On Forward Induction and Evolutionary and Strategic Stability*

Esther Hauk† Sjaak Hurkens‡

Department of Economics and Business,

Universitat Pompeu Fabra,

Ramon Trias Fargas 25-27,

08005 Barcelona, Spain

fax: +34-93-542 1746

First version July 1999, this version April 2001

running title: Forward Induction and Stability

*Hauk acknowledges financial support from DGES, grant BEC 2000-1029. Hurkens gratefully acknowledges partial financial support from CIRIT, Generalitat de Catalunya, grant 1997SGR 00138 and from DGES, grant BEC 2000-1029. We would like to thank Jean-Francois Mertens, Klaus Ritzberger, Hari Govindan, Robert Wilson, two anonymous referees and the seminar audiences at UPF and ESEM 1999 for helpful discussions.

†e-mail: esther.hauk@econ.upf.es

‡corresponding author: e-mail: sjaak.hurkens@econ.upf.es. Tel: +34-93 542 2556

Abstract

We analyze which normal form solution concepts capture the notion of forward induction, as defined by van Damme [1] 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 [24]) is always consistent with forward induction.

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

Journal of Economic Literature Classification Number: C72

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 [1], 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. [1] 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 <u>forward</u> <u>induction</u> <u>equilibrium</u>.

We have presented this forward induction logic as an <u>extensive</u> <u>form</u> argument. [18] show that the same logic can also be presented in the <u>normal</u> <u>form</u> by means of their definition of normal form subgames. Several authors, and in particular [16], 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 [16] 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 [1]'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." [16, 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 do select the forward induction equilibrium. In fact, whenever the subgame is a generic 2×2 game with three Nash equilibria this is the case. Unfortunately, stability does not select the forward induction equilibrium in all outside option games. Indeed, [1] showed that the outside option outcome can be KM stable. We replicate his example in Fig. 1 below.

[include Figure 1 around here]

The subgame has three Nash equilibria. The pure equilibrium $s_1 = (T, L)$ yields player 1 a payoff of 3, while the other equilibria, $s_2 = ((\frac{2}{3}, \frac{1}{3}), (\frac{1}{2}, \frac{1}{2}, 0))$ and $s_3 = ((\frac{1}{3}, \frac{2}{3}), (0, \frac{1}{2}, \frac{1}{2}))$, yield him a payoff of 1.5. Hence, the forward induction equilibrium is to enter the game and to play (T, L). This equilibrium is indeed stable. However, as [1] 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]

As [1] already pointed out, these sets are neither fully nor hyperstable.[2] They are not even essential.[3] [26] argued that this is the main reason for the unintuitive result in [1]'s example and suggested to restrict attention only to essential components. If we insist on essentiality, the outside option component will not be selected in [1]'s example. Whether essentiality, full stability or hyperstability always select the forward induction equilibrium remained an open question. We show, by means of two generic examples in Section 3, that they do not. This is our main negative result.

More recently [20,21] and [12] introduced new stability concepts that satisfy all of the desirable properties listed by [16].[4] [2] suggested, in a chapter prepared for the Handbook of Game Theory with Economic Applications, that perhaps Mertens stability implies forward induction. However, [14], in a different chapter of the same Handbook, showed this conjecture to be false. As a matter of fact, they used [1]'s game from Fig. 1 as their counterexample: The outside option component contains two Mertens stable sets, namely $\{(Out, (q, 1-q, 0) \mid \frac{1}{3} \le q \le \frac{2}{3})\}$ and $\{(Out, (0, q, 1-q) \mid \frac{1}{3} \le q \le \frac{2}{3})\}$. These Mertens stable sets are the convex hulls of the KM stable sets.[5] Since Mertens stable sets are known to contain Hillas stable sets (see e.g. [13]), 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 [13]), 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. However, we find an evolutionary solution concept that is always consistent with the forward induction logic, namely equilibrium evolutionarily stable sets (EES sets, [24]). This is the main positive result of the paper.

In order to prove both our negative and positive results we will make use of index theory (see section 2). As a by-product of our investigation whether or not there exists some strategic stability concept which always captures forward induction, we resolve an open issue in index theory which is an important result in its own right. [22, Theorem 4] has first shown that every Nash equilibrium component with nonzero index is essential. [9] conjectured that the reverse also holds. Our first counterexample which shows that the outside option can be essential also shows that their conjecture is not true: the essential outside option component has index zero. However, the outside option component fails to be hyperstable, i.e. essential in all equivalent games. Therefore [9]'s original question is now qualified to whether a component that is essential in all equivalent games with the same reduced normal form must have non-zero index.

The remainder of the paper is organized as follows. In Section 2 we recall [1]'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 and show that essential components may have index zero. 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 Γ. [1] argues that any solution concept that is consistent with forward induction should at least satisfy the following property (see also [5, Def. 11.8, p. 464]:

Property 1 *For any outside option game $g^{out} \in \Gamma$ 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 solution concept (or 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 [1]'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 $g^{out} \in \Gamma$. To illustrate this point, consider the outside option game in Fig. 2.

[include Figure 2 around here]

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 [1]'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. However, since that equilibrium is in mixed 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 strategy 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. 2 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. 2 suggests that the class of games Γ 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 $\bar{\Gamma}$.

2.2 Index theory

To prove some of our results in Sections 3 and 4 we will make use of index theory. [22] defined an index for Nash equilibrium components. We will rely on recent advancements by [3] and [8,9,10] 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 is defined for a vector field on the strategy space (i.e. for a smooth map from the space of mixed strategy combinations Σ to its tangent space.) The index does not depend on the particular vector field as long as it is inward pointing at the boundary, continuous in payoffs and vanishes on all Nash equilibria. Moreover, the index and the degree of a component agree and depend only on the reduced normal form of the game.[7]

The index can be calculated as follows: The index of a regular equilibrium is the sign of the determinant of $-J$, where J is the associated Jacobian matrix. 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 is no equilibrium nearby the component, the index of the component is zero. That is, any inessential component has index zero. Also, 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. [8,9,10] 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 thus remains an open question whether components that are essential in every equivalent game can have zero index.

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 outside option games in $\overline{\Gamma}$ have 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 contains a hyperstable set.)

3 Essential and hyperstable sets

In this section we will prove our negative results: we first provide an example of an outside option game with a strict forward induction equilibrium where the outside option component is essential. This component has index zero which shows that it is not a necessary condition for essentiality to have a nonzero index (proving the conjecture of [8,9,10] wrong). In section 3.2 we show that the outside option component of our first example is not hyperstable; therefore, it remains an open question whether it is a necessary condition for hyperstable components to have a nonzero index. We provide a second example (where the forward induction equilibrium is in mixed strategies) in which the outside option component is hyperstable. Of course, this second example would have been sufficient to prove that essentiality does not always capture forward induction. We include the first example, because (i) it has a strict forward induction equilibrium in which case the forward induction argument is most compelling, (ii) because its proof is constructive and does not rely on index theory unlike the proof of our second example, and most importantly (iii) because the example shows (as a by-product) that essential components may have zero index, which resolves an open issue in index theory.

3.1 Essential sets

[26] argued that the unintuitive result in [1]'s example (see Fig. 1) is due to the fact that the outside option component is not essential. In [26]'s point of view KM stability has the drawback to insist on admissibility above essentiality. He 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.[8] In particular, the outside option component in [1]'s game (Fig. 1) 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 to always select the forward induction equilibrium. 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. [15, Theorem 4]. A Nash equilibrium s^* is essential if the set $\{s^*\}$ is essential.*

It is well known that not every game admits an essential equilibrium while for any game the full set of Nash equilibria is always essential. The notion of essentiality has bite when considering minimal or connected essential sets.

Consider the (generic) game in Fig. 3a. 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 $(-7,2)$ and two other mixed equilibria $s^4 = ((\frac{2}{3},\frac{1}{3},0),(\frac{1}{3},\frac{2}{3},0))$ yielding $(\frac{-25}{3},\frac{8}{3})$ and $s^5 = ((\frac{1}{3},0,\frac{2}{3}),(\frac{2}{3},0,\frac{1}{3}))$ yielding $(\frac{-1}{3},\frac{8}{3})$.

[include Figures 3a and 3b around here]

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 strategic stability. However, we shall show that this game admits an essential set in which the outside option is chosen.

Let Θ_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. 4. 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. 4. 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.

[include Figure 4 around here]

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. 3b.).[9] We normalize $\varepsilon_2 = 0$. This is without loss of generality since only relative perturbations matter.

Proposition 1 *The outside option component of the game defined in Fig. 3 is essential.*

Proof: It would be sufficient to show that there is always an equilibrium close to either P_1 or P_2 or P_3. But for completeness we list all equilibria near the extreme points of the outside option component for all relevant perturbations.

perturbations		equilibrium
$\varepsilon_3 \leq \frac{1}{2}\varepsilon_1 \leq 0$	$\Longleftrightarrow$	$\left(\frac{-\varepsilon_1}{4-\varepsilon_1}T + \frac{4}{4-\varepsilon_1}Out, P_1\right)$
$\varepsilon_3 \geq \max\left\{\frac{1}{2}\varepsilon_1, \varepsilon_1\right\}$	$\Longleftrightarrow$	$\left(\frac{\varepsilon_3-\varepsilon_1}{2+\varepsilon_3-\varepsilon_1}T + \frac{2}{2+\varepsilon_3-\varepsilon_1}Out, P_2\right)$
$\varepsilon_3 \leq \min\left\{\frac{5}{4}\varepsilon_1, \varepsilon_1\right\}$	$\Longleftrightarrow$	$\left(\frac{\varepsilon_1-\varepsilon_3}{1+\varepsilon_1-\varepsilon_3}B + \frac{1}{1+\varepsilon_1-\varepsilon_3}Out, P_3\right)$
$\varepsilon_3 \leq \min\left\{\frac{5}{4}\varepsilon_1, \frac{7}{8}\varepsilon_1\right\}$	$\Longleftrightarrow$	$\left(\frac{5\varepsilon_1-4\varepsilon_3}{12(1+\varepsilon_1-\varepsilon_3)}M + \frac{7\varepsilon_1-8\varepsilon_3}{12(1+\varepsilon_1-\varepsilon_3)}B + \frac{1}{1+\varepsilon_1-\varepsilon_3}Out, P_4\right)$
$\varepsilon_3 \geq \max\left\{0, \frac{7}{8}\varepsilon_1\right\}$	$\Longleftrightarrow$	$\left(\frac{\varepsilon_3}{7+\varepsilon_3}M + \frac{7}{7+\varepsilon_3}Out, P_5\right)$
$0 \geq \max\{\varepsilon_1, \varepsilon_3\}$	$\Longleftrightarrow$	(Out, P_6)

Fig. 5 illustrates the set of perturbations that allow for equilibria close to the different extreme points of the outside option component. Clearly, all perturbations are considered. Hence, for any perturbation of the game there exists an equilibrium close to the outside option component. $\square$

[include Figures 5a-f around here]

From Fig. 5 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 we will show that it does not contain any hyperstable set. Some authors have mistakenly claimed that any essential set must contain a hyperstable set (e.g. [23, 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. [16]*

We will now show that there is a game equivalent to the one in Fig. 3, 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 4×4 normal form game indicated in Fig. 6 where $\varepsilon > 0$. (We only perturb player 2's payoff after (Out, z) into ε.)

[include Figure 6 around here]

Proposition 2 *The unique Nash equilibrium of the game in Fig. 6 is* (T, L).

Proof: Suppose $s \neq (T, L)$ is a Nash equilibrium of this game. Let $s_i(k)$ be the weight player i puts on pure strategy k.

Step 1: $s_1(Out) > 0$ and therefore $u_1(s_1, s_2) = 0$.

Suppose not. Then the payoff perturbation ε is irrelevant. Also, $u_1(s_1, s_2) \geq 0$. Let $\widetilde{s_2}$ be defined by

$$\widetilde{s_2} = (s_2(L) + \frac{61}{81}s_2(z), s_2(C) + \frac{21}{81}s_2(z), s_2(R), 0).$$

Then $(s_1, \widetilde{s_2})$ is also a Nash equilibrium of the same game and it gives player 1 the same nonnegative payoff as s. Since no player uses his fourth strategy in $(s_1, \widetilde{s_2})$, we can view this profile as a strategy profile of the subgame g spanned by $\{T, M, B\} \times \{L, C, R\}$. However, the only equilibrium of g that yields player 1 nonnegative payoff is (T, L) (see section 3.1). So $s_1 = T$ and s_2 must be L, since it is the only best reply to T, which contradicts the assumption that $s \neq (T, L)$.

Step 2: $s_2(R) > 0$.

Suppose not. 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 $s_1 = T$ and $s_2 = L$. A contradiction.

Step 3: $s_1(T) = 0$ and $s_2(L) = 0$.

Suppose $s_1(T) > 0$. Since $u_1(s_1, s_2) = 0$ by step 1, 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. 4). 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 which contradicts step 2. Hence $s_1(T) = 0$. This implies that $s_2(L) = 0$ since some mixtures between R and z do better than L.

Step 4: $s_2(C) > 0$.

Suppose not. Let $\alpha \in (0, 1)$ be such that $\alpha R + (1 - \alpha)z$ is on the line piece between P_4 and P_5. Such an α exists since we have chosen z to lie slightly to the right of the line through R and P_4 (see Fig. 4). 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, which contradicts step 2.

Step 5: $s_1(M) > 0$ and $s_1(B) > 0$

If $s_1(M) = 0$, C is dominated by z which contradicts step 4.

If $s_1(B) = 0$, R is dominated by z which contradicts step 2.

Conclusion: Player 1 uses M, B and Out so that 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, since z lies to the right of the line through R and P_4. □

Proposition 2 establishes that the outside option component of the game in Fig. 3 does not contain a hyperstable set. Hence, hyperstability uniquely selects the forward induction equilibrium in the game of Fig. 3. 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. 7. 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 3 *In the game of Fig. 7 both the forward induction equilibrium and the outside option outcome are hyperstable.*

Proof. [8,9] 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.[10] 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. 7 both equilibrium outcomes are hyperstable. □

[include Figure 7 around here]

3.3 Remarks

1. Note that if the forward induction equilibrium is strict it has index +1, so that the outside option component must have index zero.[11] [9]'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. 3. The example shows that essential components may have zero index.

2. Strategic stability concepts are typically defined on the <u>mixed</u> <u>strategy</u> reduced normal form. There has been some doubts that the mixed strategy reduced normal form is the right vehicle for reflecting extensive form arguments: When moving from the <u>pure</u> <u>strategy</u> reduced normal form to the mixed strategy reduced normal form, information sets present in the original extensive form might be destroyed or new normal form information sets can appear ([18]). This might be one reason why strategic stability notions do not always capture forward induction.[12]

3. The above results together with earlier results in the literature have shown that none of the existing stability concepts uniquely select the forward induction outcome. One possible approach to solve this problem is to introduce new and stronger stability notions. Following [26]'s suggestion to first restrict attention to essential components and only then make a further selection using admissibility or other stability criteria, we define the notion of an essential Mertens stable set as a Mertens stable set that is contained in an essential component. Essential Mertens stable sets exist because (i) every game has equilibrium components with nonzero index and (ii) every equilibrium component with nonzero index contains a Mertens stable set [4, Theorem 2].

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. 7. Namely, the outside option component has index +2 and is thus essential. It must contain a Mertens stable set. Note that the forward induction equilibrium is "viable" (see Property 1) since it also forms an essential Mertens stable set.

4 Evolutionary Stability

[24] 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 satisfies 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 mixed 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 $s' \in \Sigma_1 \times \Sigma_2$*

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

[24, 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 by changing the payoff of player 1 after (B, R) from 7 to 8. The resulting game (denoted by g^{out}) is depicted in Fig. 8.

[include Figure 8 around here]

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.[13]

At first sight this example seems to show that EES sets are not consistent with forward induction. However, the example is problematic: it does not belong to the subclass $\bar{\Gamma}$ [14] and it is nongeneric. In fact, Proposition 4 shows that the latter is not a coincidence.

Proposition 4 *Suppose the outside option component of g^{out} is an EES set. Then the subgame g is nongeneric.*

Before giving the proof of the proposition (which is quite involved) let us discuss this result. Proposition 4 states that in generic games choosing the outside option is not equilibrium evolutionarily stable. Since g^{out} has only two equilibrium components and EES sets are maximal connected sets of Nash equilibria, the only other possible candidate for constituting an EES set is the forward induction equilibrium.

Lemma 1 *The forward induction equilibrium s^* in a generic outside option game g^{out} constitutes a singleton EES set if and only if $\{s^*\}$ is an EES set in the subgame g.*

Proof: Suppose $\{s^*\}$ is an EES set in g^{out}. Then Out is not a weak best response for player 1 against this set. Since EES sets satisfy the NWBR property (see [24, Theorem 6]), then $\{s^*\}$ must also be an EES set after deleting Out, that is in the subgame g. Conversely, let $\{s^*\}$ be an EES set in the subgame g. Since EES sets are invariant w.r.t. addition of NWBR (see [24, remark after Theorem 6]), it is immediate that $\{s^*\}$ is also an EES set in the game where we add Out, that is in g^{out}. $\square$

Hence, from Proposition 4 and Lemma 1 it follows that if the outside option game is generic and the preferred equilibrium in the subgame constitutes an EES set, then EES sets uniquely select the forward induction equilibrium. Of course, EES sets do not always exist. In particular, the preferred equilibrium s^* might fail to be an EES set in the subgame g. However, in this case the forward induction argument in Property 1 does not apply, since s^* would not be viable. Hence, we can argue that EES sets are consistent with forward induction.

Proof of Proposition 4

We will proceed as follows. We start by assuming that there exists a generic outside option game g^{out} for which the outside option component is an EES set. 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. We then show in Lemma 2 (and here is where EES sets play a role) that no new equilibria appear in $\tilde{g}^{out}$. Hence, $\tilde{g}^{out}$ has only 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 3). This implies that the forward induction equilibrium must have zero index. This can only occur in nongeneric cases. Hence, our hypothesis was wrong: the outside option component cannot be an EES set in generic games.

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.

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

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.

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

We can then rewrite X_2 as:

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

Construction of $\tilde{g}^{out}$

Note that $s_2^* \notin X_2$. Hence, there exists $e_i \in E'$ 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 ε.

We will construct $\tilde{g}^{out}$ from g^{out} by adding extra strategies for player 1 only. For each $e_i \in E'$ we add one pure strategy $\tilde{e}_i$ for player 1. Let $\tilde{E}$ 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

$$\begin{aligned}
\tilde{u}_k(Out, f_j) &= u_k(Out, f_j) && (k = 1, 2) \\
\tilde{u}_k(e_i, f_j) &= u_k(e_i, f_j) && (e_i \in E,\ k = 1, 2) \\
\tilde{u}_1(\tilde{e}_i, f_j) &= K^i\,(u_1(e_i, f_j) - \varepsilon) && (\tilde{e}_i \in \tilde{E}) \\
\tilde{u}_2(\tilde{e}_i, f_j) &= u_2(e_i, f_j) && (\tilde{e}_i \in \tilde{E})
\end{aligned}$$

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 for any strategy outside this inflated set there is at least one additional strategy which yields a positive payoff against it. 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 2 characterizes the equilibria of the newly constructed game for large values of K.

Lemma 2 *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 ε, 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 $\tilde{e}_i \in \widetilde{E}$ is in the support of s_1, then $\tilde{u}_1(e_i, s_2) = \varepsilon$.

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

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

Note that for any other $e_r \in E'$ with $u_1(e_r, s_2) > \varepsilon$ we have $r < i$ so that for large K we have

$$\tilde{u}_1(\tilde{e}_i, s_2) = K^i \left(u_1(e_i, s_2) - \varepsilon\right) > K^r \left(u_1(e_r, s_2) - \varepsilon\right) = \tilde{u}_1(\tilde{e}_r, s_2).$$

Clearly, 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 $\mathcal{B}_1(s_2^K) = \{\tilde{e}_i\}$. Because of the genericity of the game player 2 must have a unique best reply against any pure strategy of player 1. 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.[15] 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 $e_i \in E'$ 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 $e_i \in E'$ 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 $e_i \in E'$. □

(ii) We conclude from (i) that s_2 must be on the boundary of X_2^ε. For small ε s_2 must be very close to (but outside of) X_2. Hence any strategy $e_i \in E \setminus E'$ yields a negative payoff against s_2. □

(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. □

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 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 α_i, β_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 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 2. □

Lemma 3 *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 [10] 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 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 indexes 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. $\square$

Since the sum of indexes over all equilibrium components must equal $+1$ we get as an immediate corollary to Lemma 3.

Corollary 1 *Index* $(s^*) = 0$.

However, in generic games an isolated equilibrium must have index $+1$ or -1. Therefore g cannot be generic if the outside option component is an EES set. This completes the proof of Proposition 4. $\square$

5 Conclusion

When introducing a new equilibrium refinement, various authors have checked that their new refinement satisfies some form of forward induction. (E.g. [19].) 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 [1] (defined in Section 2) in the class of generic 2-person normal form games preceded by an outside option. 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 yield the forward induction equilibrium.

On the other hand, 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 equilibrium 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.

We will now try to provide some intuition for these results. The key issue is not so much whether the forward induction equilibrium is strategically stable or an EES set but rather whether the outside option component can be strategically stable or an EES set. Notice that in general, it is easier for a set to be strategically stable than to be evolutionarily stable.[16] In particular, in order for the outside option component to be strategically stable there just needs to be *some* belief (within the component) that justifies player 1's choice of *Out* for every possible strategy or payoff perturbation, where these beliefs are allowed to vary with the perturbations. On the other hand, for the outside option component to be evolutionarily stable it must hold that *none* of the strategy profiles in the set can be invaded by *any* equilibrium entrant. Hence, in principle, EES sets have a greater potential to select the forward induction equilibrium than strategic stability concepts. Of course, this does not yet explain why EES sets work while strategic stability does not. In order to shed some light on this difference let us briefly review the proof of Proposition 4 for EES sets.

In this proof we constructed a new game from the original outside option game by adding extra strategies for player 1 that are never a weak best response to the outside option component. This was done in such a way that (i) the outside option component was preserved and (ii) the forward induction equilibrium was destroyed. In general, the new game may have new equilibria. In fact, whenever the game is generic, i.e. whenever the forward induction equilibrium does not have a zero index, this must necessarily be the case since the sum of the indexes of all new equilibria must be equal to the index of the forward induction equilibrium.[17] We showed how to construct a strategy profile from these new equilibria that would be an equilibrium entrant that could invade the outside option component. Hence, the outside option component can only be an EES set if the forward induction equilibrium has zero index. In this case, the EES set must have index +1.

On the other hand, if the forward induction equilibrium s^* constitutes a (singleton) EES set, it must also have index +1. To see this point, consider the game where players are restricted to use only strategies that are best replies to s^*. s^* is the unique Nash equilibrium of this restricted game, since any other Nash equilibrium would be able to enter (see condition (S) in the definition of EES sets). Hence, the index of s^* in the restricted game is +1. Since the index of an equilibrium can be calculated from the payoff matrices restricted to the support of the equilibrium (see [10]), the index of s^* in the original game is also +1. Hence, any EES set in the class of outside option games $\overline{\Gamma}$ has index +1. This is not true for strategic stability concepts as they might also be contained in components of index -1, 0 or +2. This difference is very important since there is a connection between having index +1 and *dynamic stability*.

Although EES sets certainly have an evolutionary flavor, no dynamics is known to yield EES sets as dynamically stable sets.[18] Nevertheless, our results seem to carry through in a dynamic setting. [4, Theorem 1] have shown that a necessary condition for asymptotic stability is that the index of a component agrees with its Euler characteristic. In two player games all equilibria components are contractible and, therefore, have Euler characteristic +1. Thus, in the class of outside option games under consideration any asymptotically stable component must have index +1, like EES sets. Hence, within this class both asymptotically stable sets and EES sets are stronger than strategic stability (as components with index +1 contain Mertens- and hyperstable sets). Also, as with EES sets, the outside option component can only be asymptotically stable if the game is nongeneric. Similar to EES sets, asymptotically stable sets might not exist.

References

1. E. van Damme, "Stable Equilibria and Forward Induction," *J. Econ. Theory* **48** (1989), 476-496.

2. E. van Damme, Strategic Equilibrium, *prepared for*: Handbook of Game Theory with Economic Applications, vol. 3 (R.J. Aumann and S. Hart, Eds.) In press, 1996.

3. S. DeMichelis and F. Germano, "On the Indices of Zeros of Nash Fields," University of California at San Diego Discussion Paper 1996/33, 1996.

4. S. DeMichelis and K. Ritzberger, "From Evolutionary to Strategic Stability", mimeo, 2000.

5. D.Fudenberg and J. Tirole, "Game Theory". MIT Press, Cambridge, MA. Third edition , 1993.

6. S. Govindan, "Stability and the Chain Store Paradox," *J. Econ. Theory* **66** (1995), 536-547.

7. S. Govindan and A. Robson, "Forward Induction, Public Randomization, and Admissibility," *J. Econ. Theory* **82** (1998), 451-457.

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

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

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

11. F. Gul and D.G. Pearce , "Forward Induction and Public Randomization," *J. Econ. Theory* **70** (1996), 43-65.

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

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

14. J. Hillas and E. Kohlberg, Foundations of Strategic Equilibrium, *prepared for*: Handbook of Game Theory with Economic Applications, vol. 3 (R.J. Aumann and S. Hart, Eds.) In press, 1996.

15. Jiang Jia-he, "Essential Component of the Set of Fixed Points of the Multivalued Mappings and its Application to the Theory of Games," *Scientia Sinica* **12** (1963), 951-964.

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

17. D. Kreps and R. Wilson , "Reputation and Information," *J. Econ. Theory* **27** (1982), 253-279.

18. G.J. Mailath, L. Samuelson and J.M. Swinkels, "Extensive Form Reasoning in Normal Form Games," *Econometrica* **61** (1993), 273-302.

19. A. Matsui, "Best Response Dynamics and Socially Stable Strategies," *J. Econ. Theory* **57** (1992), 343-362.

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

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

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

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

24. J. Swinkels, "Evolutionary Stability with Equilibrium Entrants," *J. Econ. Theory* **57** (1992), 306-332.

25. J. Swinkels, "Evolution and Strategic Stability: From Maynard Smith to Kohlberg and Mertens," *J. Econ. Theory* **57** (1992), 333-342.

26. R. Wilson, 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, 1997.

Footnotes

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 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. [16] showed that any hyperstable set contains a fully stable set, which in turn contains a KM stable set.

3. Essentiality is, like hyperstability, defined with respect to payoff perturbations, but, unlike hyperstability, need not be invariant.

4. 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.

5. This reflects the desirable connectedness property from [16]. 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.

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. For an illustrative example see [8].

7. [3] have shown this for n-player games, while the proof of [9] only applies to two player games.

8. Related results are obtained by [6] and [7]. [6] shows that in the chain store paradox of [17] the unique Mertens stable outcome is the "intuitive" one, while KM stable sets may contain "unintuitive" outcomes. [7] show that admissibility restores the

forward induction logic in an example discussed by [11], who argued that forward induction loses its power when continuous public signals are added to the game.

9. To see why, first notice that perturbations to player 2's payoffs in the subgame do not affect the outside option component. Next consider player 1. If we perturb player 1's payoffs in the subgame, the outside option component is slightly shifted but has almost the same shape. The new extreme points are very close to the old extreme points. If we perturb player 1's payoffs from choosing the outside option, we can normalize the game to achieve a zero payoff for player 1 when choosing the outside option. This corresponds to perturbing player 1's payoffs in the subgame, which, as mentioned before, shifts the outside option component only slightly.

10. Since s_1 is regular, the replicator dynamics could be used to compute the index. However, we will use an alternative method developed in [10] which is only valid for two player games and can be derived directly from the payoff matrix of the game. This index of a Nash equilibrium (σ, τ) is sgn $(-1)^{k+1}(\det A')(\det B')$ where k are the number of strategies used in this Nash equilibrium and A' (B') the payoff matrix of player 1 (2) after deleting those strategies that are not used in the Nash equilibrium (σ, τ). In our above example the index of the forward induction equilibrium is sgn $(-1)^3 \left(\det \begin{pmatrix} 6 & 2 \\ 2 & 6 \end{pmatrix}\right)^2$.

11. An alternative, more constructive way of verifying that the outside option component has index zero in the game of Fig. 3 goes as follows: Consider the perturbation $\varepsilon_1 = 3\varepsilon_3 < 0$. As can be seen from Fig. 5. the perturbed game has two equilibria close to the outside option component: P_6 and one close to P_2. It is easily verified that the first has index $+1$ and the second index -1, which shows that the outside option component has index zero.

12. We thank an anonymous referee for pointing this out to us.

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

14. 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.

15. The mutant strategy could enter when the population is at some state (Out, θ_2) where θ_2 is such that $u_1(e_i, \theta_2) = 0 > u_1(e_k, \theta_2)$ for all $e_i \neq e_k$.

16. [25] has shown that for 2-player games every EES set contains a KM stable set, a fully stable set, a hyperstable set and a Hillas stable set.

17. Recall that the sum of indexes over all equilibrium components must equal +1.

18. Closest to a dynamic foundation of EES sets is [19]'s best response dynamics that yields cyclically stable sets (CSS). The main difference between the concepts of EES and CSS is that the latter do not need to consist of Nash equilibria, which by the way is the reason why CSS always exist while EES sets need not exist.