# On Forward Induction and Evolutionary and Strategic Stability[*]

Esther Hauk      Sjaak Hurkens

Department of Economics, Universitat Pompeu Fabra,

Ramon Trias Fargas 25-27, 08005 Barcelona, Spain

First version July 1999, this version September 1999

## Abstract

We analyze which normal form solution concepts capture the notion of forward induction, as defined by van Damme (JET, 1989) in the class of generic two player normal form games preceded by an outside option. We find that none of the known strategic stability concepts (including Mertens stable sets and hyperstable sets) captures this form of forward induction. On the other hand, we show that the evolutionary concept of EES set (Swinkels, JET, 1992) is always consistent with forward induction.

**Keywords:** Forward induction, strategic stability, EES sets, index theory.

# 1  Introduction

The purpose of this paper is to analyze which normal form solution concepts capture the extensive form notion of forward induction, as defined by van Damme (1989), in the class of generic two player normal form games preceded by an outside option. In this class of games player 1 first decides between 'In' and 'Out'. If he chooses 'Out' he receives his outside option and the game ends. If he chooses 'In' both players engage in a simultaneous move game. We restrict ourselves to those games for which the subgame has several Nash equilibria but only one of them yields player 1 a higher payoff than his outside option. Clearly, any game in this class has two Nash equilibrium components or outcomes: One component consists of those strategy profiles where player 1 chooses 'Out' and player 2 chooses any (mixed) strategy that makes player 1's choice optimal. The second component is a singleton that corresponds to player 1 choosing 'In' and both players continue in the subgame with the equilibrium preferred by player 1. Van Damme (1989) has argued that rational players would only play the second component using the following forward induction argument: When player 1 chooses to enter the subgame and thereby foregoes his sure outside option, player 2 should realize that a rational player 1 will: (1) play a Nash equilibrium in the subgame and (2) only enter the subgame if he expects to obtain a higher payoff than his outside option. Hence, the fact that the subgame has been entered unambiguously signals that play should continue with the equilibrium preferred by player 1. Consequently, player 1 will decide to enter the subgame. We will, henceforth, call player 1's preferred equilibrium the *forward induction equilibrium*.

Clearly, this forward induction logic is an *extensive form* argument. Several authors, and in particular Kohlberg and Mertens (1986), have forcefully argued that games should be analyzed in the *normal form* since the normal form represents all the strategically relevant information. In their seminal paper Kohlberg and Mertens list several desirable properties a good normal form solution concept should have and define three notions of stability (KM stability, full stability and hyperstability) that satisfy some, but not all, of those properties. One of those properties, which is satisfied by all three stability notions,

is motivated by an extensive form argument similar to van Damme's forward induction logic, as they write:

> "A subgame should not be treated as a separate game, because it was preceded by a very specific form of preplay communication - the play leading to the subgame." (Kohlberg and Mertens, 1986, p. 1013.)

They call this argument, as well as the property it motivates, "forward induction" but we prefer to call their mathematical formulation of the (normal form) property the never-weak-best-reply property (NWBR). It says that if $S$ is a stable set and strategy $s_j$ is never a weak best reply against $S$, then $S$ should contain a stable set of the reduced game where $s_j$ has been deleted.

In many outside option games stable sets (or any other solution concepts satisfying NWBR) do select the forward induction equilibrium. Consider for example the game in Fig. 1.

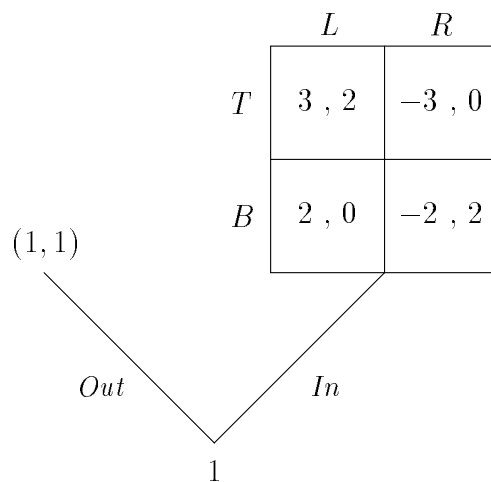|     | $L$       | $R$        |
| --- | --------- | ---------- |
| $T$ | $3$ , $2$ | $-3$ , $0$ |
| $B$ | $2$ , $0$ | $-2$ , $2$ |

$(1,1)$

*Out*        *In*

$1$

Fig. 1: NWBR selects the Forward Induction Equilibrium.

Consider the equilibria in the outside option component. In those equilibria player 1 chooses 'Out' and player 2 chooses $L$ with probability of at most $\frac{2}{3}$. Against this component the strategy $B$ is never a weak best response. (It yields at most $\frac{2}{3}$.) When we delete this strategy from the game, however, the only reasonable solution is of course $(T, L)$, which is not contained in the outside option component. The outside option component is therefore not stable. Hence, stability (or NWBR) selects the forward induction equilibrium in the example of Fig. 1.

Unfortunately, NWBR does not select the forward induction equilibrium in all outside option games. In fact, van Damme (1989) showed that the outside option outcome can be KM stable. We replicate his example in Fig. 2 below.

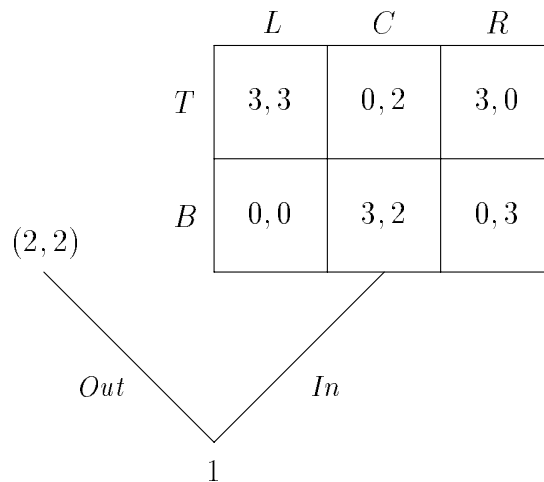|   | $L$ | $C$ | $R$ |
|---|-----|-----|-----|
| $T$ | $3,3$ | $0,2$ | $3,0$ |
| $B$ | $0,0$ | $3,2$ | $0,3$ |

$(2,2)$

*Out*          *In*

$1$

Fig. 2: Not every Stable Outcome is Consistent with Forward Induction.

The subgame has three Nash equilibria, one of which yields player 1 a payoff of 3, while the other equilibria yield a payoff strictly less than 2. Hence, the forward induction equilibrium is to enter the game and to play (T,L). This equilibrium is indeed stable. However, as van Damme (1989) shows, there exist two KM stable sets in which the outside option is chosen, namely $\{(Out, (\frac{2}{3}, \frac{1}{3}, 0)), (Out, (\frac{1}{3}, \frac{2}{3}, 0))\}$ and $\{(Out, (0, \frac{1}{3}, \frac{2}{3})), (Out, (0, \frac{2}{3}, \frac{1}{3}))\}$.[1]

---

[1] A KM stable set is a minimal set of Nash equilibria such that for any strategy perturbation of the game, there exists an equilibrium close to the set. In this case, the first (second) element of the set

As van Damme already pointed out, these sets are neither fully nor hyperstable. Whether the latter stability notions always do select the forward induction equilibrium remained an open question. We show, by means of a generic example in Section 3, that the outside option outcome may be hyperstable (and, therefore, also fully stable).[2] This is our main negative result.

More recently Mertens (1989, 1991) and Hillas (1990) introduced new stability concepts that satisfy all of the desirable properties listed by Kohlberg and Mertens.[3] Van Damme (1994) suggested, in a chapter prepared for the *Handbook of Game Theory with Economic Applications*, that perhaps Mertens stability implies forward induction. However, Hillas and Kohlberg (1996), in a different chapter of the same Handbook, showed this conjecture to be false. As a matter of fact, they used van Damme's game from Fig. 2 as their counterexample: The outside option component contains two Mertens stable sets, namely $\left\{\left(Out, (q, 1-q, 0) \mid \frac{1}{3} \leq q \leq \frac{2}{3}\right)\right\}$ and $\left\{\left(Out, (0, q, 1-q) \mid \frac{1}{3} \leq q \leq \frac{2}{3}\right)\right\}$. These Mertens stable sets are the convex hulls of the KM stable sets.[4] Since Mertens stable sets are known to contain Hillas stable sets (see e.g. Hillas *et al.* (1999)), the same negative result holds for Hillas stable sets.

The above results, together with the known relationships between all existing stability notions (including homotopy stable sets and Q-stable sets, see Hillas *et al.* (1999)), imply that none of the existing strategic stability notions uniquely selects the forward induction equilibrium. Given that the forward induction argument seems to require a high degree of sophistication from the players (players should play a Nash equilibrium and expect others to do likewise), strategic stability concepts presented themselves as prominent candidate solutions to capture this logic, but they fail to do so. Surprisingly, we do find

corresponds to strategy perturbations that put relatively high probability on 'T' ('B').

[2]Hyperstable sets are defined with respect to payoff perturbation, whereas fully stable sets are defined with respect to strategy perturbations. Kohlberg and Mertens (1986) showed that any hyperstable set contains a fully stable set, which in turn contains a KM stable set.

[3]Mertens stability is, like KM stability, defined with respect to strategy perturbations. Instead, Hillas stability is defined with respect to perturbations of the best response correspondence.

[4]This reflects the desirable connectedness property from Kohlberg and Mertens. The equilibria of the outside option component that are not contained in any of the above sets are not normal form perfect, and can therefore not be contained in any Mertens stable set.

4

an evolutionary solution concept that is always consistent with the forward induction logic, namely equilibrium evolutionarily stable sets (EES sets, Swinkels (1992)). This is the main positive result of the paper.

The remainder of the paper is set up as follows. In Section 2 we recall van Damme's definition of forward induction and argue that for a small subclass of outside option games the forward induction argument is not fully compelling. This subclass consists of those outside option games in which the preferred equilibrium is in mixed strategies and not all of player 2's best replies yield player 1 a payoff higher than the outside option. We introduce notation and derive some preliminary results relating to index theory that will prove to be helpful in the remainder of the paper. In Section 3 we present our negative results by means of two examples. The example that shows that the outside option outcome may be hyperstable has a forward induction equilibrium in mixed strategies. We provide a second example to show that even if the forward induction equilibrium is strict, the outside option outcome may be essential.[5] Section 4 contains the positive result that EES sets are always consistent with forward induction. Section 5 discusses the relation of forward induction and some adjustment dynamics and concludes.

## 2  Preliminaries

### 2.1  Forward Induction

Let $g = (E, F, u_1, u_2)$ be a generic 2-person game with pure strategy sets $E = \{e_1, \ldots, e_m\}$ and $F = \{f_1, \ldots, f_n\}$ and payoff functions $u_i : E \times F \to \mathbb{R}$. Let $s^*$ be player 1's preferred Nash equilibrium in $g$ and let $x$ be a number such that $u_1(s^*) > x > u_1(s)$ for any equilibrium $s \neq s^*$. Let $g^{out}$ denote the reduced normal form of the game where player 1 decides between an outside option, yielding him a payoff of $x$, or to enter the subgame $g$. We denote the class of these outside option games by , . Van Damme (1989) argues that any solution concept that is consistent with forward induction should at least

---

[5]Essentiality is, like hyperstability, defined with respect to payoff perturbations, but, unlike hyperstability, need not be invariant. Whether the outside option outcome can be hyperstable when the forward induction equilibrium is strict remains an open question.

satisfy the following property (see also Fudenberg and Tirole, 1993, Def. 11.8, p. 464):

**Property 1** *For any outside option game $g^{out} \in$ , for which $s^*$ is "viable", only the outcome in which player 1 chooses to enter the subgame $g$ and $s^*$ is played in $g$ is plausible.*

Which equilibria are viable depends on the theory of rationality that we are considering. E.g. if we believe players adhere to a theory that says that (weakly) dominated strategies are irrational, $s^*$ is only viable if it does not use (weakly) dominated strategies. Similarly, if we use stability as a selection criterion $s^*$ is only viable if it is itself stable.

We believe that van Damme's argument is very compelling. However, he defined property 1 for a slightly too large class of outside option games: the intuitive argument for forward induction does not work for all $g^{out} \in$ , . To illustrate this point, consider the outside option game in Fig. 3.

|       | $L$      | $C$      | $R$      |
|-------|----------|----------|----------|
| $T$   | 25 , 2   | 20 , 6   | 68 , 4   |
| $M$   | 33 , 20  | 32 , 12  | 48 , 17  |
| $B$   | 17 , 4   | 16 , 8   | 80 , 5   |

$(37\frac{1}{2}, 0)$
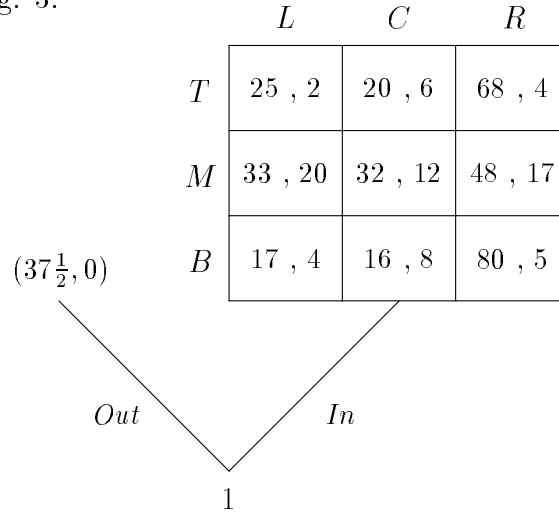
*Out*          *In*

1

Fig. 3: Outside Option Game.

The subgame $g$ has three equilibria, namely $s^1 = ((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}))$, yielding $(\frac{113}{3}, \frac{26}{3})$, $s^2 = ((\frac{3}{5}, \frac{2}{5}, 0), (\frac{5}{7}, 0, \frac{2}{7}))$ yielding $(\frac{261}{7}, \frac{46}{5})$ and $s^3 = ((0, 1, 0), (1, 0, 0))$ yielding $(33, 20)$. Since $\frac{113}{3} > 37.5 > \frac{261}{7} > 33$, player 1 prefers $s^1$. According to van Damme's notion of forward induction player 1, by entering the subgame, unambiguously signals that he will continue with the only equilibrium that yields more than the outside option, that is, with $s_1^1$, and that player 2 should play his part of the equilibrium, $s_2^1$.

It is true that player 1, by entering the subgame, unambiguously signals his intention to play his part of the preferred equilibrium. But, since that equilibrium is in *mixed*

6

strategies, player 2 has multiple pure best replies. Even if player 2 is convinced, by observing that player 1 entered the subgame, that player 1 will play his part of the preferred equilibrium, it is not at all obvious that player 2 will respond with his equilibrium strategies or any other mixed best reply that yields player 1 a *high* payoff. Since some of the alternative best replies of the game in Fig. 3 yield player 1 a payoff lower than the outside option, player 1 should not enter. The intuitive argument for forward induction fails here. However, if all best replies of player 2 yielded player 1 a payoff higher than the outside options, the intuitive argument for forward induction would be restored.

The example of Fig. 3 suggests that the class of games , should be restricted to those outside option games for which any mixed best reply of player 2 against $s^*$ yields player 1 a higher payoff than the outside option. We denote this restricted class by $\bar{,}$ .

## 2.2   Index theory

To prove some of our results in Sections 3 and 4 we will make use of index theory. Shapley (1974) introduced the index for isolated Nash equilibria and Ritzberger (1994) defined an index for Nash equilibrium components. We will rely on recent advancements by Govindan and Wilson (1996, 1997*a*,*b*) on the relation between the index and the degree of an equilibrium component. We here summarize the relevant results and explain how we will make use of them in the next sections.

With each component of Nash equilibria of a game are associated its index and its degree. The degree of a component is the local degree of the projection map from the Nash graph to the space of games.[6] The index of an equilibrium component is defined for a given map from the strategy space to itself whose fixed points are the equilibria of the game. The index does not depend on the particular fixed point map.[7] Moreover, the index and the degree of a component agree and depend only on the reduced normal form of the game (Govindan and Wilson, 1997*a*).

---

[6]Loosely speaking, the local degree counts the number of cycles around the original game traversed by the image of a cycle in the graph around the component. The degree is negative if the cycles have reversed orientations.

[7]Govindan and Wilson (1997*b*) proved this for two person games while DeMichelis and Germano (1998) did so for $n$ player games.

The index can be calculated as follows: The index of an isolated equilibrium is the sign of an associated jacobian. It follows straightforwardly that any strict Nash equilibrium has index $+1$. The index of a component is the sum of the indexes of the nearby equilibria of any generic nearby game (measured in distance between payoff vectors). In particular, if for some nearby generic game there are no equilibria nearby the component, the index of the component is zero. That is, any inessential component has index zero and any component with non-zero index is essential. Since the index of a component is invariant, it follows that components with non-zero index are essential in any equivalent game, and thus contain a hyperstable set. Govindan and Wilson (1996, 1997$a$,$b$) conjecture that a component is essential if and only if it has non-zero index. One of the examples in Section 3 shows that this is false: essential components may have index zero. However, this component is not essential in every equivalent game. It remains an open question whether components that are essential in every equivalent game can have zero index.[8]

An important and very useful property of the indexes of the equilibrium components of a game is that they must sum to $+1$. This is a consequence of the Poincaré-Hopf theorem. We will use this property repeatedly in the next sections. The class of outside option games under consideration in this paper has two equilibrium components: the outside option component and the forward induction equilibrium $s^*$. If $s^*$ is strict it has index $+1$ and, by the property mentioned above, the outside option component must have index zero. When $s^*$ is in mixed strategies its index may be either $+1$ or $-1$. In the latter case the outside option component must have index $+2$. (In particular, the outside option component will then be essential and contain a hyperstable set.)

## 3   Essential and hyperstable sets

In this section we will prove our negative results: we first provide an example in which the forward induction equilibrium is strict but the set of equilibria that are consistent with the outside option outcome is shown to be essential but not hyperstable since it

---

[8]It turns out that this open question is closely related to the open question posed in Section 3 whether the outside option outcome can be hyperstable when the forward induction equilibrium is strict.

is not robust to payoff perturbations of all equivalent games. We then show that the outside option can be hyperstable: this example has a forward induction equilibrium in mixed strategies.

Wilson (1997) argued that the unintuitive result in van Damme's (1989) example (see Fig. 2) is due to the fact that the outside option component is not essential. In Wilson's point of view KM stability has the drawback to insist on admissibility above essentiality. Wilson argues that admissibility should not be invoked when selecting an equilibrium component, although it may be useful to select among some of the equilibria within the selected (essential) component. Indeed, in many examples where KM stability yields unintuitive results, intuition is restored when only essential components are considered.[9] In particular, the outside option component in Van Damme's game (Fig. 2) is not essential. Hence, if we insist on essentiality, the outside option component will not be selected. However, as we will show now, essentiality is not restrictive enough either to always select the forward induction equilibrium.

## 3.1 Essential sets

In this subsection we will show that the outside option component can be essential. We first define the notion of an essential set.

**Definition 1** *A closed set of Nash equilibria $S$ is* essential *if for any small payoff perturbation of the normal form, there exists a Nash equilibrium close to $S$. (Jiang Jia-he, 1963).*

Consider the (generic) game in Fig. 4a. The subgame has 5 equilibria: the 2 pure equilibria are $s^1 = ((1, 0, 0), (1, 0, 0))$ yielding $(5, 4)$ and $s^2 = ((0, 1, 0), (0, 1, 0))$ yielding $(-1, 8)$. There is one completely mixed equilibrium $s^3 = ((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}))$, yielding

---

[9]Related results are obtained by Govindan (1995) and Govindan and Robson (1998). Govindan (1995) shows that in the chain store paradox of Kreps and Wilson (1982) the unique Mertens stable outcome is the "intuitive" one, while KM stable sets may contain "unintuitive" outcomes. Govindan and Robson (1998) show that admissibility restores the forward induction logic in an example discussed by Gul and Pearce (1996), who argued that forward induction loses its power when continuous public signals are added to the game.

$(-7,2)$ and two other mixed equilibria $s^4 = \left(\left(\frac{2}{3},\frac{1}{3},0\right),\left(\frac{1}{3},\frac{2}{3},0\right)\right)$ yielding $\left(\frac{-25}{3},\frac{8}{3}\right)$ and $s^5 = \left(\left(\frac{1}{3},0,\frac{2}{3}\right),\left(\frac{2}{3},0,\frac{1}{3}\right)\right)$ yielding $\left(\frac{-1}{3},\frac{8}{3}\right)$.



Fig. 4a: Outside Option Game in Extensive Form.
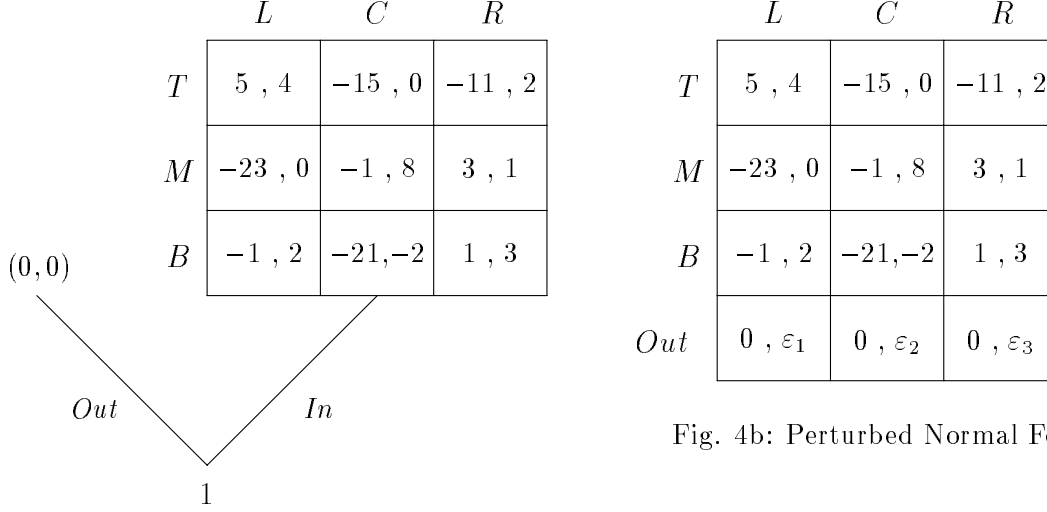
Fig. 4b: Perturbed Normal Form.

Obviously, $s^1$ is strict and it is the only equilibrium that gives a positive payoff, i.e. a payoff higher than the outside option. Clearly, since $s^1$ is strict it constitutes a singleton stable set, for any notion of stability. However, we shall show that this game admits an essential set in which the outside option is chosen.

Let $\Theta_2$ be the set of mixed strategies for player 2 against which the best reply for player 1 is his outside option. This set is illustrated in Fig. 5. Only the 6 extreme points of this set are possible candidates for members of a minimal essential set in which the outside option is chosen. We shall refer to these extreme points as $P_1$ through $P_6$ in accordance to Fig. 5. Namely, $P_1$ and $P_2$ are the extreme points that make player 1 indifferent between $T$ and $Out$. At $P_3$ player 1 is indifferent between $B$ and $Out$. At $P_4$ player 1 is indifferent between $M$, $B$, and $Out$. At $P_5$ player 1 is indifferent between $M$ and $Out$. At $P_6$ player 1 strictly prefers $Out$.
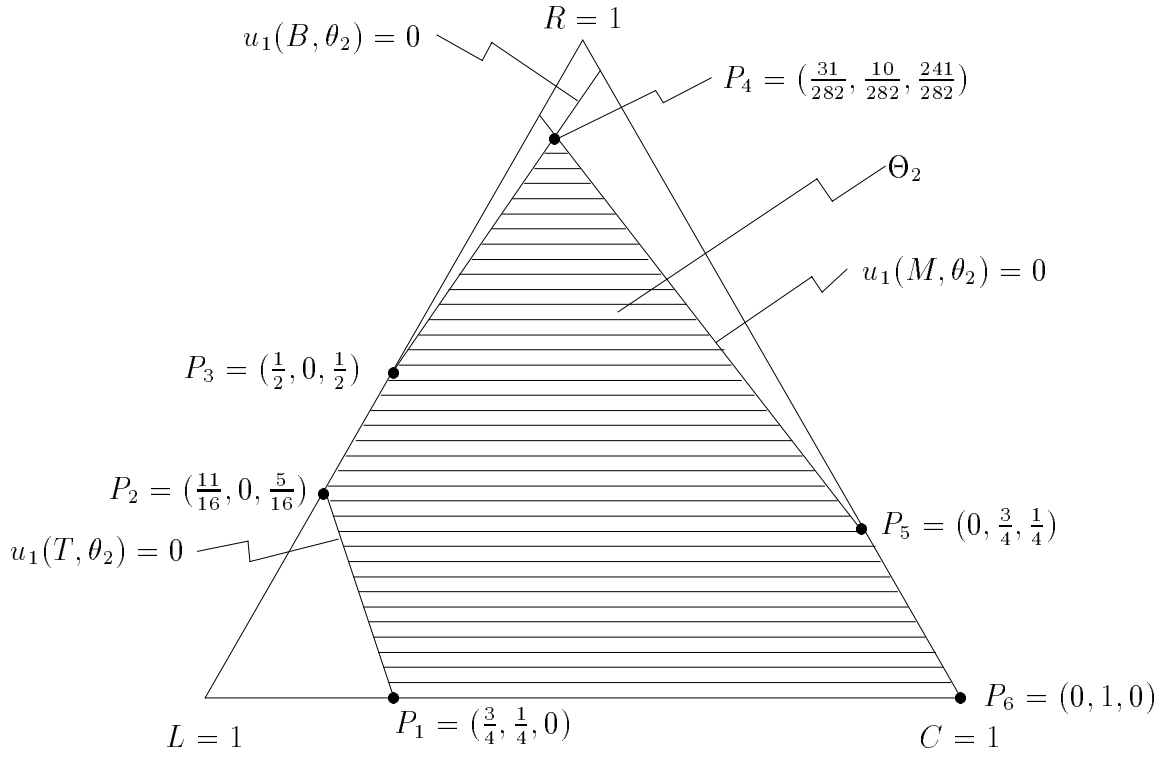
10

Fig. 5.

In order to check whether these extreme points can be part of an essential set, we need to check if there exist equilibria close to those extreme points in the game where players' payoffs are perturbed slightly. Because of the genericity of the game, the only payoff perturbations that matter are those that affect the payoff of player 2 when player 1 chooses $Out$. (See Fig. 4b.)

**Proposition 1** *The outside option component of the game defined in Fig. 4 is essential.*

*Proof:* The following table lists all equilibria near the (extreme points of the) outside option component for all relevant perturbations:

perturbations　　　　　　　　　　　　　　　　　　　equilibrium

$$\left.\begin{array}{c} \varepsilon_2 \geq \varepsilon_1 \\[2mm] \varepsilon_2 \geq 2\varepsilon_3 - \varepsilon_1 \end{array}\right\} \implies \left(\frac{\varepsilon_2 - \varepsilon_1}{4 + \varepsilon_2 - \varepsilon_1} T + \frac{4}{4 + \varepsilon_2 - \varepsilon_1} Out, P_1\right)$$

$$\left.\begin{array}{c} \varepsilon_2 \leq 2\varepsilon_3 - \varepsilon_1 \\[2mm] \varepsilon_3 \geq \varepsilon_1 \end{array}\right\} \implies \left(\frac{\varepsilon_3 - \varepsilon_1}{2 + \varepsilon_3 - \varepsilon_1} T + \frac{2}{2 + \varepsilon_3 - \varepsilon_1} Out, P_2\right)$$

$$\left.\begin{array}{c} \varepsilon_3 \leq \varepsilon_1 \\[2mm] 5\varepsilon_1 \geq 4\varepsilon_3 + \varepsilon_2 \end{array}\right\} \implies \left(\frac{\varepsilon_1 - \varepsilon_3}{1 + \varepsilon_1 - \varepsilon_3} B + \frac{1}{1 + \varepsilon_1 - \varepsilon_3} Out, P_3\right)$$

$$\left.\begin{array}{c} 5\varepsilon_1 \geq 4\varepsilon_3 + \varepsilon_2 \\[2mm] 7\varepsilon_1 \geq 8\varepsilon_3 - \varepsilon_2 \end{array}\right\} \implies \left(\frac{5\varepsilon_1 - \varepsilon_2 - 4\varepsilon_3}{12(1 + \varepsilon_1 - \varepsilon_3)} M + \frac{7\varepsilon_1 + \varepsilon_2 - 8\varepsilon_3}{12(1 + \varepsilon_1 - \varepsilon_3)} B + \frac{1}{1 + \varepsilon_1 - \varepsilon_3} Out, P_4\right)$$

$$\left.\begin{array}{c} 7\varepsilon_1 \leq 8\varepsilon_3 - \varepsilon_2 \\[2mm] \varepsilon_3 \geq \varepsilon_2 \end{array}\right\} \implies \left(\frac{\varepsilon_3 - \varepsilon_2}{7 + \varepsilon_3 - \varepsilon_2} M + \frac{7}{7 + \varepsilon_3 - \varepsilon_2} Out, P_5\right)$$

$$\left.\begin{array}{c} \varepsilon_3 \leq \varepsilon_2 \\[2mm] \varepsilon_2 \geq \varepsilon_1 \end{array}\right\} \implies (Out, P_6)$$

Fig. 6 illustrates the conditions for the different extreme points to be equilibrium strategies. In this figure we take $\varepsilon_2 = 0$. This is without loss of generality since only relative perturbations matter.
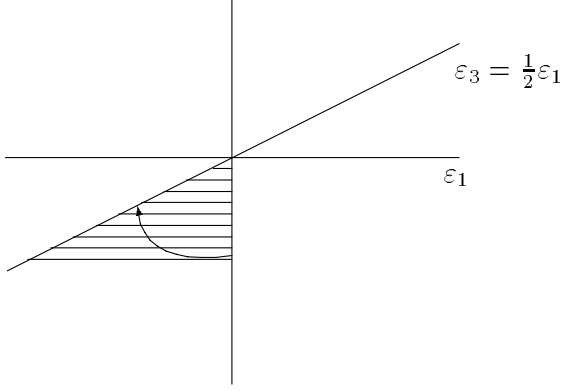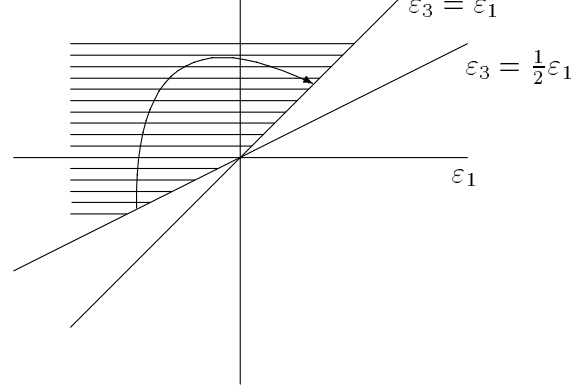
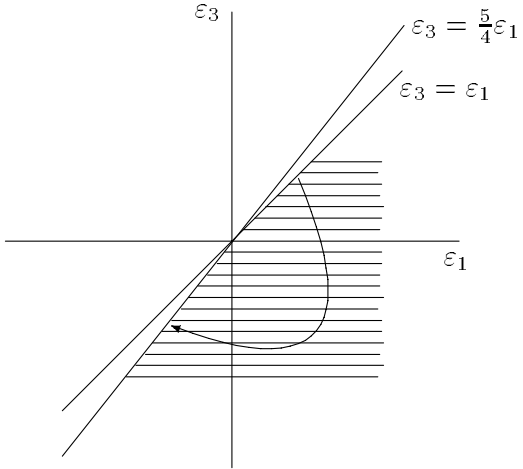Fig. 6a: Equilibria close to $P_1$



Fig. 6b: Equilibria close to $P_2$
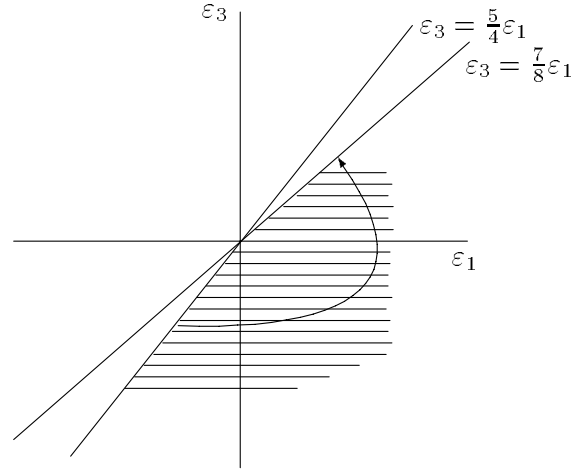


Fig. 6c: Equilibria close to $P_3$
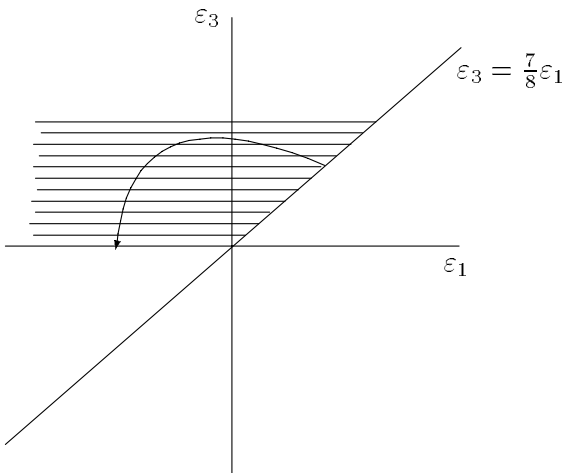


Fig. 6d: Equilibria close to $P_4$



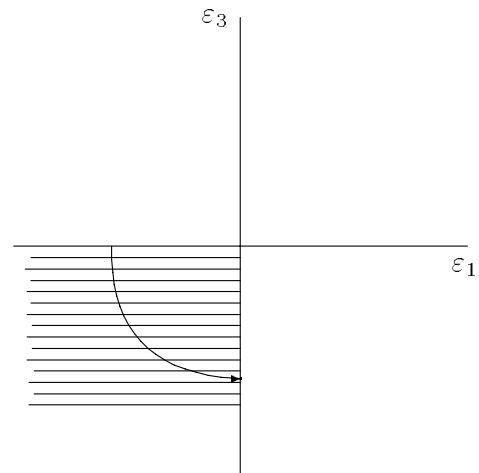Fig. 6e: Equilibria close to $P_5$



Fig. 6f: Equilibria close to $P_6$

13

From Fig. 6 it is easy to see that each of $\{P_1, P_2, P_3\}$, $\{P_2, P_3, P_6\}$, $\{P_3, P_5, P_6\}$ and $\{P_4, P_5, P_6\}$ cover all possible perturbations. Thus the sets $X_1 = \{Out\} \times \{P_1, P_2, P_3\}$, $X_2 = \{Out\} \times \{P_2, P_3, P_6\}$, $X_3 = \{Out\} \times \{P_3, P_5, P_6\}$ and $X_4 = \{Out\} \times \{P_4, P_5, P_6\}$ are (minimal) essential.

## 3.2   Hyperstable sets

In the example above the set of equilibria supporting the outside option is essential, and contains many minimal essential sets, but it does not contain any hyperstable set. Some authors have mistakenly claimed that any essential set must contain a hyperstable set (e.g. Ritzberger and Weibull, 1995, Prop. 4b and Coroll. 4). However, hyperstability also requires an invariance property. In particular, every equivalent game, obtained by adding a mixed strategy as an additional pure strategy, should have the same solution. Formally,

**Definition 2** *S is* hyperstable *in a game G if it is minimal with respect to the following property: S is a closed set of Nash equilibria of G such that, for any equivalent game $G'$ and for any small payoff perturbation of the normal form of $G'$, there exists a Nash equilibrium close to S. (Kohlberg and Mertens, 1986).*

We will now show that there is a game equivalent to the one in Fig. 4, in which the outside option component is not essential.

Let $z = \frac{61}{81}L + \frac{20}{81}C$. We add the mixed strategy $z$ as an additional pure strategy for player 2 to the game. When we now perturb the payoffs of the game slightly, we obtain the $4 \times 4$ normal form game indicated in Fig. 7 where $\varepsilon > 0$. (We only perturb player 2's payoff after $(Out, z)$ into $\varepsilon$.)

|   | $L$ | $C$ | $R$ | $z$ |
|---|-----|-----|-----|-----|
| $T$ | 5 , 4 | $-15$ , 0 | $-11$ , 2 | $\frac{5}{81}$ , $\frac{244}{81}$ |
| $M$ | $-23$ , 0 | $-1$ , 8 | 3 , 1 | $-\frac{1423}{81}$ , $\frac{160}{81}$ |
| $B$ | $-1$ , 2 | $-21$,$-2$ | 1 , 3 | $-\frac{481}{81}$ , $\frac{82}{81}$ |
| $Out$ | 0 , 0 | 0 , 0 | 0 , 0 | 0 , $\varepsilon$ |

Fig. 7: Perturbed Equivalent Game.

The unique Nash equilibrium of this game is $(T, L)$. To see why, consider first some Nash equilibrium that does not use $R$. Then neither $M$ nor $B$ can be used since they are dominated by $Out$. Hence, only $L$ and $z$ can be used in the equilibrium. That, however, implies that $T$ is the unique best reply for player 1 so that any equilibrium that does not use $R$ is $(T, L)$.

Next consider a Nash equilibrium in which $R$ is used. Then also $Out$ must be used, since any equilibrium different from $(T, L)$ that does not use $Out$ yields a negative payoff to player 1. Hence, the equilibrium must yield zero to player 1. If $T$ were used in this equilibrium, then the mixed strategy of player 2 (seen as distribution over $L$, $C$, and $R$) must be on the line piece between $P_1$ and $P_2$. (See Fig. 5.) Since both $M$ and $B$ yield negative payoffs to any of those points, they cannot be used in the equilibrium. But this means that $R$ will not be used either, a contradiction.

Hence, $T$ is not used. This implies that $L$ is not used either (since some mixtures between $R$ and $z$ do better than $L$). Suppose that $C$ is not used. Let $\alpha \in (0, 1)$ be such that $\alpha R + (1 - \alpha)z$ is on the line piece between $P_4$ and $P_5$. Then $M$ yields zero against this strategy while $B$ yields a negative payoff. Hence, player 1 uses $M$ and $Out$ only. But then some mixtures of $C$ and $z$ do strictly better than $R$, so that $R$ cannot be used, another contradiction. So $C$ and $R$ must be used, which implies that both $M$ and $B$ must be used. Hence, the mixed strategy employed by player 2 (seen as distribution

over $L$, $C$, and $R$) must be $P_4$. However, no convex combination of $C$, $R$, and $z$ yields the same distribution as $P_4$.

This establishes that the outside option component of the game in Fig. 4 does not contain a hyperstable set. Hence, hyperstability uniquely selects the forward induction equilibrium in the game of Fig. 4. However, our next example shows that this is not always the case and that the outside option outcome can be hyperstable.

Consider the game with outside option in Fig. 8. The subgame has three Nash equilibria: $s^1 = (1/2, 1/2, 0), (1/2, 1/2, 0))$ yielding player 1 a payoff of 4, $s^2 = (1/3, 1/3, 1/3), (1/3, 1/3, 1/3))$ yielding player 1 a payoff of 8/3, and $s^3 = (B, R)$ yielding player 1 a payoff of 1. Hence, $s^1$ is the only equilibrium yielding player 1 more than the outside option of 3. Moreover, if player 2 is convinced that player 1 plays his first two strategies with equal probability, *any* best reply of player 2, i.e. any mixture between $L$ and $R$ will yield player 1 a payoff of 4.

**Proposition 2** *In the game of Fig. 8 both the forward induction equilibrium and the outside option outcome are hyperstable.*

*Proof.* Govindan and Wilson (1996,1997a) showed that an equilibrium component with non-zero index is essential in all equivalent games, and thus contains a hyperstable set. It is easily verified that the index of the forward induction equilibrium is $-1$. Since the sum of indexes of all components must equal $+1$, this means that the component of equilibria supporting the outside option has index $+2$. Hence, in the game of Fig. 8 both equilibrium outcomes are hyperstable. □

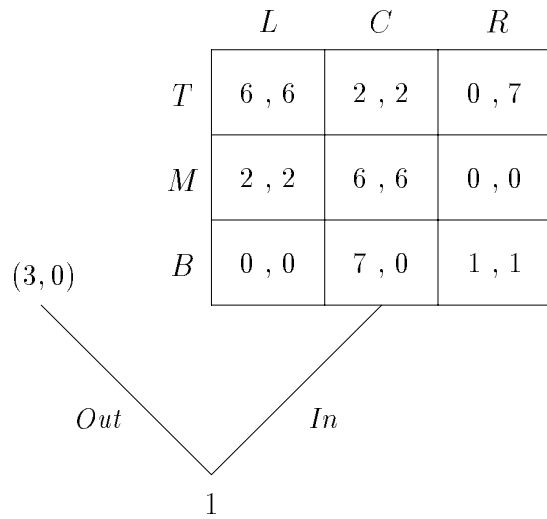|   | L | C | R |
|---|---|---|---|
| T | 6 , 6 | 2 , 2 | 0 , 7 |
| M | 2 , 2 | 6 , 6 | 0 , 0 |
| B | 0 , 0 | 7 , 0 | 1 , 1 |

$(3,0)$

*Out*      *In*

1

Fig. 8: Hyperstability does not satisfy Forward Induction.

## 3.3    Discussion

Note that if the forward induction equilibrium is strict it has index $+1$, so that the outside option component must have index zero. Govindan and Wilson's sufficient condition for equilibrium components to be essential and invariant is violated. For this reason we had to give a constructive proof of the essentiality of the outside option component in the game of Fig. 4. The example shows that essential components may have zero index. This partially resolves an open question posed by Govindan and Wilson. It remains an open question whether components that are essential in every equivalent game (and therefore contain a hyperstable set) can have zero index. This question would be answered affirmatively if one would find an example of an outside option game with a strict forward induction equilibrium in which the outside option outcome is hyperstable.

Govindan and Wilson's (1996) original sufficient condition for an essential component to be invariant is that the local degree of the projection map from the Nash graph to the space of payoff perturbations be non-zero. In the example of Fig. 4 this degree is zero, as can be most easily seen from Fig. 6: A cycle in the Nash graph around the outside option component, $(P_1 \!-\! P_2 \!-\! P_3 \!-\! P_4 \!-\! P_5 \!-\! P_6)$ corresponds to the following path in the

perturbation space: first (more than) a full turn clockwise (Figs. 6a,b,c) and then a (more than) full turn counter-clockwise (Figs. 6d,e,f). The projection of the cycle $P_1$—$P_6$ consists of two paths with opposite orientations so that they cancel each other out. The local degree of the projection is thus zero.

Govindan and Wilson's condition of a non-zero degree is similar to the one used by Mertens (1989, 1991) to define Mertens stability. However, Mertens considers the projection map from the Nash graph to the space of *strategy* perturbations. From van Damme's example in Fig. 2 it is clear that the distinction between payoff and strategy perturbations matters, since the outside option outcome is Mertens stable but not hyperstable. (However, for the game of Fig. 4 one can check, by explicitly calculating the Nash graph for the relevant small strategy perturbations, that this degree is also zero.) By incorporating the non-zero degree condition into the definition of Mertens stable sets, these sets are guaranteed to have certain desirable properties but the proof of existence becomes very hard. We believe, however, that his proof can be readily adapted to show the following conjecture.

**Conjecture.** *Every game has a Mertens stable set that lies within a component with positive index.*

If this Conjecture is true, it follows that every game has a Mertens stable set that lies within an essential component. Let us call such sets essential Mertens stable sets. As we wrote at the beginning of this section, Wilson (1997) argued that one should first restrict attention to essential components and only then make a further selection using admissibility or other stability criteria. The essential Mertens stable set seems to be as much as one can do in this respect. However, even this very strong notion of strategic stability is not restrictive enough to uniquely select the forward induction equilibrium in the game of Fig. 8. Namely, the outside option component has index +2 and is thus essential. It must contain a Mertens stable set since it is the only component with a positive index. Note that the forward induction equilibrium is "viable" (see Property 1) since it forms an essential Mertens stable set.

# 4 Evolutionary Stability

Swinkels (1992) introduced the notion of an equilibrium evolutionarily stable set (EES set). An EES set is a closed and non-empty set of Nash equilibria that is stable under a dynamic evolutionary process. The evolutionary process is not modeled explicitly. Instead, the stability requirement is described by an entry condition for mutants. Roughly speaking, this entry condition says that a small portion of mutants can enter (and 'survive') if the strategy they play is a best response to the strategy of the post-entry population.

Let $\Sigma = \Sigma_1 \times \Sigma_2$ denote the set of mixed strategy profiles and let $\mathcal{B}(s) = \mathcal{B}_1(s) \times \mathcal{B}_2(s)$ denote the set of best replies against $s$.

**Definition 3** $\Theta \subset \Sigma$ *is an equilibrium evolutionarily stable set (EES set) if it is a minimal closed and non-empty set of Nash equilibria that satisfies*

**(S)** *there exists $\delta' > 0$ such that for all $\delta \in (0, \delta')$, for all $s \in \Theta$ and for all $s' \in \Sigma_1 \times \Sigma_2$*

$$s' \in \mathcal{B}((1 - \delta)s + \delta s') \quad \Rightarrow \quad (1 - \delta)s + \delta s' \in \Theta.$$

Swinkels (1992, Fig. 10, p. 330-331) noted that an EES set need not contain an element consistent with forward induction. However, his example is not really adequate to test forward induction, since the subgame has a continuum of equilibria and there does not exist a number $x$ such that $u_1(s^*) > x > u_1(s)$, for all equilibria $s$ different from the preferred equilibrium $s^*$. The example can be corrected. Namely, consider the game $g^{out}$ from Fig. 9.

|       | $L$  | $C$   | $R$   |
|------:|:----:|:-----:|:-----:|
| $T$   | 0,8  | 15,0  | 0,15  |
| $M$   | 0,7  | 0,15  | 15,0  |
| $B$   | 5,5  | 7,1   | 8,0   |
| $Out$ | 6,6  | 6,6   | 6,6   |

Fig. 9: Outside Option Game in Normal Form.

The subgame $g$ has two equilibria, which demonstrates the nongenericity of the example: $s = (B, L)$ and $s^* = (\frac{1}{2}T + \frac{1}{2}M, \frac{1}{2}C + \frac{1}{2}R)$. The latter equilibrium yields 7.5 and is the only one that yields a payoff higher than the outside option. It is easily verified that $\{s^*\}$ is not an EES set since $s$ could enter that set. The set of all Nash equilibria in which player 1 plays his outside option is an EES set.[10]

At first sight this example seems to show that EES sets are not consistent with forward induction. However, the example is problematic for several reasons. First, it is nongeneric and it does not belong to the subclass $\bar{,}$.[11] Second, the preferred equilibrium itself is not evolutionarily stable in the subgame. Hence, we may argue that $s^*$ is not viable (when EES set is the solution concept employed) so that the forward induction argument does not apply. In order to investigate whether EES sets capture the notion of forward induction we should therefore only consider those outside option games $g^{out} \in \bar{,}$ for which the preferred equilibrium $s^*$ constitutes an EES set in the subgame $g$. For such games it is straightforward to show that $\{s^*\}$ is then also an EES set of $g^{out}$.[12] Hence, in order to see that EES sets uniquely select the forward induction equilibrium (in this subclass of outside option games), we need to show that $g^{out}$ has no other EES sets. Since EES sets are maximal connected sets of Nash equilibria and since $g^{out}$ has only two equilibrium components, this amounts to showing that the outside option component is not an EES set. That is what we do in this section.

We will proceed as follows. We start by assuming that there exists an outside option game $g^{out} \in \bar{,}$ for which both equilibrium components are EES sets. We will derive a contradiction in a number of steps. First we construct a new game $\tilde{g}^{out}$ by adding certain strategies for player 1. We show that the additional strategies are such that the outside option component remains an equilibrium component while the forward induction equilibrium is destroyed and no new equilibria appear (Lemma 1). Hence, $\tilde{g}^{out}$ has only

---

[10] This is due to the fact that at no strategy of the outside option component both $T$ and $M$ are (simultaneous) best replies.

[11] When player 1 plays his part of the preferred equilibrium, $s_1^*$, and player 2 best responds by choosing $L$, 1's payoff is less than the outside option.

[12] EES sets are invariant w.r.t addition and deletion of NWBR strategies. See Swinkels (1992, Thm. 6).

one equilibrium component and its index must be $+1$. We then argue that the index of the outside option component in the original game must also be $+1$ (Lemma 2). On the other hand, if the preferred equilibrium itself constitutes an EES set, it must have index $+1$ (Lemma 3), which implies that the outside option component in the original game must have zero index (Corollary 1). The contradiction between Lemma 2 and Corollary 1 thus shows that our hypothesis was wrong: the outside option component cannot be an EES set. Thus we then have shown

**Proposition 3** *Let $g^{out} \in \bar{,}$ and suppose the forward induction equilibrium $s^*$ constitutes a singleton EES set. Then it is the unique EES set so that the outside option outcome is not equilibrium evolutionarily stable.*

Without loss of generality we will assume that the outside option yields player 1 a payoff of 0 so that $u_1(s^*) > 0 > u_1(s)$ for any Nash equilibrium of the subgame $s$ different from $s^*$. We introduce some notation. Let $X$ denote the outside option component of $g^{out}$, i.e.

$$X = \{Out\} \times X_2 = \{Out\} \times \{\theta_2 \in \Sigma_2 : u_1(e_i, \theta_2) \leq 0 \ \ \forall e_i \in E\}.$$

Let $E'$ denote the set of pure strategies of player 1 that correspond to the binding constraints in the definition of $X_2$, i.e.

$$E' = \{e_i \in E : \exists \theta_2 \in \Sigma_2 \text{ s.t. } u_1(e_i, \theta_2) = 0 = \max_{e_k \in E} u_1(e_k, \theta_2)\}.$$

Finally, let $I' = \{i' : e_{i'} \in E'\}$.

<div align="center">*Construction of $\tilde{g}^{out}$*</div>

Note that $s_2^* \notin X_2$. Hence, there exists $i' \in I'$ such that $u_1(e_{i'}, s_2^*) > 0$. Let $\varepsilon \in (0, u_1(e_{i'}, s_2^*))$ be small. Note that $u_1(s^*) > \varepsilon > 0$ and thus, by assumption, $g$ does not have any Nash equilibrium yielding a payoff of $\varepsilon$.

We will construct $\tilde{g}^{out}$ from $g^{out}$ by adding extra strategies for player 1 only. For every $i' \in I'$ we add a pure strategy $\tilde{e}_{i'}$ for player 1. Let $\tilde{E} = \{\tilde{e}_{i'} : i' \in I'\}$ denote the set of those additional strategies. Let $K > 1$. The payoffs in $\tilde{g}^{out} = (E \cup \tilde{E} \cup \{Out\}, F, \tilde{u}_1, \tilde{u}_2)$

are given by

$$\tilde{u}_k(e_i, f_j) = u_k(e_i, f_j) \qquad\qquad (e_i \in E \cup \{Out\}, \ k = 1, 2)$$

$$\tilde{u}_1(\tilde{e}_{i'}, f_j) = K^{i'}\left(u_1(e_{i'}, f_j) - \varepsilon\right) \qquad\qquad (\tilde{e}_{i'} \in \tilde{E})$$

$$\tilde{u}_2(\tilde{e}_{i'}, f_j) = u_2(e_{i'}, f_j) \qquad\qquad (\tilde{e}_{i'} \in \tilde{E})$$

The idea behind this construction is as follows. Consider the "inflated" set

$$X_2^\varepsilon = \{\theta_2 \in \Sigma_2 \ : \ u_1(e_{i'}, \theta_2) \leq \varepsilon \ \ \forall e_{i'} \in E'\}.$$

Clearly, $X_2 \subset X_2^\varepsilon$. Note that all additional strategies yield negative payoffs to player 1 when playing against a strategy in the interior of this inflated set, while at least one additional strategy will yield a positive payoff against strategies outside this inflated set. In particular, for large values of $K$ player 1's best reply against a strategy outside the inflated set must be one of his additional strategies. Since $s_2^*$ is outside the inflated set, the forward induction equilibrium of $g^{out}$ will not be an equilibrium in $\tilde{g}^{out}$. Lemma 1 characterizes the equilibria of the newly constructed game for large values of $K$.

**Lemma 1** *For large $K$, $\tilde{g}^{out}$ has a unique equilibrium component, namely*

$$\tilde{X} = \{Out\} \times \tilde{X}_2 = \{Out\} \times \{\theta_2 \in \Sigma_2 \ : \ u_1(e_i, \theta_2) \leq 0 \ \ \forall e_i \in E\},$$

*Proof.* Note that for all strategies $s_2 \in \Sigma_2$

$$e_{i'} \in E' \text{ and } u_1(e_{i'}, s_2) \leq 0 \Rightarrow \tilde{u}_1(\tilde{e}_{i'}, s_2) \leq -\varepsilon K^{i'} < 0.$$

Hence, it is clear that the set $\tilde{X}$ is a component of equilibria in $\tilde{g}^{out}$ for any $K > 0$.

Suppose now, on the contrary, that for all large $K$ there is some equilibrium $s^K \notin \tilde{X}$. Without loss of generality (by taking a convergent subsequence) we may assume that the sequence $\{s^K\}_K$ converges to some strategy profile $s$. We may also assume that the support of $s^K$ is constant, although of course the support of $s$ may be a strict subset of the support of $s^K$. We will show (in Claim 2 below) that we can construct a strategy $\hat{s}_1$ such that $(\hat{s}_1, s_2)$ is an equilibrium in $g^{out}$ which yields player 1 a payoff of $\varepsilon$, contradicting the assumption that the only equilibrium with positive payoffs in that game is $s^*$ while $u_1(s^*) > \varepsilon$.

We first determine the best replies for player 1 in $\tilde{g}^{out}$ against the limit strategy $s_2$.

**Claim 1:**

(i) $\max_{e_{i'} \in E'} \tilde{u}_1(e_{i'}, s_2) = \varepsilon$.

(ii) $\max_{e_i \in E \setminus E'} \tilde{u}_1(e_i, s_2) < \varepsilon$.

(iii) If $\tilde{e}_{i'}$ is in the support of $s_1$, then $\tilde{u}_1(e_{i'}, s_2) = \varepsilon$.

*Proof.* (i) By contradiction. Suppose first that there is $e_{i'} \in E'$ with $\tilde{u}_1(e_{i'}, s_2) = u_1(e_{i'}, s_2) > \varepsilon$. Of all such strategies take the one that has the highest index $i'$. Suppose it is $e_{i'}$. Then

$$\tilde{u}_1(\tilde{e}_{i'}, s_2) = K^{i'}\left(u_1(e_{i'}, s_2) - \varepsilon\right) \to \infty \text{ as } K \to \infty.$$

For $\bar{K}$ large enough, we have for all $K > \bar{K}$, $\mathcal{B}_1(s_2) = \{\tilde{e}_{i'}\}$. That is, for large $K$ there is a unique best reply against $s_2$. But then, for large enough $K$, we must also have that $\mathcal{B}_1(s_2^K) = \{\tilde{e}_{i'}\}$. Let $\{f_j\} = \mathcal{B}_2(\tilde{e}_{i'}) = \mathcal{B}_2(e_{i'})$. Then $s^K = (\tilde{e}_{i'}, f_j) = s$. However, then $(e_{i'}, f_j)$ is a mutant strategy that could enter $X$ (in $g^{out}$), and since $u_1(e_{i'}, f_j) > \varepsilon > 0$ it would actually take the population out of the set $X$, which contradicts the assumption that $X$ is an EES set.[13] Hence, $u_1(e_{i'}, s_2) \leq \varepsilon$ for all $e_{i'} \in E'$.

Suppose now that $\tilde{u}_1(e_{i'}, s_2) < \varepsilon$ for all $e_{i'} \in E'$. Then $\tilde{u}_1(\tilde{e}_{i'}, s_2) \to -\infty$ and thus $\tilde{e}_{i'}$ is not in the support of $s_1^K$ for any $i' \in I'$ when $K$ is large. But that means that $s^K = s$ (viewed as strategy in $g^{out}$) is an equilibrium of $g^{out}$ which yields player 1 a positive payoff. Hence, $s = s^*$. However, recall from the remark before the construction of $\tilde{g}^{out}$ that there exists some $i' \in I'$ such that $u_1(e_{i'}, s_2^*) > \varepsilon > u_1(e_{i'}, s_2)$ which contradicts our previous conclusion that $s = s^*$. Hence, $\tilde{u}_1(e_{i'}, s_2) = \varepsilon$ for some $e_{i'} \in E'$. $\square$

(ii) We conclude from (i) that $s_2$ must be very close to (but outside of) $X_2$ so that no strategy $e_i \in E \setminus E'$ will be a best reply to $s_2$. $\square$

(iii) It follows from (i) that all additional strategies $\tilde{e}_{i'}$ yield player 1 a payoff of at most zero. Moreover, $\tilde{e}_{i'}$ is not in the support of $s_1^K$ (and, hence, not in the support of $s_1$) if it yields player 1 a strictly negative payoff against $s_2$. $\square$

---

[13]The mutant strategy could enter when the population is at some state $(Out, \theta_2)$ where $\theta_2$ is such that $u_1(e_{i'}, \theta_2) = 0 > u_1(e_i, \theta_2)$ for all $e_i \neq e_{i'}$.

Let $I = \{i : \tilde{u}_1(e_i, s_2) = \varepsilon\}$. From (i) and (ii) of Claim 1 we know that $I \neq \emptyset$ and that $I \subset I'$. Hence, if $e_i$ is in the support of $s_1$ then $i \in I$. Similarly, from Claim 1(iii) we know that if $\tilde{e}_i$ is in the support of $s_1$ then $i \in I$. Hence, there are nonnegative numbers $\alpha_i, \beta_i$ ($i \in I$) such that

$$s_1 = \sum_{i \in I} \alpha_i \tilde{e}_i + \sum_{i \in I} \beta_i e_i.$$

Define

$$\hat{s}_1 = \sum_{i \in I} (\alpha_i + \beta_i) e_i.$$

**Claim 2:** $(\hat{s}_1, s_2)$ is an equilibrium of $g^{out}$ yielding player 1 a payoff of $\varepsilon > 0$.


*Proof.* Each pure strategy in the support of $s_2$ is a best reply (in $\tilde{g}^{out}$) against $s_1^K$ for all $K$, and by continuity it must also be a best reply against $s_1$. But since $\tilde{u}_2(\tilde{e}_i, f_j) = \tilde{u}_2(e_i, f_j) = u_2(e_i, f_j)$, we must also have that each pure strategy in the support of $s_2$ is a best reply against $\hat{s}_1$ (in $g^{out}$), hence that $s_2 \in \mathcal{B}_2(\hat{s}_1)$.

Every pure strategy $e_{i'} \in E'$ in the support of $s_1$ is a best reply against $s_2^K$, and by continuity also against $s_2$. For every pure strategy $\tilde{e}_{i'}$ in the support of $s_1$ we know by Claim 1(iii) that $\tilde{u}_1(e_{i'}, s_2) = \varepsilon$. Hence, for every $\tilde{e}_{i'}$ in the support of $s_1$, $e_{i'}$ is a best reply against $s_2$. This implies that $\hat{s}_1 \in \mathcal{B}_1(s_2)$. □

This concludes the proof of Lemma 1. □

**Lemma 2** $index(\tilde{X}) = index(X) = +1$.

*Proof.* Since $\tilde{X}$ is the unique equilibrium component of $\tilde{g}^{out}$ its index must be $+1$. (See Section 2.2.)

In order to show that the index of $X$ and $\tilde{X}$ must be equal we use a result by Govindan and Wilson (1997) that explains how to calculate the index of a component. The index of the component can be calculated as follows: Perturb the payoffs of the normal form of the game slightly in a generic way. Compute all equilibria of the perturbed game that are nearby the component. Calculate the index of each of those equilibria and sum them. The resulting sum is the index of the component.

Since the added strategies $\tilde{e}_i \in \tilde{E}$ are not best replies against the component, they will not be used in equilibria of the perturbed game that are nearby the component when

24

the payoff perturbations are small enough. That means that when using the same payoff perturbation in $g^{out}$ and $\tilde{g}^{out}$, the set of nearby equilibria of the perturbed games will be identical, and also the indices of each of those nearby equilibria will be the same (since those only depend on the payoff matrix restricted to the strategies actually used with positive probability). Therefore, the index of $X$ and the index of $\tilde{X}$ must coincide. $\square$

On the other hand, we have

**Lemma 3** *Let $\{s'\}$ be an EES set of a two person game. Then $index(s') = +1$.*

*Proof.* Namely, consider the game where players are restricted to use only strategies from $\mathcal{B}(s')$. This game must have $s'$ as the unique Nash equilibrium since any other Nash equilibrium would be able to enter. (See condition (S) in the definition of an EES set.) The index of the unique equilibrium of a game must always equal $+1$, since the sum of the indexes over all equilibrium components equals $+1$. Hence, the index of $s'$ in the restricted game is $+1$. But then the index of $s'$ in the original game is also equal to $+1$. (This follows from the fact that the index of an equilibrium can be calculated from the payoff matrices restricted to the support of the equilibrium. See Govindan and Wilson, 1997$b$.) $\square$

**Corollary 1** $index(X) = 0$.

*Proof.* From Lemma 3 we know that $index(s^*) = +1$. Again using the argument that the sum of indexes over all equilibrium components must equal $+1$ implies then that the only other component, $X$, must have zero index. $\square$

Clearly, Lemma 2 and Corollary 1 contradict each other. They were derived under the hypothesis that $X$ is an EES set. (This was used in the proof of Lemma 1, which in turn was used to prove Lemma 2.) Hence, $X$ is not an EES set and the proof of Proposition 3 is complete: EES sets uniquely select the forward induction equilibrium.

# 5 Conclusion

When introducing a new equilibrium refinement, various authors have checked that their new refinement satisfies some form of forward induction. (Eg. Matsui (1992).) However, they usually chose an outside option game in which repeated elimination of dominated

strategies already uniquely selects the forward induction equilibrium. In this paper we have asked ourselves which solution concepts satisfy the stronger notion of forward induction proposed by van Damme (defined in Section 2) in the class of generic 2-person normal form games preceded by an outside option. This notion of forward induction requires more rationality on the part of the players than just playing rationalizable strategies. It basically requires players to play a Nash equilibrium and to expect others to play a Nash equilibrium.

Because of this required level of sophistication, it seemed obvious to examine solution concepts that are characterized by strong rationality assumptions, like the strategic stability concepts. Van Damme had already shown that KM stability is not restrictive enough to capture forward induction but suggested that perhaps some stronger forms of stability would satisfy forward induction. However, we have seen that neither Mertens stability (and thus neither Hillas stability) nor hyperstability (or essentiality) always capture forward induction. Even combining Mertens stability with essentiality does not seem to yield the forward induction equilibrium.

Surprisingly, we found that an evolutionary concept, EES set, does satisfy forward induction: if the preferred equilibrium in the subgame is a singleton EES set, then it is the unique EES set of the outside option game. In other words, choosing the outside option is not evolutionarily stable. This result is perhaps not completely satisfactory since EES sets may fail to exist, even in generic games. However, if the preferred equilibrium itself is not a singleton EES set, we may argue that it is not "viable" (see Property 1) so that we should not expect it to be the plausible solution.

Given the level of sophistication required for the forward induction argument, it is surprising that deductive solution concepts (like stability) do not uniquely select the forward induction equilibrium whereas an eductive solution concept does. It seems to suggest that learning and/or adjustment dynamics might lead to the forward induction equilibrium. We should be cautious in reaching this conclusion. Although EES sets certainly have an evolutionary flavor, they are defined by means of a static entry condition for mutants. This static entry condition requires some rationality on part of the mutants: they have to be equilibrium entrants. This seems to be the key why EES sets select

the forward induction equilibrium, since forward induction also requires deviations to be interpreted as equilibrium play. Usually, rationality requirements of evolutionary or learning dynamics are less stringent than the requirement of equilibrium entrants. Indeed, no dynamics is known to yield EES sets as dynamically stable sets.

Several authors have examined the forward induction property by directly considering certain dynamics. Nöldeke and Samuelson (1993) consider a dynamics driven by stochastic learning and rare mutations. They get some support for forward induction of the limiting distribution of such dynamics. However, they only consider games in which the preferred equilibrium is strict.[14] Ritzberger and Weibull (1995) examine which sets of strategies are asymptotically stable under the wide class of so called sign-preserving selection dynamics. They find that the only asymptotically stable sets are (entire) faces of the strategy space that are *curb*, i.e. closed under rational behavior.[15] A product set of mixed strategies is curb when for every strategy profile in this set, all best replies against it are also contained in the set. (See Basu and Weibull (1991) for details.) It is easily verified that in the case of an outside option game where the preferred equilibrium is strict, the only minimal curb set is the singleton containing that equilibrium. However, when the preferred equilibrium is in mixed strategies, the only curb set may contain all strategy profiles.

Ritzberger (1994) takes a differentiable viewpoint of normal form games by considering dynamics given by the so called Nash field (vaguely related to replicator dynamics) and regular interior approximations thereof. Using these tools he is able to assign an index to each component of Nash equilibria and argues that components with non-zero index have nice properties, and suggests that the property of having a non-zero index be used for equilibrium selection. It follows from the work of Govindan and Wilson (1997*a*,*b*) that those indexes are the same as the ones we have used in this paper. If the preferred equilibrium has an index of $+1$, the outside option component has index

---

[14]In fact, they consider outside option games where the subgame may have several strict equilibria that yield a higher payoff than the outside option. Van Damme's notion of forward induction does not really apply since player 1 cannot unambiguously signal which equilibrium he will play in the subgame.

[15]In particular, an EES set which lies in the relative interior of a face cannot be asymptotically stable under such dynamics.

zero and Ritzberger's method selects the forward induction equilibrium. However, if the preferred equilibrium has index $-1$ (which of course implies that the equilibrium is in mixed strategies and does not form a singleton EES set), the outside option component has index $+2$ and cannot be excluded. An easy way out of this problem consists of restricting attention to equilibrium components with strictly positive index. Such a "solution concept" will obviously satisfy our notion of forward induction[16] and, in contrast to the EES set, it will always exist.

The above discussion suggests that when the preferred equilibrium is in mixed strategies the details of the dynamics seem to be important. In the case where the preferred equilibrium is strict all roads lead to Rome, that is, many different dynamics will lead to the forward induction equilibrium.[17]

# References

Basu, K. and J.W. Weibull (1991). "Strategy subsets closed under rational behavior," *Economics Letters*, **36**, 141-146.

Binmore, K. and L. Samuelson (1999), "Evolutionary Drift and Equilibrium Selection," *Review of Economic Studies*, **66**, 363-393.

van Damme, E. (1989). "Stable Equilibria and Forward Induction," *Journal of Economic Theory,* **48**, 476-496.

van Damme, E. (1994). "Strategic Equilibrium," prepared for: *Handbook of Game*

---

[16]Note, if we use this criterion for equilibrium selection and the preferred equilibrium has index $-1$, it is simply not viable and forward induction does not apply. Also note that an equilibrium may have index $+1$ while it does not form a singleton EES set. Hence, this new solution concept selects a forward induction equilibrium more often than EES sets do.

[17]We should point out one notable exception. Binmore and Samuelson (1999) argue that drift may matter, especially in large equilibrium components that correspond to play at unreached subgames. They show that even in an outside option game where the preferred equilibrium is strict and a simple repeated elimination of dominated strategies argument suffices to select the forward induction equilibrium, some specification of drift may in fact stabilize the outside option outcome.

*Theory with Economic Applications, vol. 3* (R.J. Aumann and S. Hart (Eds.) In press.

DeMichelis, S. and F. Germano (1998). "On the Indices of Zeros of Nash Fields," mimeo CORE.

Fudenberg, D. and J. Tirole (1993). *Game Theory.* MIT Press, Cambridge, MA. Third edition. (First edition 1991).

Govindan, S. (1995). "Stability and the Chain Store Paradox," *Journal of Economic Theory* **66**, 536-547.

Govindan, S. and A. Robson (1998). "Forward Induction, Public Randomization, and Admissibility," *Journal of Economic Theory,* **82**, 451-457.

Govindan, S. and R. Wilson (1996). "A Sufficient Condition for Invariance of Essential Components," *Duke Mathematical Journal,* **81**, 39-46.

Govindan, S. and R. Wilson (1997a). "Equivalence and Invariance of the Index and Degree of Nash Equilibria," *Games and Economic Behavior,* **21**, 56-61.

Govindan, S. and R. Wilson (1997b). "Uniqueness of the Index for Nash Equilibria of Two-Player Games," *Economic Theory,* **10**, 541-549.

Gul, F. and D.G. Pearce (1996). "Forward Induction and Public Randomization," *Journal of Economic Theory,* **70**, 43-65.

Hillas, J. (1990). "On the Definition of Strategic Stability," *Econometrica,* **58**, 1365-1390.

Hillas, J., Jansen, M., Potters, J. and D. Vermeulen (1999). "On the relation among some definitions of strategic stability", mimeo.

Hillas, J. and E. Kohlberg. (1994) "Foundations of Strategic Equilibrium," prepared for: *Handbook of Game Theory with Economic Applications, vol. 3* (R.J. Aumann and S. Hart (Eds.) In press.

Jiang Jia-he (1962). "Essential Fixed Points of the Multivalued Mappings," *Scientia Sinica*, **11** 293-298.

Kohlberg, E. and J.-F. Mertens (1986). "On the Strategic Stability of Equilibria," *Econometrica,* **54**, 1003-1037.

Kreps, D. and R. Wilson (1982). "Reputation and Information," *Journal of Economic Theory* **27**, 253-279.

Matsui, A. (1992). "Best Response Dynamics and Socially Stable Strategies," *Journal of Economic Theory,* **57**, 343-362.

Mertens, J.F. (1989). "Stable Equilibria - A Reformulation, Part I: Definition and Basic Properties," *Mathematics of Operations Research,* **14**, 575-625.

Mertens, J.F. (1991). "Stable Equilibria - A Reformulation, Part II: Discussion of the Definition, and Further Results," *Mathematics of Operations Research,* **16**, 694-753.

Nöldeke, G. and L. Samuelson (1993). "An Evolutionary Analysis of Backward and Forward Induction," *Games and Economic Behavior* **5**, 425-454.

Ritzberger, K. (1994). "The Theory of Normal Form Games from the Differentiable Viewpoint," *International Journal of Game Theory,* **23**, 207-236.

Ritzberger, K. and J. Weibull (1995). "Evolutionary Selection in Normal-Form Games," *Econometrica,* **63**, 1371-1399.

Shapley, L. (1974). "A Note on the Lemke-Howson Algorithm," *Mathematical Programming Study,* **1**, 175-189.

Swinkels, J. (1992). "Evolutionary Stability with Equilibrium Entrants," *Journal of Economic Theory,* **57**, 306-332.

Wilson. R. (1997). "Admissibility and Stability," in *Understanding Strategic Behavior: Essays in Honor of Reinhard Selten,* W. Albers, W. Güth, P. Hammerstein, P. Moldovanu and E. van Damme (Eds.) pp. 85-99. Berlin: Springer-Verlag.