1, \beta_{h^1}) > 0$ but

$$\beta_{h^2}(x^*) \neq \frac{\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(x^* \mid h^1, \beta_{h^1})}{\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(h^2 \mid h^1, \beta_{h^1})}$$

[1] These beliefs play an explicit role in sequential equilibrium, whereas used implicitly in extensive form perfect equilibrium.

for some node $x^* \in h^2$. Since both $\beta_{h^2}(\cdot)$ and $\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(\cdot \mid h^1, \beta_{h^1})/\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(h^2 \mid h^1, \beta_{h^1})$ are probability distributions on the set of nodes at h^2, we can choose x^* such that

$$\beta_{h^2}(x^*) < \frac{\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(x^* \mid h^1, \beta_{h^1})}{\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(h^2 \mid h^1, \beta_{h^1})}. \tag{2.1}$$

The reader may verify that h^1 and h^2 can always be chosen in such a way that Eq. (2.1) holds and there is no further player i information set between h^1 and h^2.

By perfect recall, there is a unique sequence $h_1, \ldots, h_K$ of player i information sets with the following properties: (1) h_k follows h_{k-1} for all k, (2) there is no player i information set between h_{k-1} and h_k for all k, (3) there is no player i information set before h_1, and (4) $h_{K-1} = h^1$ and $h_K = h^2$. We define the player i payoffs following h_k by induction on k.

We first define the player i payoffs following $h_K = h^2$. Let a_K be some action at h_K with $\sigma_{ih^2}(a_K) < 1$. Such an action exists since by assumption there are at least two actions at h_K. For every terminal node z following node x^* (see (2.1)) and action a_K, set $u_i(z) = 1$. For all terminal nodes z following action a_K but not following node x^*, set $u_i(z) = 0$. For every terminal node z following h_K but not following action a_K, set $u_i(z) = \beta_{h^2}(x^*)$.

Now, suppose that $k < K$ and that the player i payoffs $u_i(z)$ have been defined for all terminal nodes z following h_{k+1}. We define the player i payoffs following h_k but not following h_{k+1}in the following way. Let a_k be the unique action at h_k that leads to h_{k+1}. For every terminal node z following a_k but not following h_{k+1}, set $u_i(z) = 0$. By (a_k, σ_{-h_k}) we denote the strategy profile in which player i chooses action a_k with probability one at h_k, and players act according to σ at all other information sets. Let $u_i((a_k, \sigma_{-h_k}) \mid h_k, \beta_{h_k})$ be the expected payoff induced by (a_k, σ_{-h_k}) at h_k, given the beliefs β_{h_k} and the payoffs following a_k, which have already been defined above. For all terminal nodes z following h_k but not action a_k, we set $u_i(z) = u_i((a_k, \sigma_{-h_k}) \mid h_k, \beta_{h_k})$.

Finally, for all terminal nodes not covered by the procedure above, we set $u_i(z) = 0$. For all players $j \neq i$, we set $u_j(z) = 0$ for all terminal nodes z.

It can be verified easily that the assessment (σ, β) is locally sequentially rational, given the payoff vector u. Note that the payoffs are constructed in such a way that at every information set h_k, for $k = 1, \ldots, K$, player i is indifferent between action a_k and all other actions available at h_k, given his beliefs β_{h_k}, and given σ_{-h_k}.

If a player i information set h does not belong to $\{h_1, \ldots, h_K\}$, then, by construction of the payoffs, for every node $x \in h$ all payoffs following x are equal, and hence local sequential rationality follows trivially.

We finally show that (σ, β) is not sequentially rational at $h^1 = h_{K-1}$. For every action a at h^1 we have, by construction of the payoffs following h^1, that $u_i((a, \sigma_{-h^1}) \mid h^1, \beta_{h^1}) = u_i((a_{K-1}, \sigma_{-h^1}) \mid h^1, \beta_{h^1})$. Hence, $u_i(\sigma \mid h^1, \beta_{h^1}) =$

$u_i((a_{K-1}, \sigma_{-h^1}) \mid h^1, \beta_{h^1})$. After choosing a_{K-1} at h^1, the only feasible payoffs different from zero are the ones following h^2. By definition of the payoffs following $h^2 = h_K$, we have that $u_i((a_{K-1}, \sigma_{-h^1}) \mid h^1, \beta_{h^1})$ equals

$$\begin{aligned} &\mathbb{P}_{(a_{K-1},\sigma_{-h^1})}\big(x^* \bigm| h^1, \beta_{h^1}\big)\sigma_{ih^2}(a_K)1 \\ &\quad + \mathbb{P}_{(a_{K-1},\sigma_{-h^1})}\big(h^2 \bigm| h^1, \beta_{h^1}\big)\big(1 - \sigma_{ih^2}(a_K)\big)\beta_{h^2}(x^*). \end{aligned} \tag{2.2}$$

We may thus conclude that $u_i(\sigma \mid h^1, \beta_{h^1})$ is equal to (2.2).

Let σ_i' be the player i strategy defined as follows: (1) at h^2, it chooses with probability one the action a_K defined above, (2) at information set h^1 it chooses with probability one the action a_{K-1} leading to h^2, and (3) at all other player i information sets it coincides with σ_i. It can be verified that $u_i((\sigma_i', \sigma_{-i}) \mid h^1, \beta_{h^1})$ equals $\mathbb{P}_{(a_{K-1},\sigma_{-h^1})}(x^* \mid h^1, \beta_{h^1})$.

By Eq. (2.1), there exists a strategy σ_i^* with $\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(h^2 \mid h^1, \beta_{h^1}) > 0$. Since there is no player i information set between h^1 and h^2, and a_{K-1} is the unique action that leads from h^1 to h^2, it follows that $\mathbb{P}_{(a_{K-1},\sigma_{-h^1})}(h^2 \mid h^1, \beta_{h^1}) > 0$. We know that the ratio in (2.1) does not depend upon the choice of σ_i^*, as long as $\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(h^2 \mid h^1, \beta_{h^1}) > 0$. Hence,

$$\beta_{h^2}(x^*) < \frac{\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(x^* \mid h^1, \beta_{h^1})}{\mathbb{P}_{(\sigma_i^*,\sigma_{-i})}(h^2 \mid h^1, \beta_{h^1})} = \frac{\mathbb{P}_{(a_{K-1},\sigma_{-h^1})}(x^* \mid h^1, \beta_{h^1})}{\mathbb{P}_{(a_{K-1},\sigma_{-h^1})}(h^2 \mid h^1, \beta_{h^1})},$$

which implies that

$$\mathbb{P}_{(a_{K-1},\sigma_{-h^1})}\big(x^* \bigm| h^1, \beta_{h^1}\big) > \beta_{h^2}(x^*)\mathbb{P}_{(a_{K-1},\sigma_{-h^1})}\big(h^2 \bigm| h^1, \beta_{h^1}\big). \tag{2.3}$$

Since $\sigma_{ih^2}(a_K) < 1$, it follows from (2.3) that

$$u_i\big((\sigma_i', \sigma_{-i}) \bigm| h^1, \beta_{h^1}\big) = \mathbb{P}_{(a_{K-1},\sigma_{-h^1})}\big(x^* \bigm| h^1, \beta_{h^1}\big) > (2.2) = u_i\big(\sigma \bigm| h^1, \beta_{h^1}\big).$$

Hence, (σ, β) is not sequentially rational at h^1. □

3. One-deviation property for updating systems

In this section, we turn to the context in which players are no longer assumed to hold correct conjectures about the opponents' future behavior. As mentioned in the introduction, we restrict our attention to the case of two players. For the sake of convenience, we further assume that there are no chance moves. Before stating the result, we need some terminology.

3.1. Updating systems

A pure strategy for player i is a vector $s_i = (s_i(h))_{h \in H_i}$, where $s_i(h) \in A(h)$ for all $h \in H_i$. Let S_i be the set of pure strategies for player i. Every player i

holds at each of his information sets a conjecture about the opponent's behavior that is compatible with the event of reaching this information set. Such vectors of conjectures are called *updating systems* (cf. Battigalli, 1997). Formally, for every $h \in H$ and both players i, let $S_i(h) = \{s_i \in S_i \mid \exists s_j \in S_j$ such that (s_i, s_j) reaches $h\}$ be the set of player i strategies that are compatible with the event of reaching h. Here, we always assume that $i \neq j$. By perfect recall, it holds that a strategy profile (s_1, s_2) reaches h if and only if $s_1 \in S_1(h)$ and $s_2 \in S_2(h)$. An *updating system* for player i is a vector $c_i = (c_{ih})_{h \in H_i}$ where c_{ih} is a probability distribution on $S_j(h)$ for every $h \in H_i$.

3.2. Sequential rationality

Let s_i be a player i strategy and $h \in H_i$. By $s_i|_h$ we denote the strategy that at every $h' \in H_i$ preceding h chooses the unique action at h' leading to h, and at all other information sets coincides with s_i. By construction, $s_i|_h \in S_i(h)$. We say that the strategy s_i is *sequentially rational* with respect to the updating system c_i if for all $h \in H_i$ it holds that $u_i(s_i|_h, c_{ih}) = \max_{s_i'} u_i(s_i'|_h, c_{ih})$. Here, $u_i(s_i|_h, c_{ih})$ is the expected payoff induced by $s_i|_h$ and c_{ih}. For every $s_i \in S_i$, information set $h \in H_i$ and action $a \in A(h)$, let (a, s_{i-h}) be the player i strategy that chooses a at h and coincides with s_i at all other information sets. We say that s_i is *locally sequentially rational* with respect to the updating system c_i if for all $h \in H_i$ it holds that $u_i(s_i|_h, c_{ih}) = \max_{a \in A(h)} u_i((a, s_{i-h})|_h, c_{ih})$.

3.3. One-deviation property

Let $\mathcal{S}$ be an extensive form structure and c_i an updating system for player i in $\mathcal{S}$. We say that c_i satisfies the *one-deviation* property if for every payoff function u and every strategy $s_i \in S_i$ the following holds: s_i is sequentially rational with respect to c_i in the game $\Gamma = (\mathcal{S}, u)$ if and only if s_i is locally sequentially rational with respect to c_i in Γ.

3.4. Updating consistency

Let $\mu^1, \mu^2 \in \Delta(S_i)$ be two mixed strategies for player i and let $T_j \subseteq S_j$. We say that μ^1 is *equivalent* to μ^2 on T_j if for every $s_j \in T_j$ the probability distributions on the terminal nodes induced by (μ^1, s_j) and (μ^2, s_j) are identical. Let c_i be an updating system and $h^1, h^2 \in H_i$ be such that h^2 follows h^1 and $c_{ih^1}(S_j(h^2)) > 0$. Here, we use the convention $c_{ih^1}(S_j(h^2)) = \sum_{s_j \in S_j(h^2)} c_{ih^1}(s_j)$. By $c_{ih^1|h^2}$ we denote the conditional probability distribution on $S_j(h^2)$ given by

$$c_{ih^1|h^2}(s_j) = \frac{c_{ih^1}(s_j)}{c_{ih^1}(S_j(h^2))}$$

for all $s_j \in S_j(h^2)$. We say that an updating system c_i is *updating consistent* if for every two information sets $h^1, h^2 \in H_i$ where h^2 follows h^1 and $c_{ih^1}(S_j(h^2)) > 0$ it holds that c_{ih^2} is equivalent to $c_{ih^1|h^2}$ on $S_i(h^2)$.

The intuition of updating consistency is basically the same as in Section 2: if player i's conjecture at h^1 about the opponent's behavior is compatible with the event of reaching h^2, then his conjecture at h^2 should be induced by his conjecture at h^1, up to "inessential differences." By the latter we mean that player i, when updating his conjecture, is allowed to shift weight from one opponent's strategy to some other, as long as it does not affect the expected outcome conditional on h^2 being reached.

Updating consistency is somewhat weaker than the notion of *consistent* updating systems, as used by Battigalli (1997). An updating system c_i is called *consistent* if for all $h^1, h^2 \in H_i$ it holds that c_{ih^2} is equal to $c_{ih^1|h^2}$ whenever h^2 comes after h^1 and $c_{ih^1}(S_j(h^2)) > 0$. Clearly, consistency implies updating consistency, but the reverse is not true.

In order to illustrate the difference between updating consistency and consistency, consider the example in Fig. 1.

Let h^1, h^2 be the first and the second information set controlled by player 2, respectively. Let $c_2 = (c_{2h^1}, c_{2h^2})$ be player 2's updating system given by $c_{2h^1} = \frac{1}{2}(a,e,g,k) + \frac{1}{2}(b,e,h,k)$ and $c_{2h^2} = (a,e,g,l)$. Here, $\frac{1}{2}(a,e,g,k) + \frac{1}{2}(b,e,h,k)$ denotes the probability distribution which assigns equal probability to the strategies (a,e,g,k) and (b,e,h,k). The updating system is not consistent, since $c_{2h^1|h^2} = (a,e,g,k) \neq c_{2h^2}$. However, the updating system is updating consistent since $c_{2h^1|h^2}$ and c_{2h^2} are equivalent if player 2 chooses from $S_2(h^2)$, that is, if player 2 plays c.

Beforing stating the theorem, we briefly outline how the setup could be generalized to games with more than two players. In this case, player i's updating system would be a vector $c_i = (c_{ih})_{h \in H_i}$ assigning to every information set $h \in H_i$ some probability distribution $c_{ih} \in \Delta(S_{-i}(h))$ on the set of opponents' strategy

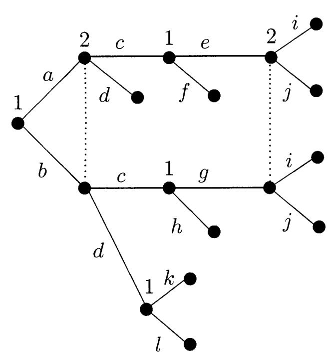

Fig. 1.

profiles $S_{-i}(h)$ leading to h. If we assume that conjectures are *uncorrelated* then c_{ih} should be the product of probability distributions on the opponents' individual strategy spaces. If conjectures are allowed to be *correlated* then c_{ih} may be any probability distribution on $S_{-i}(h)$. The remaining concepts can be generalized in a straightforward fashion to games with more than two players in both the correlated and the uncorrelated case. (See also Battigalli, 1997).

Theorem 3.1. *Let $\mathcal{S}$ be an extensive form structure with two players. Then, an updating system in $\mathcal{S}$ satisfies the one-deviation property if and only if it is updating consistent.*

Proof. (a) We first show that every updating system which is updating consistent satisfies the one-deviation property. Let the updating system c_i be updating consistent and let the strategy s_i be locally sequentially rational with respect to c_i in some game $\Gamma = (\mathcal{S}, u)$. We show that s_i is sequentially rational with respect to c_i. Let s_i' be an arbitrary pure strategy for player i. We prove that

$$u_i(s_i'|_h, c_{ih}) \leqslant u_i(s_i|_h, c_{ih}) \tag{3.1}$$

at every information set $h \in H_i$. We procede by induction on the number of player i information sets that follow h.

If h is not followed by any other information set of player i then the above inequality holds by local sequential rationality and the observation that $u_i(s_i'|_h, c_{ih})$ depends only on the action prescribed by s_i' at h. Now, let $k \in \mathbb{N}$ and assume that (3.1) holds for all player i information sets that are followed by at most k other player i information sets. Let $h \in H_i$ be followed by at most $k+1$ player i information sets. Let $H_i^*(s_i')$ be the set of player i information sets h' with the following properties: (1) h' follows h, (2) $s_i'|_h \in S_i(h')$ and (3) there is no player i information set between h and h'.

Let $S_j^0(h, s_i')$ be the set of strategies $s_j \in S_j(h)$ for which $(s_i'|_h, s_j)$ does not reach any $h' \in H_i^*(s_i')$. For every $s_j \in S_j(h) \backslash S_j^0(h, s_i')$, the strategy profile $(s_i'|_h, s_j)$ reaches exactly one $h' \in H_i^*(s_i')$. Using perfect recall, it may be checked that $(s_i'|_h, s_j)$ reaches $h' \in H_i^*(s_i')$ if and only if $s_j \in S_j(h')$. Moreover, we claim that the sets $S_j(h')$ are disjoint for $h' \in H_i^*(s_i')$. In order to see this, assume that $s_j \in S_j(h^1) \cap S_j(h^2)$ for two different $h^1, h^2 \in H_i^*(s_i')$. Hence, there exist s_i^1, s_i^2 such that (s_i^1, s_j) reaches h^1 and (s_i^2, s_j) reaches h^2. By construction of the set $H_i^*(s_i')$, all paths to h^1 and h^2 pass through h and contain the action prescribed by s_i' at h. Hence, all paths to h^1 and h^2 contain the same sequence of player i actions. But this implies that (s_i^1, s_j) and (s_i^2, s_j) should lead to the same information set in $H_i^*(s_i')$, which is a contradiction. We may therefore conclude that every $s_j \in S_j(h)$ either belongs to $S_j^0(h, s_i')$ or belongs to exactly one $S_j(h')$

with $h' \in H_i^*(s_i')$. Consequently,

$$\begin{aligned} u_i(s_i'|_h, c_{ih}) &= \sum_{s_j \in S_j(h)} c_{ih}(s_j) u_i(s_i'|_h, s_j) \\ &= \sum_{h' \in H_i^*(s_i')} \sum_{s_j \in S_j(h')} c_{ih}(s_j) u_i(s_i'|_h, s_j) \\ &\quad + \sum_{s_j \in S_j^0(h, s_i')} c_{ih}(s_j) u_i(s_i'|_h, s_j) \\ &= \sum_{\substack{h' \in H_i^*(s_i') \\ c_{ih}(S_j(h'))>0}} c_{ih}\big(S_j(h')\big) \sum_{s_j \in S_j(h')} \frac{c_{ih}(s_j)}{c_{ih}(S_j(h'))} u_i(s_i'|_h, s_j) \\ &\quad + \sum_{s_j \in S_j^0(h, s_i')} c_{ih}(s_j) u_i(s_i'|_h, s_j) \\ &= \sum_{\substack{h' \in H_i^*(s_i') \\ c_{ih}(S_j(h'))>0}} c_{ih}\big(S_j(h')\big) u_i(s_i'|_h, c_{ih|h'}) \\ &\quad + \sum_{s_j \in S_j^0(h, s_i')} c_{ih}(s_j) u_i(s_i'|_h, s_j). \end{aligned}$$

Since c_i is updating consistent, we have that $u_i(s_i'|_h, c_{ih|h'}) = u_i(s_i'|_h, c_{ih'})$ for all $h' \in H_i^*(s_i')$ with $c_{ih}(S_j(h')) > 0$. Hence,

$$\begin{aligned} u_i(s_i'|_h, c_{ih}) &= \sum_{\substack{h' \in H_i^*(s_i') \\ c_{ih}(S_j(h'))>0}} c_{ih}\big(S_j(h')\big) u_i(s_i'|_h, c_{ih'}) \\ &\quad + \sum_{s_j \in S_j^0(h, s_i')} c_{ih}(s_j) u_i(s_i'|_h, s_j) \\ &= \sum_{h' \in H_i^*(s_i')} c_{ih}\big(S_j(h')\big) u_i(s_i'|_h, c_{ih'}) \\ &\quad + \sum_{s_j \in S_j^0(h, s_i')} c_{ih}(s_j) u_i(s_i'|_h, s_j). \end{aligned}$$

For all $h' \in H_i^*(s_i')$ it holds, by definition, that $s_i'|_h \in S_i(h')$, and hence $s_i'|_{h'} = s_i'|_h$ for all $h' \in H_i^*(s_i')$. Consequently,

$$u_i(s_i'|_h, c_{ih}) = \sum_{h' \in H_i^*(s_i')} c_{ih}\big(S_j(h')\big) u_i(s_i'|_{h'}, c_{ih'})$$

$$+ \sum_{s_j \in S_j^0(h, s_i')} c_{ih}(s_j) u_i(s_i'|_h, s_j). \tag{3.2}$$

Since every $h' \in H_i^*(s_i')$ is followed by at most k player i information sets, we know by induction assumption that $u_i(s_i'|_{h'}, c_{ih'}) \leqslant u_i(s_i|_{h'}, c_{ih'})$ for all $h' \in H_i^*(s_i')$. This implies that

$$u_i(s_i'|_h, c_{ih}) \leqslant \sum_{h' \in H_i^*(s_i')} c_{ih}\big(S_j(h')\big) u_i(s_i|_{h'}, c_{ih'}) + \sum_{s_j \in S_j^0(h, s_i')} c_{ih}(s_j) u_i(s_i'|_h, s_j). \tag{3.3}$$

Let s_i'' be the player i strategy which coincides with s_i' at h and coincides with s_i at all other information sets. It can be checked that $H_i^*(s_i') = H_i^*(s_i'')$ and that $S_j^0(h, s_i') = S_j^0(h, s_i'')$. Moreover, $u_i(s_i|_{h'}, c_{ih'}) = u_i(s_i''|_{h'}, c_{ih'})$ for all $h' \in H_i^*(s_i')$ and $u_i(s_i'|_h, s_j) = u_i(s_i''|_h, s_j)$ for all $s_j \in S_j^0(h, s_i'')$. Together with (3.3) we obtain that

$$\begin{aligned} u_i(s_i'|_h, c_{ih}) &\leqslant \sum_{h' \in H_i^*(s_i'')} c_{ih}\big(S_j(h')\big) u_i(s_i''|_{h'}, c_{ih'}) \\ &\quad + \sum_{s_j \in S_j^0(h, s_i'')} c_{ih}(s_j) u_i(s_i''|_h, s_j) \\ &= u_i(s_i''|_h, c_{ih}), \end{aligned}$$

where the last equality follows from substituting s_i' by s_i'' in (3.2). Since s_i'' differs only at h from s_i and s_i is locally sequentially rational with respect to c_i, it holds that $u_i(s_i''|_h, c_{ih}) \leqslant u_i(s_i|_h, c_{ih})$. This implies that $u_i(s_i'|_h, c_{ih}) \leqslant u_i(s_i|_h, c_{ih})$, which completes the proof of (a).

(b) Next, we prove that every updating system which is not updating consistent does not satisfy the one-deviation property. Let c_i be an updating system in $\mathcal{S}$ which is not updating consistent. We show that there is a payoff vector u for the terminal nodes and a pure strategy s_i^* for player i such that s_i^* is locally sequentially rational with respect to c_i in the extensive form game $\Gamma = (\mathcal{S}, u)$ but not sequentially rational.

Since c_i is not updating consistent, there are $h^1, h^2 \in H_i$ where h^2 follows h^1 and $c_{ih^1}(S_j(h^2)) > 0$ such that c_{ih^2} is not equivalent to $c_{ih^1|h^2}$ on $S_i(h^2)$. Hence, there is some $s_i \in S_i(h^2)$ such that the probability distributions on the terminal nodes induced by (s_i, c_{ih^2}) and $(s_i, c_{ih^1|h^2})$ are different. Let these probability distributions be denoted by $\mathbb{P}_{(s_i, c_{ih^2})}$ and $\mathbb{P}_{(s_i, c_{ih^1|h^2})}$, respectively. We thus can find a terminal node z^* with

$$\mathbb{P}_{(s_i, c_{ih^2})}(z^*) < \mathbb{P}_{(s_i, c_{ih^1|h^2})}(z^*). \tag{3.4}$$

The above inequality implies that $\mathbb{P}_{(s_i, c_{ih^1|h^2})}(z^*) > 0$, and hence s_i necessarily chooses all the player i actions on the path to z^*. Since $s_i \in S_i(h^2)$ and $c_{ih^1|h^2} \in \Delta(S_j(h^2))$, it follows that $(s_i, c_{ih^1|h^2})$ passes h^2 with probability one, and hence z^* follows information set h^2.

By perfect recall, there is a unique sequence $h_1, \ldots, h_K$ of player i information sets with the following properties: (1) h_k follows h_{k-1} for all k, (2) there is no player i information set between h_{k-1} and h_k for all k, (3) there is no player i information set before h_1 and (4) $h_K = h^2$.

For every $k < K$, let a_k be the unique action at h_k which leads to h_{k+1}, and let a_K be the unique action at $h_K = h^2$ which leads to z^*. Since we know that s_i chooses all the player i actions that lead to z^*, it holds that s_i chooses action a_k at h_k for all $k = 1, \ldots, K$.

Since h^1 precedes $h^2 = h_K$, it must hold that $h^1 = h_{k^*}$ for some $k^* \in \{1, \ldots, K-1\}$. Let b_{k^*} be some action different from a_{k^*} at h^1, and let b_K be some action different from a_K at h^2. Let s_i^* be the player i strategy which chooses b_{k^*} at h^1, chooses b_K at h^2 and coincides with s_i at all other information sets.

We now define the player i payoffs u_i following h_k by induction on k. We start with the player i payoffs following $h_K = h^2$. Set $u_i(z^*) = 1$. For every terminal node z following h_K but not following action a_K, set $u_i(z) = \mathbb{P}_{(s_i|_{h^2}, c_{ih^2})}(z^*)$. For every terminal node $z \neq z^*$ following action a_K, we set $u_i(z) = 0$.

Now, suppose that the player i payoffs $u_i(z)$ have been defined for all terminal nodes z following h_{k+1}. We define the player i payoffs following h_k but not following h_{k+1}in the following way. For every terminal node z following action a_k but not following h_{k+1}, set $u_i(z) = 0$. Let $(a_k, s^*_{i-h_k})$ be the player i strategy that chooses a_k at h_k and coincides with s_i^* at all other player i information sets. Since all player i payoffs following a_k have already defined, the expression $u_i((a_k, s^*_{i-h_k})|_{h_k}, c_{ih_k})$ is well-defined. For all terminal nodes z following h_k but not following action a_k, set $u_i(z) = u_i((a_k, s^*_{i-h_k})|_{h_k}, c_{ih_k})$.

Finally, for all terminal nodes not covered by the procedure above, we set $u_i(z) = 0$. For player j, we set $u_j(z) = 0$ for all terminal nodes z.

It may be verified easily that s_i^* is locally sequentially rational with respect to c_i in the game $\Gamma = (\mathcal{S}, u)$. The payoffs are constructed in such a way that at every information set h_k, with $k = 1, \ldots, K$, player i is indifferent between action a_k and any other action available, given that s_i^* is played at all other player i information sets, and given the conjecture c_{ih_k}. If h is a player i information set that follows h^2, we distinguish two cases.

If h lies on the path to z^*, then there is a unique action a^* at h that leads to z^*. By construction of the payoffs following action a_K at h^2, action a^* at h leads to payoff 1, if z^* is reached, and leads to payoff zero otherwise, and all other actions at h lead to payoff zero for sure. Since s_i^* coincides with s_i at h, and s_i chooses a^* here, we have that s_i^* chooses a^* at h, which is optimal.

If h does not lie on the path to z^*, then, by construction of the payoffs, all payoffs following h are equal, and hence local sequential rationality follows trivially.

If a player i information set h does not follow h^2 nor precede h^2, all payoffs following h are equal, and local sequential rationality follows trivially.

We finally show that s_i^* is not sequentially rational. To this purpose, we prove that $u_i(s_i|_{h^1}, c_{ih^1}) > u_i(s_i^*|_{h^1}, c_{ih^1})$. Since s_i^* chooses b_{k^*} at h^1, we have, by definition of the payoffs following b_{k^*}, that

$$u_i(s_i^*|_{h^1}, c_{ih^1}) = u_i\big((a_{k^*}, s_{i-h^1}^*)|_{h^1}, c_{ih^1}\big). \tag{3.5}$$

Note that $(a_{k^*}, s_{i-h^1}^*)$ chooses all the actions a_k, for $k = 1, \ldots, K-1$, that lead to information set h^2. Hence, by construction of the payoffs, the only terminal nodes which are feasible for $((a_{k^*}, s_{i-h^1}^*)|_{h^1}, c_{ih^1})$ and have payoffs different from zero, are the ones following h^2. Since $(a_{k^*}, s_{i-h^1}^*) \in S_i(h^2)$, the probability of $((a_{k^*}, s_{i-h^1}^*)|_{h^1}, c_{ih^1})$ reaching h^2 is equal to $c_{ih^1}(S_j(h^2))$. Recall that $(a_{k^*}, s_{i-h^1}^*)$ chooses b_K at h^2 and that all terminal nodes following b_K have payoff $\mathbb{P}_{(s_i|_{h^2}, c_{ih^2})}(z^*)$. Gathering all these insights leads to the observation that

$$u_i\big((a_{k^*}, s_{i-h^1}^*)|_{h^1}, c_{ih^1}\big) = c_{ih^1}\big(S_j\big(h^2\big)\big)\mathbb{P}_{(s_i|_{h^2}, c_{ih^2})}(z^*).$$

Since $s_i \in S_i(h^2)$ we have that $s_i|_{h^2} = s_i$. Together with (3.5) it implies that

$$u_i(s_i^*|_{h^1}, c_{ih^1}) = c_{ih^1}\big(S_j\big(h^2\big)\big)\mathbb{P}_{(s_i, c_{ih^2})}(z^*). \tag{3.6}$$

On the other hand, s_i chooses all the player i actions a_k, for $k = 1, \ldots, K-1$, that lead to h^2. Hence, the only terminal nodes feasible for $(s_i|_{h^1}, c_{ih^1})$ and having payoffs different from zero are the ones following h^2. Recall that s_i chooses a_K at h^2, and that the only terminal node following a_K with non-zero payoff is z^*, with $u_i(z^*) = 1$. Hence, $u_i(s_i|_{h^1}, c_{ih^1}) = u_i(s_i, c_{ih^1}) = \mathbb{P}_{(s_i, c_{ih^1})}(z^*)$. Since z^* follows h^2, it holds that only player j strategies $s_j \in S_j(h^2)$ can lead to z^*, and therefore $\mathbb{P}_{(s_i, c_{ih^1})}(z^*) = c_{ih^1}(S_j(h^2))\ \mathbb{P}_{(s_i, c_{ih^1|h^2})}(z^*)$. It follows that

$$u_i(s_i|_{h^1}, c_{ih^1}) = c_{ih^1}\big(S_j\big(h^2\big)\big)\mathbb{P}_{(s_i, c_{ih^1|h^2})}(z^*). \tag{3.7}$$

Since, by assumption, $c_{ih^1}(S_j(h^2)) > 0$, and $\mathbb{P}_{(s_i, c_{ih^2})}(z^*) < \mathbb{P}_{(s_i, c_{ih^1|h^2})}(z^*)$, it follows from (3.6) and (3.7) that $u_i(s_i^*|_{h^1}, c_{ih^1}) < u_i(s_i|_{h^1}, c_{ih^1})$. $\square$

Acknowledgments

This paper has been written while the author was visiting the Department of Quantitive Economics at Maastricht University. The author wishes to thank their hospitality. He would also like to thank an anonymous referee for his suggestions.

References

Battigalli, P., 1997. On rationalizability in extensive games. J. Econ. Theory 74, 40–61.

Battigalli, P., Bonanno, B., 1999. Recent results on belief, knowledge and the epistemic foundations of game theory. Res. Econ. 53, 149–225.

Ben-Porath, E., 1997. Rationality, Nash equilibrium, and backwards induction in perfect information games. Rev. Econ. Stud. 64, 23–46.

Bernheim, D., 1984. Rationalizable strategic behavior. Econometrica 52, 1007–1028.

Dekel, E., Fudenberg, D., Levine, D., 1999. Payoff information and self-confirming equilibrium. J. Econ. Theory 89, 165–185.

Dekel E., Fudenberg D., Levine D., 2000. Subjective Uncertainty over Behavior Strategies: A Correction. Harvard University.

Fudenberg, D., Levine, D., 1993. Self-confirming equilibrium. Econometrica 61, 523–545.

Greenberg, J., 1996. Stable (Incomplete) Agreements in Dynamic Games: Worlds Apart—but Acting Together. McGill University.

Hendon, E., Jacobsen, J., Sloth, B., 1996. The one-shot-deviation principle for sequential rationality. Games Econ. Behav. 12, 274–282.

Kreps, D., Wilson, R., 1982. Sequential equilibria. Econometrica 50, 863–894.

Kuhn, H.W., 1953. Extensive form games and the problem of information. Ann. Math. Stud. 28, 193–216.

Pearce, D., 1984. Rationalizable strategic behavior and the problem of perfection. Econometrica 52, 1029–1050.

Selten, R., 1965. Spieltheoretische Behandlung eines Oligopolmodells mit Nachfragezeit. Z. Gesam. Staatswiss. 121, 301–324, 667–689.

Selten, R., 1975. Reexamination of the ‘perfectness concept for equilibrium points in extensive games.’ Internat. J. Game Theory 4, 25–55.