

## Fall Back Equilibrium

Kleppe, J.; Borm, P.E.M.; Hendrickx, R.L.P.

*Publication date:*  
2008

[Link to publication](#)

*Citation for published version (APA):*

Kleppe, J., Borm, P. E. M., & Hendrickx, R. L. P. (2008). *Fall Back Equilibrium*. (CentER Discussion Paper; Vol. 2008-31). Tilburg: Operations research.

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

### Take down policy

If you believe that this document breaches copyright, please contact us providing details, and we will remove access to the work immediately and investigate your claim.

No. 2008–31

**FALL BACK EQUILIBRIUM**

By John Kleppe, Peter Borm, Ruud Hendrickx

March 2008

ISSN 0924-7815

# Fall Back Equilibrium

John Kleppe<sup>1,2</sup>

Peter Borm<sup>1</sup>

Ruud Hendrickx<sup>1</sup>

March 17, 2008

## Abstract

Fall back equilibrium is a refinement of the Nash equilibrium concept. In the underlying thought experiment each player faces the possibility that, after all players decided on their action, his chosen action turns out to be blocked. Therefore, each player has to decide beforehand on a back-up action, which he plays in case he is unable to play his primary action.

In this paper we introduce the concept of fall back equilibrium and show that the set of fall back equilibria is a non-empty and closed subset of the set of Nash equilibria. We discuss the relations with other equilibrium concepts, and among other results it is shown that each robust equilibrium is fall back and for bimatrix games also each proper equilibrium is a fall back equilibrium. Furthermore, we show that for bimatrix games the set of fall back equilibria is the union of finitely many polytopes, and that the notions of fall back equilibrium and strictly fall back equilibrium coincide. Finally, we allow multiple actions to be blocked, resulting in the notion of *complete fall back equilibrium*. We show that the set of complete fall back equilibria is a non-empty and closed subset of the set of proper equilibria.

**Keywords:** strategic game, equilibrium refinement, blocked action, fall back equilibrium

**JEL Classification Number:** C72

## 1 Introduction

The notion of equilibrium for strategic games, introduced by Nash (1951), is the fundamental concept in non-cooperative game theory. The set of Nash equilibria, however, may be very large and can contain counterintuitive outcomes. In order to overcome these drawbacks Selten (1975) developed the concept of perfectness as a refinement of the Nash equilibrium concept. In the thought experiment underlying perfectness all players make mistakes in such a way that each action is played with positive probability. The notions of properness (cf. Myerson (1978)), robustness (cf. Okada (1983)), strict perfectness (cf. Okada (1984)) and many others originated from Selten's work. Although these refinements differ in their exact concept, the common underlying idea is that an equilibrium should be stable against perturbations in the strategies due to mistakes made by the players of the game. This line of research culminated into the concept of stable sets (cf. Kohlberg

---

<sup>1</sup>CentER and Department of Econometrics and Operations Research, Tilburg University.

<sup>2</sup>Corresponding author: PO Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: J.Kleppe@uvt.nl.

and Mertens (1986)). A partial overview of this literature can be found in Van Damme (1991).

In this paper we introduce a new equilibrium concept in which the strategy perturbations are based on another type of thought experiment. The idea is that each player faces a small but positive probability that, after all players decided on their action, the action chosen by him is blocked. Therefore, each player has to choose beforehand a back-up action, which he plays in case his first choice action, called primary action, is blocked.

The probability with which a player is unable to play his primary action and has to rely on his back-up is assumed to be independent of the particular choice he makes. This probability may, however, vary between the players. It is important to notice that, contrary to the perfectness concept in which players randomly play all other actions by mistake, in our setting players choose their back-up action strategically.

Consider the following  $2 \times 4$  bimatrix game, where both players are allowed to randomise between their actions:

$$\begin{matrix} e_1^1 \\ e_2^1 \end{matrix} \begin{bmatrix} e_1^2 & e_2^2 & e_3^2 & e_4^2 \\ 1, 7 & 0, 0 & 1, 5 & 1, 6 \\ 1, 7 & 1, 6 & 1, 5 & 0, 2 \end{bmatrix} .$$

Following Borm (1992) we analyse this game graphically in Figure 1.1. In this figure

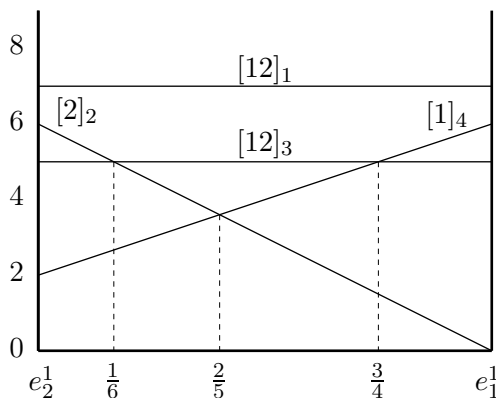


Figure 1.1: Graphical representation of a  $2 \times n$  bimatrix game.

the horizontal axis represents the strategy space of player 1, and each line describes player 2's payoff function corresponding to a particular action (indicated by the subindex). Each label displays player 1's set of pure best replies (either action 1, 2, or both) against the corresponding action of player 2.

In addition to the two proper equilibria on the boundaries of player 1's strategy space, there is a third proper equilibrium  $(\frac{2}{5}e_1^1 + \frac{3}{5}e_2^1, e_3^2)$ . Here player 1's strategy is the unique strategy for which player 2 is indifferent between his second and fourth action. Hence,

although these actions are dominated by both  $e_1^2$  and  $e_3^2$  if player 1 plays (a strategy in the neighbourhood of)  $\frac{2}{5}e_1^1 + \frac{3}{5}e_2^1$ , the coordination point of these two actions determines a proper equilibrium. The reason is that the concept of proper equilibrium, like many other concepts, assumes full rationality of all players, which in this particular example implies that player 1 has to anticipate the possibility that player 2 makes the mistake of playing  $\frac{1}{2}e_2^2 + \frac{1}{2}e_4^2$ . Clearly, such an analysis requires a high level of rationality by the players on less relevant payoff levels in the game.

We assume that players are boundedly rational in the sense that they only take into account the possibility of a single back-up for each player. This is modelled by only allowing the primary action of a player to be blocked, and not also his back-up action. The set of fall back equilibria of this game is given by  $\{(e_1^1, e_1^2)\} \cup \{(e_2^1, e_1^2)\} \cup \text{Conv}(\{\frac{1}{6}e_1^1 + \frac{5}{6}e_2^1, \frac{3}{4}e_1^1 + \frac{1}{4}e_2^1\}) \times \{e_2^2\}$ , where  $\text{Conv}(A)$  denotes the convex hull of a set  $A$ .

The idea behind fall back equilibrium can also be applied to extensive form games. In that framework one has to make a distinction between the setup with mixed strategies and the setup with behavioural strategies. In the first one players decide on all their (possible) actions beforehand, whereas in the second setup players determine each choice at the moment they actually have to make it. In games with perfect recall, this distinction has no effect on the set of Nash equilibria (cf. Kuhn (1953)). However, for the notion of fall back equilibria it does matter whether a player faces the possibility of blocked actions once at the beginning of the game, or at each choice moment separately.

In this paper, however, we have chosen to restrict attention to mixed extensions of finite non-cooperative games in strategic form. In the thought experiment players act by choosing both a primary and a back-up strategy. These strategies together define a strategy in the fall back game. Given that a player can choose between  $m$  actions in the original game, the fall back game has  $m(m - 1)$  actions to choose from, as players are not allowed to choose the same action both as primary and as back-up. The payoffs in the fall back game are the expected payoffs in the original game given the blocking probabilities. In the fall back game players are also allowed to use mixed strategies.

Consider the  $3 \times 3$  bimatrix game, which is due to Myerson (1978), given by:

$$\begin{array}{c} e_1^1 \\ e_2^1 \\ e_3^1 \end{array} \left[ \begin{array}{ccc} e_1^2 & e_2^2 & e_3^2 \\ 1, 1 & 0, 0 & -9, -9 \\ 0, 0 & 0, 0 & -7, -7 \\ -9, -9 & -7, -7 & -7, -7 \end{array} \right].$$

Out of the three Nash equilibria of this game,  $(e_1^1, e_1^2)$ ,  $(e_2^1, e_2^2)$  and  $(e_3^1, e_3^2)$ , only the last one is not a perfect equilibrium. The focal point of this game is  $(e_1^1, e_1^2)$ , and this strategy profile is also the unique proper equilibrium, as the probability that players play the third row/column by mistake is significantly smaller than the probability of making any other mistake.

The strategy profile  $(e_1^1, e_1^2)$  is also the unique fall back equilibrium of this game. In our framework, however, this is due to the fact that back-up actions are chosen strategically. Therefore, the strategy profile in which both players choose their first action as primary

strategy and their second action as back-up strategy forms the unique equilibrium in the fall back game, which supports  $(e_1^1, e_1^2)$  as the unique fall back equilibrium.

The first result we obtain in this paper is that the set of fall back equilibria is a non-empty and closed subset of the set of Nash equilibria. We also analyse the relation between fall back equilibrium on one hand and the equilibrium concepts of perfect, proper, strictly perfect and robust on the other. We prove that each robust equilibrium is a fall back equilibrium. Furthermore, for bimatrix games also each proper equilibrium is a fall back equilibrium, and consequently the intersection between the sets of fall back and perfect equilibria is non-empty. For games with more players this relation between proper and fall back equilibrium does not hold. The relation between the sets of fall back and strictly perfect equilibria is restricted to  $2 \times 2$  bimatrix games. For these games the two sets coincide, otherwise the intersection can be empty.

Similar to the way Okada (1984) refined perfectness in strict perfectness we define the concept of strictly fall back equilibrium. It turns out that the sets of fall back and strictly fall back equilibria coincide for bimatrix games. However, for games with more than two players the set of strictly fall back equilibria can be empty.

For bimatrix games also the structure of the set of fall back equilibria is analysed. The main result is that the set of fall back equilibria is the union of finitely many polytopes.

In the thought experiment underlying fall back equilibrium we assume that only one action of each player can be blocked. In the final section of this paper we analyse the equilibrium concept that emerges when we allow multiple actions of each player to be blocked. The main result provided for this concept, called *complete fall back equilibrium*, is that the set of complete fall back equilibria is a non-empty and closed subset of the set of proper equilibria.

This paper is organised as follows. In Section 2 we set up notation, formally introduce and characterise the concept of fall back equilibrium for strategic games, and present some basic results. In Section 3 we discuss the concept of strictly fall back equilibrium, while in Section 4 we consider the relations between fall back equilibrium and other equilibrium concepts. In Section 5 we discuss the structure of the set of fall back equilibria for bimatrix games and Section 6 covers the analysis of the concept of complete fall back equilibrium.

## 2 Fall back equilibrium

A non-cooperative game in strategic form is given by  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$ , with  $N = \{1, \dots, n\}$  the player set,  $\Delta_{M^i}$  the mixed strategy space of player  $i \in N$ , with  $M^i = \{1, \dots, m^i\}$  the set of pure strategies, and  $\pi^i : \prod_{j \in N} \Delta_{M^j} \rightarrow \mathbb{R}$  the von Neumann Morgenstern expected payoff function of player  $i$ . A pure strategy  $k \in M^i$  of player  $i$  is alternatively denoted by  $e_k^i$ , a typical element of  $\Delta_{M^i}$  by  $x^i$ . We denote the probability which  $x^i$  assigns to pure strategy (action)  $k$  by  $x_k^i$ . The set of all strategy profiles is given by  $\Delta = \prod_{i \in N} \Delta_{M^i}$ , a typical element of  $\Delta$  by  $x$ .

A strategy profile  $\hat{x}$  is a *Nash equilibrium* (cf. Nash (1951)) of  $G$ , denoted by  $\hat{x} \in NE(G)$ , if  $\pi^i(\hat{x}) \geq \pi^i(x^i, \hat{x}^{-i})$  for all  $x^i \in \Delta_{M^i}$  and all  $i \in N$ . Here  $(x^i, \hat{x}^{-i})$  is the frequently used

shorthand notation for the strategy profile  $(\hat{x}^1, \dots, \hat{x}^{i-1}, x^i, \hat{x}^{i+1}, \dots, \hat{x}^n)$ .

The carrier of a strategy  $x^i$  is given by  $C(x^i) = \{k \in M^i \mid x_k^i > 0\}$ , the pure best reply correspondence of player  $i$  by  $PB^i(x^{-i}) = \{k \in M^i \mid \pi^i(e_k^i, x^{-i}) \geq \pi^i(e_\ell^i, x^{-i}) \text{ for all } \ell \in M^i\}$ . Clearly,  $\hat{x} \in NE(G)$  if and only if  $C(\hat{x}^i) \subseteq PB^i(\hat{x}^{-i})$  for all  $i \in N$ .

The action set for player  $i$  in the associated fall back game (only defined if  $m^j \geq 2$  for all  $j \in N$ ) is given by  $\tilde{M}^i = \{(k, \ell) \in M^i \times M^i \mid k \neq \ell\}$ . Hence, the total number of actions in the fall back game for player  $i$  is  $\tilde{m}^i = m^i(m^i - 1)$ . An action  $(k, \ell) \in \tilde{M}^i$  consists of a *primary* action  $k$  and a *back-up* action  $\ell$ . Let  $\varepsilon = (\varepsilon^1, \dots, \varepsilon^n)$  be an  $n$ -tuple of (small) non-negative probabilities. The interpretation of player  $i$ 's action  $(k, \ell)$  in the fall back game is that he plays in the original game with probability  $1 - \varepsilon^i$  primary action  $k$  and with probability  $\varepsilon^i$  back-up action  $\ell$ .

The fall back game  $\tilde{G}(\varepsilon)$  is given by  $\tilde{G}(\varepsilon) = (N, \{\Delta_{\tilde{M}^i}\}_{i \in N}, \{\pi_\varepsilon^i\}_{i \in N})$ , with  $\pi_\varepsilon^i : \prod_{j \in N} \Delta_{\tilde{M}^j} \rightarrow \mathbb{R}$  the extended expected payoff function of player  $i$ . Let pure strategy  $(k, \ell) \in \tilde{M}^i$  be alternatively denoted by  $e_{k\ell}^i$ . Then, the payoff function  $\pi_\varepsilon^i$  is formally defined by

$$\pi_\varepsilon^i((e_{k^j \ell^j}^j)_{j \in N}) = \sum_{S \subseteq N} \left( \prod_{j \in S} (1 - \varepsilon^j) \prod_{j \in N \setminus S} \varepsilon^j \right) \pi^i((e_{k^j}^j)_{j \in S}, (e_{\ell^j}^j)_{j \in N \setminus S}).$$

A typical element of  $\Delta_{\tilde{M}^i}$  is denoted by  $\rho^i$ , where  $\rho_{k\ell}^i$  is the probability which  $\rho^i$  assigns to pure strategy  $(k, \ell)$ . Note that  $\rho^i$  assigns probabilities to pure strategies  $(k, \ell)$  of the fall back game, not to primary and back-up actions separately. The set of all strategy profiles is given by  $\tilde{\Delta} = \prod_{i \in N} \Delta_{\tilde{M}^i}$ , an element of  $\tilde{\Delta}$  will be denoted by  $\rho$ .

**Definition** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. A strategy profile  $x \in \Delta$  is called a *fall back equilibrium* of  $G$  if there exists a sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  of  $n$ -tuples of positive real numbers converging to zero, and a sequence  $\{\rho_t\}_{t \in \mathbb{N}}$  such that  $\rho_t \in NE(\tilde{G}(\varepsilon_t))$  for all  $t \in \mathbb{N}$ , converging to  $\rho \in \tilde{\Delta}$ , with  $x_k^i = \sum_{\ell \in M^i \setminus \{k\}} \rho_{k\ell}^i$  for all  $k \in M^i$  and all  $i \in N$ . The set of fall back equilibria of  $G$  is denoted by  $FBE(G)$ .

In the thought experiment underlying the concept of fall back equilibrium each player faces the small but positive probability that, after all players decided on their action, the action chosen by him is blocked. In that case the player plays a back-up action, he choose beforehand. This is modelled by letting players play the fall back game in which each action consists of a primary action, played with a probability close to one, and a back-up action, played with the remaining probability. A fall back equilibrium of the original game is then deduced from the limit point of a sequence of Nash equilibria of the corresponding fall back games when the blocking probabilities converge to zero.

**Theorem 2.1** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. Then  $FBE(G)$  is a non-empty and closed subset of  $NE(G)$ .

**Proof:** We first show non-emptiness. Let  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  be a sequence of  $n$ -tuples of positive real numbers converging to zero. Take a sequence  $\{\rho_t\}_{t \in \mathbb{N}}$  such that  $\rho_t \in NE(\tilde{G}(\varepsilon_t))$  for all  $t \in \mathbb{N}$ . Because the strategy spaces are compact there exists a subsequence of  $\{\rho_t\}_{t \in \mathbb{N}}$

converging to, say,  $\rho \in \Delta$ . Define  $x \in \Delta$  by  $x_k^i = \sum_{\ell \in M^i \setminus \{k\}} \rho_{k\ell}^i$  for all  $k \in M^i$  and all  $i \in N$ . By definition  $x \in FBE(G)$ .

Next we prove that each fall back equilibrium is a Nash equilibrium. Take  $x \in FBE(G)$ . We prove that  $x \in NE(G)$  by showing that  $C(x^i) \subseteq PB^i(x^{-i})$  for all  $i \in N$ . Take a sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  of  $n$ -tuples of positive real numbers converging to zero and a sequence  $\{\rho_t\}_{t \in \mathbb{N}}$  such that  $\rho_t \in NE(\tilde{G}(\varepsilon_t))$  for all  $t \in \mathbb{N}$ , converging to  $\rho \in \tilde{\Delta}$ , with  $x_k^i = \sum_{\ell \in M^i \setminus \{k\}} \rho_{k\ell}^i$  for all  $k \in M^i$  and all  $i \in N$ . Let  $i \in N$  and  $k \in C(x^i)$ . Then for sufficiently large  $t \in \mathbb{N}$  we have that  $(k, \ell) \in C(\rho_t^i)$  for some  $\ell \in M^i \setminus \{k\}$ . Hence,

$$\pi_{\varepsilon_t}^i(e_{k\ell}^i, \rho_t^{-i}) \geq \pi_{\varepsilon_t}^i(e_{rs}^i, \rho_t^{-i})$$

for every  $(r, s) \in \tilde{M}^i$ . Taking  $t$  to infinity, we find by continuity of  $\pi_{\varepsilon}^i$ ,

$$\pi_0^i(e_{k\ell}^i, \rho^{-i}) \geq \pi_0^i(e_{rs}^i, \rho^{-i})$$

for every  $(r, s) \in \tilde{M}^i$ . Since  $\pi_0^i(e_{k\ell}^i, \rho^{-i}) = \pi^i(e_k^i, x^{-i})$  and similarly  $\pi_0^i(e_{rs}^i, \rho^{-i}) = \pi^i(e_r^i, x^{-i})$ , it follows that

$$\pi^i(e_k^i, x^{-i}) \geq \pi^i(e_r^i, x^{-i})$$

for every  $r \in M^i$  and hence  $k \in PB^i(x^{-i})$ .

Finally we show that  $FBE(G)$  is closed. Take a converging sequence  $\{x_t\}_{t \in \mathbb{N}}$  with  $x_t \in FBE(G)$  for all  $t \in \mathbb{N}$ , with limit  $x$ . For all  $t \in \mathbb{N}$  there exists a sequence  $\{\varepsilon_{tr}\}_{r \in \mathbb{N}}$  of  $n$ -tuples of positive real numbers converging to zero and a sequence  $\{\rho_{tr}\}_{r \in \mathbb{N}}$  converging to  $\rho_t$ , with  $x_{t,k}^i = \sum_{\ell \in M^i \setminus \{k\}} \rho_{t,k\ell}^i$  for all  $k \in M^i$  and all  $i \in N$ , such that

$$x_{tr} \in NE(\tilde{G}(\varepsilon_{tr}))$$

for all  $r \in \mathbb{N}$ . Considering the sequences  $\{\varepsilon_{tt}\}_{t \in \mathbb{N}}$  and  $\{x_{tt}\}_{t \in \mathbb{N}}$  one readily establishes that  $x \in FBE(G)$ .  $\square$

Although the definition of fall back equilibrium is natural in its interpretation, the fact that the size of the payoff matrices is larger in the fall back game than in the original game makes further analysis complicated. Therefore, we now provide an alternative characterization of fall back equilibrium.

For a (sufficiently small) blocking vector  $\delta \in \mathbb{R}_+^N$ , the blocking game  $G(\delta) = (N, \{\Delta_{M^i}(\delta^i)\}_{i \in N}, \{\pi^i\}_{i \in N})$  is defined to be the game which only differs from  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  in the sense that the strategy spaces are restricted to

$$\Delta_{M^i}(\delta^i) = \{x^i \in \Delta_{M^i} \mid x_k^i \leq 1 - \delta^i \text{ for all } k \in M^i\}$$

for all  $i \in N$  and the domains of the payoff functions are restricted accordingly. Define the set of all strategy profiles of the blocking game by  $\Delta(\delta) = \prod_{j \in N} \Delta_{M^j}(\delta^j)$ .

Note that the strategy spaces of the blocking game, with  $\delta > 0$ , restrict each player to play at least two of his original actions with positive probability, but also allow him to play some actions with zero probability.



**Lemma 2.2** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. Let  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  and  $\{\delta_t\}_{t \in \mathbb{N}}$  be sequences of  $n$ -tuples of positive real numbers converging to zero such that  $\varepsilon_t = \delta_t$  for all  $t \in \mathbb{N}$ , with corresponding fall back and blocking games  $\tilde{G}(\varepsilon_t) = (N, \{\Delta_{\tilde{M}^i}\}_{i \in N}, \{\pi_{\varepsilon_t}^i\}_{i \in N})$  and  $G(\delta_t) = (N, \{\Delta_{M^i}(\delta_t)\}_{i \in N}, \{\pi^i\}_{i \in N})$  respectively.

Then for each sequence  $\{\rho_t\}_{t \in \mathbb{N}}$  converging to  $\rho$ , with  $\rho_t \in \tilde{\Delta}$  for all  $t \in \mathbb{N}$ , there exists a sequence  $\{x_t\}_{t \in \mathbb{N}}$  converging to  $x$ , with  $x_t \in \Delta(\delta_t)$  for all  $t \in \mathbb{N}$ , such that  $x_k^i = \sum_{\ell \in M^i \setminus \{k\}} \rho_{k\ell}^i$  for all  $k \in M^i$  and all  $i \in N$ , and  $\pi^i(x_t) = \pi_{\varepsilon_t}^i(\rho_t)$  for all  $i \in N$  and all  $t \in \mathbb{N}$ .

Conversely, for each sequence  $\{x_t\}_{t \in \mathbb{N}}$  converging to  $x$ , with  $x_t \in \Delta(\delta_t)$  for all  $t \in \mathbb{N}$ , there exists a sequence  $\{\rho_t\}_{t \in \mathbb{N}}$  converging to  $\rho$ , with  $\rho_t \in \tilde{\Delta}$  for all  $t \in \mathbb{N}$ , such that  $x_k^i = \sum_{\ell \in M^i \setminus \{k\}} \rho_{k\ell}^i$  for all  $k \in M^i$  and all  $i \in N$ , and  $\pi_{\varepsilon_t}^i(\rho_t) = \pi^i(x_t)$  for all  $i \in N$  and all  $t \in \mathbb{N}$ .

**Proof:** Let  $\{\rho_t\}_{t \in \mathbb{N}}$  be a sequence converging to  $\rho \in \tilde{\Delta}$ , with  $\rho_t \in \tilde{\Delta}$  for all  $t \in \mathbb{N}$ . We define the sequence  $\{x_t\}_{t \in \mathbb{N}}$  such that for all  $t \in \mathbb{N}$

$$x_{t,k}^i = (1 - \delta_t^i) \sum_{\ell \in M^i \setminus \{k\}} \rho_{t,k\ell}^i + \delta_t^i \sum_{\ell \in M^i \setminus \{k\}} \rho_{t,\ell k}^i \quad (1)$$

for all  $k \in M^i$  and all  $i \in N$ . Then  $x_t \in \Delta(\delta_t)$  for all  $t \in \mathbb{N}$  and  $\{x_t\}_{t \in \mathbb{N}}$  converges to  $x$ , with  $x_k^i = \sum_{\ell \in M^i \setminus \{k\}} \rho_{k\ell}^i$  for all  $k \in M^i$  and all  $i \in N$ . Furthermore, because  $\varepsilon_t = \delta_t$  for all  $t \in \mathbb{N}$  the strategy profile  $x_t$  puts the same probabilities on the actions of the game  $G$  as  $\rho_t$  for all  $t \in \mathbb{N}$ . Therefore,  $\pi^i(x_t) = \pi_{\varepsilon_t}^i(\rho_t)$  for all  $i \in N$  and all  $t \in \mathbb{N}$ .

The reverse statement is shown similarly, with the sequence  $\{\rho_t\}_{t \in \mathbb{N}}$  defined in such a way that equation (1) is satisfied. Note that since  $x_t \in \Delta(\delta_t)$  and  $\varepsilon_t = \delta_t$  for all  $t \in \mathbb{N}$ , and therefore at least two actions of each player  $i$  of game  $G$  are played with a probability of at least  $\varepsilon_t^i$ , it is always possible to construct such a sequence  $\{\rho_t\}_{t \in \mathbb{N}}$ .  $\square$

As a consequence of Lemma 2.2, a fall back equilibrium can also be defined in terms of Nash equilibria of blocking games.

**Theorem 2.3** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. Then, a strategy profile  $x \in \Delta$  is a fall back equilibrium of  $G$  if and only if there exists a sequence  $\{\delta_t\}_{t \in \mathbb{N}}$  of blocking vectors of positive real numbers converging to zero and a sequence  $\{x_t\}_{t \in \mathbb{N}}$  converging to  $x$  such that  $x_t \in NE(G(\delta_t))$  for all  $t \in \mathbb{N}$ .

**Proof:** We just prove the “only if” part, the reverse statement can be shown analogously. Assume  $\hat{x} \in FBE(G)$ . Then by definition there exists a sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  of  $n$ -tuples of positive real numbers converging to zero, and a sequence  $\{\hat{\rho}_t\}_{t \in \mathbb{N}}$  converging to  $\hat{\rho} \in \tilde{\Delta}$ , with  $\hat{x}_k^i = \sum_{\ell \in M^i \setminus \{k\}} \hat{\rho}_{k\ell}^i$  for every  $k \in M^i$  and all  $i \in N$ , such that  $\hat{\rho}_t \in NE(\tilde{G}(\varepsilon_t))$  for all  $t \in \mathbb{N}$ . By Lemma 2.2 there exists a sequence  $\{\hat{x}_t\}_{t \in \mathbb{N}}$  converging to  $\hat{x} \in \Delta$ , with  $\hat{x}_t \in \Delta(\delta_t)$  for all  $t \in \mathbb{N}$ , such that  $\pi^i(\hat{x}_t) = \pi_{\varepsilon_t}^i(\hat{\rho}_t)$  for all  $i \in N$  and all  $t \in \mathbb{N}$ .

Let  $i \in N$ . We show that  $\pi^i(\hat{x}_t) \geq \pi^i(x_t^i, \hat{x}_t^{-i})$  for all  $x_t^i \in \Delta_{M^i}(\delta_t^i)$  and all  $t \in \mathbb{N}$ ,

which proves that  $\hat{x}_t \in NE(G(\delta_t))$  for all  $t \in \mathbb{N}$  and therefore completes the proof. Let  $t \in \mathbb{N}$  and let  $(x_t^i, \hat{x}_t^{-i}) \in \Delta(\delta_t)$ . Then by Lemma 2.2 we can take a strategy  $(\rho_t^i, \hat{\rho}_t^{-i}) \in \tilde{\Delta}$  such that  $\pi_{\varepsilon_t}^i(\rho_t^i, \hat{\rho}_t^{-i}) = \pi^i(x_t^i, \hat{x}_t^{-i})$ .

Since  $\hat{\rho}_t \in NE(\tilde{G}(\varepsilon_t))$  we obtain

$$\begin{aligned} \pi^i(x_t^i, \hat{x}_t^{-i}) &= \pi_{\varepsilon_t}^i(\rho_t^i, \hat{\rho}_t^{-i}) \\ &\leq \pi_{\varepsilon_t}^i(\hat{\rho}_t) \\ &= \pi^i(\hat{x}_t). \end{aligned}$$

Consequently,  $\pi^i(\hat{x}_t) \geq \pi^i(x_t^i, \hat{x}_t^{-i})$  for all  $x_t^i \in \Delta_{M^i}(\delta_t^i)$  and all  $t \in \mathbb{N}$ .  $\square$

Since a blocking game with  $\delta > 0$  only excludes the possibility for any player to play an original action with probability one, we obtain the following proposition.

**Proposition 2.4** *Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game and let  $x \in \Delta$  be such that  $|C(x^i)| > 1$  for all  $i \in N$ . Then  $x \in FBE(G)$  if and only if  $x \in NE(G)$ .*

In the thought experiment underlying perfectness (cf. Selten (1975)) players also play a perturbed game. Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. A perturbation vector for player  $i \in N$  is given by  $\varepsilon^i \in \mathbb{R}^{M^i}$ , with  $\varepsilon_k^i > 0$  for all  $k \in M^i$  and  $\sum_{k \in M^i} \varepsilon_k^i \leq 1$ . Then the  $\varepsilon$ -perturbed game  $H(\varepsilon) = (N, \{\Delta_{M^i}(\varepsilon^i)\}_{i \in N}, \{\pi^i\}_{i \in N})$  is defined to be the game which only differs from  $G$  in the sense that the strategy spaces are restricted to

$$\Delta_{M^i}(\varepsilon^i) = \{x^i \in \Delta_{M^i} \mid x_k^i \geq \varepsilon_k^i \text{ for all } k \in M^i\}$$

for all  $i \in N$  and the domains of the payoff functions are restricted accordingly. Define the set of all strategy profiles of the  $\varepsilon$ -perturbed game by  $\Delta(\varepsilon) = \prod_{j \in N} \Delta_{M^j}(\varepsilon^j)$ .

**Definition** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. A strategy profile  $x \in \Delta$  is called a *perfect equilibrium* of  $G$  if there exists a sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  of perturbation vectors converging to zero, and a sequence  $\{x_t\}_{t \in \mathbb{N}}$  converging to  $x$ , such that  $x_t \in NE(H(\varepsilon_t))$  for all  $t \in \mathbb{N}$ .

The difference between fall back and perfect equilibrium is that in the thought experiment underlying perfectness all actions have to be played with positive probability and for fall back equilibria only at least two. This observation leads to the following proposition.

**Proposition 2.5** *Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game such that  $m^i = 2$  for all  $i \in N$ . Then the sets of fall back and perfect equilibria coincide.*

### 3 Strictly fall back equilibrium

Okada (1984) refined the perfectness concept to strict perfectness by requiring that for *every* sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  of perturbation vectors converging to zero there exists a sequence  $\{x_t\}_{t \in \mathbb{N}}$  of strategy profiles converging to  $x$  such that  $x_t \in NE(G(\varepsilon_t))$  for all  $t \in \mathbb{N}$ . In a similar way we introduce the concept of strictly fall back equilibrium.

**Definition** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. A strategy profile  $x \in \Delta$  is called a *strictly fall back equilibrium* of  $G$  if for *every* sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  of  $n$ -tuples of positive real numbers converging to zero there exists a sequence  $\{\rho_t\}_{t \in \mathbb{N}}$  such that  $\rho_t \in NE(\tilde{G}(\varepsilon_t))$  for all  $t \in \mathbb{N}$ , converging to  $\rho \in \tilde{\Delta}$ , with  $x_k^i = \sum_{\ell \in M^i \setminus \{k\}} \rho_{k\ell}^i$  for all  $k \in M^i$  and all  $i \in N$ . The set of strictly fall back equilibria of a game  $G$  is denoted by  $SFBE(G)$ .

Note that if we impose in Theorem 2.3 the requirement *for every* sequence  $\{\delta_t\}_{t \in \mathbb{N}}$  of blocking vectors of positive real numbers converging to zero, we get in the same way an equivalent characterisation of strictly fall back equilibrium in terms of blocking vectors.

The sets of fall back and perfect equilibria are refined to a strict concept in the same way, which means that by the use of Proposition 2.5 also the sets of strictly fall back and strictly perfect equilibria coincide for all games in which each player has only two actions available. Since the set of strictly perfect equilibria can be empty for three-player games with action sets of size two for all players, also the set of strictly fall back equilibria can be empty if the number of players is three.

However, for any strategic game with only two players, i.e. bimatrix games, the set of strictly fall back equilibria is non-empty since in that case it coincides with the non-empty set of fall back equilibria. Before we can prove this result we first have to provide a second characterisation of fall back equilibrium, which can only be applied to bimatrix games. This characterisation is convenient as it does not make use of perturbed games or converging sequences.

Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game and let  $x \in \Delta$ . Then, we first of all define the set of pure second best replies of player  $i$  by

$$PSB^i(x^{-i}) = \left\{ k \in M^i \mid \begin{aligned} &\exists \ell \in M^i : \pi^i(e_\ell^i, x^{-i}) \geq \pi^i(e_k^i, x^{-i}), \\ &\pi^i(e_k^i, x^{-i}) \geq \pi^i(e_r^i, x^{-i}) \forall r \in M^i \setminus \{\ell\} \end{aligned} \right\}.$$

Note that if  $|PB^i(x^{-i})| > 1$ , then  $PB^i(x^{-i}) = PSB^i(x^{-i})$ . Also note that the correspondences  $PB^i$  and  $PSB^i$  are upper-semi-continuous.

In the blocking game the strategy of each player is composed of primary and back-up strategies. The preferences of player  $i$  over possible actions are independent of  $\delta^i$ , as they only depend on the strategies of the other players. Furthermore, if the probability on the back-up strategies,  $\delta^j$ , of player  $j \neq i$  in a blocking game corresponding to a bimatrix game is sufficiently close to zero, the set of best replies for player  $i$  in the fall back game

is, by upper-semi-continuity of  $PB^i$  and  $PSB^i$ , the same for all  $\delta^j \in (0, \bar{\delta}^j]$ . Consequently, the best replies of both players in a blocking game corresponding to a bimatrix game are independent of the blocking vector  $\delta \in \mathbb{R}_{++}^2$  when  $\delta$  is sufficiently close to zero.

**Proposition 3.1** *Let  $G = (N, \{\Delta_{M^i}\}_{i \in \{1,2\}}, \{\pi^i\}_{i \in \{1,2\}})$  be a bimatrix game. Then a strategy profile  $x = (x^1, x^2) \in \Delta$  is a fall back equilibrium if and only if one of the following three statements is satisfied.*

1.  $|C(x^1)| > 1$ ,  $|C(x^2)| > 1$  and  $x \in NE(G)$ .
2. For  $i, j \in \{1, 2\}, i \neq j$ :  $|C(x^i)| > 1$ ,  $|C(x^j)| = 1$  and there exists a strategy  $\bar{x}^j \in \Delta_{M^j}$  such that  $C(\bar{x}^j) \cap C(x^j) = \emptyset$  and a blocking probability  $\bar{\delta}^j > 0$ , such that for all  $\delta^j \in (0, \bar{\delta}^j]$  the strategy profile  $\hat{x} = (x^i, \hat{x}^j)$ , with  $\hat{x}^j = (1 - \delta^j)x^j + \delta^j \bar{x}^j$ , satisfies

$$\begin{aligned} C(x^i) &\subseteq PB^i(\hat{x}^j), \\ C(x^j) &\subseteq PB^j(x^i), \\ C(\bar{x}^j) &\subseteq PSB^j(x^i). \end{aligned}$$

3.  $|C(x^1)| = |C(x^2)| = 1$  and there exists for all  $i \in \{1, 2\}$  a strategy  $\bar{x}^i \in \Delta_{M^i}$  such that  $C(\bar{x}^i) \cap C(x^i) = \emptyset$  and a blocking probability  $\bar{\delta}^i > 0$ , such that for all  $\delta \in \mathbb{R}_{++}^2$ , with  $\delta^i \in (0, \bar{\delta}^i]$  for all  $i \in \{1, 2\}$ , the strategy profile  $(\hat{x}^1, \hat{x}^2)$ , with  $\hat{x}^i = (1 - \delta^i)x^i + \delta^i \bar{x}^i$  for all  $i \in \{1, 2\}$ , satisfies

$$\begin{aligned} C(x^1) &\subseteq PB^1(\hat{x}^2), \\ C(x^2) &\subseteq PB^2(\hat{x}^1), \\ C(\bar{x}^1) &\subseteq PSB^1(\hat{x}^2), \\ C(\bar{x}^2) &\subseteq PSB^2(\hat{x}^1). \end{aligned}$$

**Proof:** We first prove the “if” part. We do this by distinguishing between the three cases. If the first statement is satisfied, the result follows immediately from Proposition 2.4.

Next assume that the second statement is satisfied. Let  $\{\delta_t\}_{t \in \mathbb{N}}$  be a sequence of pairs of positive real numbers converging to zero. Define the sequence  $\{\hat{x}_t\}_{t \in \mathbb{N}}$  such that

$$\begin{aligned} \hat{x}_t^i &= x^i, \\ \hat{x}_t^j &= (1 - \delta_t^j)x^j + \delta_t^j \bar{x}^j \end{aligned}$$

for all  $t \in \mathbb{N}$ . Then the sequence  $\{\hat{x}_t\}_{t \in \mathbb{N}}$  converges to  $x$  and  $x_t \in \Delta(\delta_t)$  for all  $t \in \mathbb{N}$ . Let  $t \in \mathbb{N}$  be such that  $\delta_t^j \leq \bar{\delta}^j$ . We show that  $x_t \in NE(G(\delta_t))$  by showing that both players play best replies in  $G(\delta_t)$ . Since  $C(x^i) \subseteq PB^i(\hat{x}^j)$  player  $i$  plays by playing  $\hat{x}_t^i$  a best reply against  $\hat{x}_t^j$  in  $G(\delta_t)$ . Furthermore, since  $C(x^j) \subseteq PB^j(x^i)$  player  $j$  puts by playing  $\hat{x}_t^j$  maximal probability on best reply actions against  $\hat{x}_t^i$ , and as  $C(\bar{x}^j) \subseteq PB^j(\hat{x}^i)$  he puts the remaining probability on second best reply actions against  $\hat{x}_t^i$ .

We finally assume that the third statement holds. Let  $\{\delta_t\}_{t \in \mathbb{N}}$  be a sequence of pairs of positive real numbers converging to zero. Define the sequence  $\{\hat{x}_t\}_{t \in \mathbb{N}}$  such that for all  $i \in \{1, 2\}$

$$\hat{x}_t^i = (1 - \delta_t^i)x^i + \delta_t^i \bar{x}^i$$

for all  $t \in \mathbb{N}$ . Then the sequence  $\{\hat{x}_t\}_{t \in \mathbb{N}}$  converges to  $x$  and  $x_t \in \Delta(\delta_t)$  for all  $t \in \mathbb{N}$ . Let  $t \in \mathbb{N}$  be such that  $\delta_t^i \leq \bar{\delta}^i$  for all  $i \in \{1, 2\}$ . We prove that  $x_t \in NE(G(\delta_t))$  by showing that player 1 plays a best reply in  $G(\delta_t)$ . Showing that player 2 plays a best reply can be done analogously. Since,  $C(x^1) \subseteq PB^1(\hat{x}^2)$  player 1 puts by playing  $\hat{x}_t^1$  maximal probability on best reply actions against  $\hat{x}_t^2$ , and as  $C(\bar{x}^1) \subseteq PB^1(\hat{x}^2)$  the remaining probability on second best reply actions against  $\hat{x}_t^2$ . Hence, player 1 plays a best reply in  $G(\delta_t)$ .

We now show the “only if” part. Let  $x \in FBE(G)$ . If  $|C(x^1)| > 1$  and  $|C(x^2)| > 1$  it follows from Proposition 2.4 that the first statement is satisfied. Otherwise, by Theorem 2.3 there exists a sequence of blocking vectors  $\{\delta_t\}_{t \in \mathbb{N}}$  of positive real numbers converging to zero and a sequence of strategy profiles  $\{x_t\}_{t \in \mathbb{N}}$  converging to  $x$  such that  $x_t \in NE(G(\delta_t))$  for all  $t \in \mathbb{N}$ .

Let  $i \in \{1, 2\}$ . Since  $x_t \in \Delta(\delta_t)$  for all  $t \in \mathbb{N}$ , if  $|C(x^i)| = 1$ , then for all  $t \in \mathbb{N}$  there exists a strategy  $\bar{x}_t^i \in \Delta_{M^i}$ , with  $C(\bar{x}_t^i) \cap C(x^i) = \emptyset$ , such that  $x_t^i = (1 - \xi_t^i)x^i + \xi_t^i \bar{x}_t^i$ , with  $\xi_t^i \geq \delta_t^i$ . Take  $\hat{t} \in \mathbb{N}$  sufficiently large, and define

$$\hat{x}^i(\xi^i) = (1 - \xi^i)x^i + \xi^i \bar{x}_t^i.$$

Then, by the upper-semi-continuity of  $PB^j$  and  $PSB^j$ , we obtain that  $PB^j(\hat{x}^i(\xi^i)) = PB^j(\hat{x}^i(\tilde{\xi}^i))$  and  $PSB^j(\hat{x}^i(\xi^i)) = PSB^j(\hat{x}^i(\tilde{\xi}^i))$  for all  $\xi^i, \tilde{\xi}^i \in (0, \xi_t^i]$ .

Take for all  $i \in \{1, 2\}$  some  $\delta^i \in (0, \xi_t^i]$  and define  $\hat{x}$  such that

$$\hat{x}^i = \begin{cases} x^i & \text{if } |C(x^i)| > 1, \\ \hat{x}^i(\delta^i) & \text{if } |C(x^i)| = 1 \end{cases}$$

for all  $i \in \{1, 2\}$ . Then we show that either the second or the third statement is satisfied.

Let us first assume that  $|C(x^i)| > 1$  and  $|C(x^j)| = 1$ . We show that for strategy  $\hat{x} = (x^i, \hat{x}^j(\delta^j))$  it holds that  $C(x^i) \subseteq PB^i(\hat{x}^j(\delta^j))$ ,  $C(x^j) \subseteq PB^j(x^i)$  and  $C(\bar{x}_t^j) \subseteq PSB^j(x^i)$ . Note that  $x_t^i$  is a best reply against  $x_t^j$  in  $G(\delta_t)$ . Since  $x_t^i$  is close to  $x^i$  (as  $x_t^i$  converges to  $x^i$  and  $\hat{t}$  was chosen sufficiently large) we obtain that  $C(x^i) \subseteq PB^i(x_t^j)$ . Because  $PB^i(x_t^j) = PB^i(\hat{x}^j(\delta^j))$  we obtain  $C(x^i) \subseteq PB^i(\hat{x}^j(\delta^j))$ . Then,  $x \in FBE(G)$  and therefore  $x \in NE(G)$ , which immediately gives the second result that  $C(x^j) \subseteq PB^j(x^i)$ . Furthermore, since  $x_t^j$  is a best reply against  $x_t^i$  in  $G(\delta_t)$ ,  $C(\bar{x}_t^j) \subseteq PSB^j(x_t^i)$ , and as  $PSB^j(x_t^i) = PSB^j(x^i)$  it follows that  $C(\bar{x}_t^j) \subseteq PSB^j(x^i)$ .

We now consider the case  $|C(x^1)| = |C(x^2)| = 1$ . We show that for  $\hat{x} = (\hat{x}^1(\delta^1), \hat{x}^2(\delta^2))$  it

holds that  $C(x^1) \subseteq PB^1(\hat{x}^2(\delta^2))$  and  $C(\bar{x}^1) \subseteq PSB^1(\hat{x}^2(\delta^2))$ . Showing that  $C(x^2) \subseteq PB^2(\hat{x}^1(\delta^1))$  and  $C(\bar{x}^2) \subseteq PSB^2(\hat{x}^1(\delta^1))$  can be done analogously. Since  $x_{\hat{t}}^1$  is a best reply against  $x_{\hat{t}}^2$  in  $G(\delta_{\hat{t}})$  it must hold (since  $\hat{t}$  was chosen sufficiently large) that  $C(x^1) \subseteq PB^1(x_{\hat{t}}^2)$  and  $C(\bar{x}^1) \subseteq PSB^1(x_{\hat{t}}^2)$ . As  $PB^1(x_{\hat{t}}^2) = PB^1(\hat{x}^2(\delta^2))$  and  $PSB^1(x_{\hat{t}}^2) = PSB^1(\hat{x}^2(\delta^2))$  we obtain both  $C(x^1) \subseteq PB^1(\hat{x}^2(\delta^2))$  and  $C(\bar{x}^1) \subseteq PSB^1(\hat{x}^2(\delta^2))$ .  $\square$

By the use of this proposition we can prove the following theorem.

**Theorem 3.2** Let  $G = (N, \{\Delta_{M^i}\}_{i \in \{1,2\}}, \{\pi^i\}_{i \in \{1,2\}})$  be a bimatrix game. Then the sets  $FBE(G)$  and  $SFBE(G)$  coincide.

**Proof:** Since the set of strictly fall back equilibria refines the set of fall back equilibria we only have to show that  $FBE(G) \subseteq SFBE(G)$ . Let  $x \in FBE(G)$ . Then one of the three statements of Proposition 3.1 is satisfied. Let  $\{\delta_t\}_{t \in \mathbb{N}}$  be a sequence of blocking vectors of positive real numbers converging to zero and let for all  $i \in \{1, 2\}$ ,  $\bar{x}^i \in \Delta_{M^i}$  be such that it fulfills all conditions of the satisfied statement. We define the sequence  $\{\hat{x}_t\}_{t \in \mathbb{N}}$  such that for all  $t \in \mathbb{N}$

$$\hat{x}_t^i = \begin{cases} x^i & \text{if } |C(x^i)| > 1, \\ (1 - \delta_t^i)x^i + \delta_t^i \bar{x}^i & \text{if } |C(x^i)| = 1 \end{cases}$$

for all  $i \in \{1, 2\}$ . Then the sequence  $\{\hat{x}_t\}_{t \in \mathbb{N}}$  converges to  $x$ ,  $\hat{x}_t \in \Delta(\delta_t)$  for all  $t \in \mathbb{N}$  and for  $t \in \mathbb{N}$  sufficiently large  $\hat{x}_t \in NE(G(\delta_t))$ . Consequently,  $x \in SFBE(G)$ .  $\square$

## 4 Relations to other refinements

In this section we discuss the relation of fall back equilibrium to the concepts of perfect, proper, strictly perfect and robust equilibrium. We start with the relation between fall back and proper equilibrium (cf. Myerson (1978)).

**Definition** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. A strategy profile  $x \in \Delta$  is called a *proper equilibrium* of  $G$  if there exists a sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  of positive real numbers converging to zero, and a sequence  $\{x_t\}_{t \in \mathbb{N}}$  of completely mixed strategy profiles converging to  $x$  such that  $x_t$  is  $\varepsilon_t$ -proper for all  $t \in \mathbb{N}$ , i.e.,

$$\pi^i(e_k^i, x_t^{-i}) < \pi^i(e_\ell^i, x_t^{-i}) \Rightarrow x_{t,k}^i \leq \varepsilon_t x_{t,\ell}^i \quad (2)$$

for all  $k, \ell \in M^i$  and all  $i \in N$ .

Note that by replacing  $\varepsilon_t x_{t,\ell}^i$  on the right hand side of equation (2) by  $\varepsilon_t$ , one obtains an alternative characterisation of perfect equilibrium.

The relation between fall back and proper equilibrium is such that for any bimatrix game each proper equilibrium is a fall back equilibrium, which is illustrated by the first example of Section 1. Only the sets of pure best and pure second best replies determine whether a strategy profile is a fall back equilibrium, which can also be seen in the bimatrix game

characterisation of Proposition 3.1. In the concept of proper equilibria however, all lower-level sets of best replies may be relevant as well. Hence, for bimatrix games, any strategy profile that satisfies the conditions for proper equilibrium also satisfies the conditions for fall back equilibrium.

**Theorem 4.1** Let  $G = (N, \{\Delta_{M^i}\}_{i \in \{1,2\}}, \{\pi^i\}_{i \in \{1,2\}})$  be a bimatrix game. Then, each proper equilibrium of  $G$  is a fall back equilibrium of  $G$ .

**Proof:** Let  $x \in \Delta$  be a proper equilibrium. Then by definition there exists a sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  of positive real numbers converging to zero and a sequence  $\{x_t\}_{t \in \mathbb{N}}$  of completely mixed strategy profiles converging to  $x$  such that  $x_t$  is  $\varepsilon_t$ -proper for all  $t \in \mathbb{N}$ , i.e.,

$$\pi^i(e_k^i, x_t^{-i}) < \pi^i(e_\ell^i, x_t^{-i}) \Rightarrow x_{t,k}^i \leq \varepsilon_t x_{t,\ell}^i$$

for all  $k, \ell \in M^i$  and all  $i \in N$ .

We show that for this particular  $x$  one of the three statements of Proposition 3.1 is satisfied and hence that  $x \in FBE(G)$ . If  $|C(x^1)| > 1$  and  $|C(x^2)| > 1$ , then the fact that each proper equilibrium is a Nash equilibrium gives by Proposition 2.4 that statement 1 of Proposition 3.1 is fulfilled. Otherwise, take  $\hat{t} \in \mathbb{N}$  sufficiently large. Then, by upper-semi-continuity we obtain that  $PB^i(x_{\hat{t}}^{-i}) \subseteq PB^i(x^{-i})$  and  $PSB^i(x_{\hat{t}}^{-i}) \subseteq PSB^i(x^{-i})$  for all  $i \in N$  and all  $t \geq \hat{t}$ .

In the remainder of this proof we make use of the following notation. Let  $i \in \{1,2\}$ . Then, for a given strategy  $x^i \in \Delta_{M^i}$  the vector  $x^i(-k)$  is defined such that for all  $\ell \in M^i$

$$x_\ell^i(-k) = \begin{cases} 0 & \text{if } \ell = k, \\ x_\ell^i & \text{otherwise.} \end{cases}$$

Note that  $x^i(-k)$  is not necessarily a strategy, as the probabilities might not sum up to 1. Moreover, if  $|C(x^i)| = 1$  for some  $i \in \{1,2\}$  we assume in this proof without loss of generality that  $x^i = e_1^i$  and introduce the set  $Q^i(t) = \{\ell \in M^i \setminus \{1\} \mid x_{t,\ell}^i > \varepsilon_t x_{t,r}^i \text{ for all } r \in M^i \setminus \{1\}\}$ . Let the strategy  $\bar{x}^i$  in that case be defined by

$$\bar{x}_\ell^i = \begin{cases} \frac{x_{t,\ell}^i}{\sum_{r \in Q^i(t)} x_{t,r}^i} & \text{for all } \ell \in Q^i(\hat{t}), \\ 0 & \text{otherwise.} \end{cases}$$

Now consider the case that  $|C(x^i)| > 1$  and  $|C(x^j)| = 1$ . Take  $\delta^j > 0$  sufficiently small and define  $\hat{x} = (x^i, \hat{x}^j)$ , with  $\hat{x}^j = (1 - \delta^j)e_1^j + \delta^j \bar{x}^j$ . Then we show that  $C(x^i) \subseteq PB^i(\hat{x}^j)$ ,  $C(x^j) \subseteq PB^j(x^i)$  and  $C(\bar{x}^j) \subseteq PSB^j(x^i)$  as in statement 2 of Proposition 3.1.

We first show that  $C(x^i) \subseteq PB^i(\hat{x}^j)$ . Without loss of generality let  $1 \in C(x^i)$ . Since  $x \in NE(G)$

$$\pi^i(e_1^i, e_1^j) \geq \pi^i(e_k^i, e_1^j) \tag{3}$$

for all  $k \in M^i$ . Furthermore, since the sequence  $\{x_t\}_{t \in \mathbb{N}}$  is  $\varepsilon_t$ -proper and converges to  $x$  it holds that  $1 \in PB^i(x_{\hat{t}}^j)$ , which implies that

$$\pi^i(e_1^i, x_{\hat{t}}^j) \geq \pi^i(e_k^i, x_{\hat{t}}^j)$$

for all  $k \in M^i$ . Since  $x_{\hat{t}}^j$  is sufficiently close to  $e_1^j$  we obtain that

$$\pi^i(e_1^i, x_{\hat{t}}^j(-1)) \geq \pi^i(e_k^i, x_{\hat{t}}^j(-1))$$

for all  $k \in PB^i(e_1^j)$ . Using the fact that  $x_{\hat{t},\ell}^j > \varepsilon_{\hat{t}} x_{\hat{t},r}^j$  for all  $\ell \in Q^j(\hat{t})$ ,  $r \notin Q^j(\hat{t}) \cup \{1\}$  this results in

$$\pi^i(e_1^i, \bar{x}^j) \geq \pi^i(e_k^i, \bar{x}^j) \quad (4)$$

for all  $k \in PB^i(e_1^j)$ . Combining (3) and (4) we find

$$\pi^i(e_1^i, \hat{x}^j) \geq \pi^i(e_k^i, \hat{x}^j)$$

for all  $k \in M^j$ . Hence,  $C(x^i) \subseteq PB^i(\hat{x}^j)$ .

Since  $x \in NE(G)$  we immediately obtain that  $C(x^j) \subseteq PB^j(x^i)$ . It remains to be shown that  $C(\bar{x}^j) \subseteq PSB^j(x^i)$ . Properness of  $x$  implies that  $Q^j(\hat{t}) \subseteq PSB^j(x_{\hat{t}}^i)$  and hence

$$\begin{aligned} C(\bar{x}^j) &= Q^j(\hat{t}) \\ &\subseteq PSB^j(x_{\hat{t}}^i) \\ &\subseteq PSB^j(x^i). \end{aligned}$$

Finally consider the case that  $|C(x^1)| = |C(x^2)| = 1$ . Take  $\delta^i > 0$ ,  $i \in \{1, 2\}$ , sufficiently small and define  $\hat{x} = (\hat{x}^1, \hat{x}^2)$ , with  $\hat{x}^i = (1 - \delta^i)e_1^i + \delta^i \bar{x}^i$  for all  $i \in \{1, 2\}$ . We prove that statement 3 of Proposition 3.1 is satisfied by showing that  $C(x^1) \subseteq PB^1(\hat{x}^2)$  and that  $C(\bar{x}^1) \subseteq PSB^1(\hat{x}^2)$ . Showing that  $C(x^2) \subseteq PB^2(\hat{x}^1)$  and  $C(\bar{x}^2) \subseteq PSB^2(\hat{x}^1)$  can be done analogously. The proof that  $C(x^1) \subseteq PB^1(\hat{x}^2)$  is similar to the proof that  $C(x^i) \subseteq PB^i(\hat{x}^j)$  for the previous case with  $|C(x^i)| > 1$  and  $|C(x^j)| = 1$ . Hence, we only have to show that  $C(\bar{x}^1) \subseteq PSB^1(\hat{x}^2)$ . Assume without loss of generality that  $2 \in C(\bar{x}^1)$ . Since  $C(\bar{x}^1) = Q^1(\hat{t}) \subseteq PSB^1(x_{\hat{t}}^2)$ , it holds that  $2 \in PSB^1(x_{\hat{t}}^2)$ , which implies that

$$\pi^1(e_2^1, x_{\hat{t}}^2) \geq \pi^1(e_k^1, x_{\hat{t}}^2)$$

for all  $k \in M^1 \setminus \{1\}$ . Since  $x_{\hat{t}}^2$  is close to  $e_1^2$

$$\pi^1(e_2^1, x_{\hat{t}}^2(-1)) \geq \pi^1(e_k^1, x_{\hat{t}}^2(-1))$$

for all  $k \in PSB^1(e_1^2) \setminus \{1\}$ . Using the fact that  $x_{\hat{t},\ell}^2 > \varepsilon_{\hat{t}} x_{\hat{t},r}^2$  for all  $\ell \in Q^2(\hat{t})$ ,  $r \notin Q^2(\hat{t}) \cup \{1\}$ , this results in

$$\pi^1(e_2^1, \bar{x}^2) \geq \pi^1(e_k^1, \bar{x}^2) \quad (5)$$



for all  $k \in PSB^1(e_1^2) \setminus \{1\}$ . Furthermore, since  $PSB^1(x_t^2) \subseteq PSB^1(e_1^2)$ , we know that

$$\pi^1(e_2^1, e_1^2) \geq \pi^1(e_k^1, e_1^2) \quad (6)$$

for all  $k \in M^1 \setminus \{1\}$ . As a result of equations (5) and (6)

$$\pi^1(e_2^1, \hat{x}^2) \geq \pi^1(e_k^1, \hat{x}^2)$$

for all  $k \in M^1 \setminus \{1\}$ , which implies that  $C(\bar{x}^1) \subseteq PSB^1(\hat{x}^2)$ .  $\square$

Note that since the set of proper equilibria refines the set of perfect equilibria Theorem 4.1 implies that for all bimatrix games the intersection between the sets of fall back and perfect equilibria is non-empty. The intersection between the sets of fall back and strictly perfect equilibria, however, can be empty. This follows from the example in Vermeulen and Jansen (1996) in which it is shown that not every strictly perfect equilibrium is a proper equilibrium.

The following example shows that for games with three players the set of proper equilibria need not be a subset of the set of fall back equilibria.

**Example 4.2** Consider the following three-player game in which the third player chooses the left ( $e_1^3$ ) or the right ( $e_2^3$ ) matrix:

$$\begin{array}{c} e_1^2 \quad e_2^2 \quad e_3^2 \\ e_1^1 \left[ \begin{array}{ccc} 10, 10, 1 & 5, 5, 1 & 0, 5, 1 \end{array} \right] \quad \left[ \begin{array}{ccc} e_1^2 & e_2^2 & e_3^2 \\ 0, 0, 0 & 0, 0, 0 & 0, 0, 0 \end{array} \right] \\ e_2^1 \left[ \begin{array}{ccc} 10, 0, 1 & 0, 0, 1 & 5, 0, 1 \end{array} \right] \quad \left[ \begin{array}{ccc} 0, 0, 0 & 0, 0, 0 & 0, 10, 0 \end{array} \right] \\ e_3^1 \left[ \begin{array}{ccc} 0, -1, 1 & 5, 10, 1 & 0, 0, 1 \end{array} \right] \quad \left[ \begin{array}{ccc} 0, 0, 0 & 0, 0, 0 & 0, 0, 0 \end{array} \right]. \end{array}$$

The strategy profile  $x = (e_1^1, e_2^2, e_3^3)$  is a proper equilibrium since for a sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$ , with  $\varepsilon_t = \frac{2}{t}$  for all  $t \in \mathbb{N}$ , converging to zero the sequence  $\{x_t\}_{t \in \mathbb{N}}$  converging to  $x$  is  $\varepsilon_t$ -proper for all  $t \in \mathbb{N}$ , with  $x_t$  given by  $x_t^1 = (1 - \frac{1}{t} - \frac{1}{t^2})e_1^1 + \frac{1}{t}e_2^1 + \frac{1}{t^2}e_3^1$ ,  $x_t^2 = (1 - \frac{1}{t} - \frac{1}{t^2})e_1^2 + \frac{1}{t}e_2^2 + \frac{1}{t^2}e_3^2$  and  $x_t^3 = (1 - \frac{1}{10t})e_1^3 + \frac{1}{10t}e_2^3$ . However,  $x$  is not a fall back equilibrium, which can be seen by considering a corresponding blocking game  $G(\delta_t)$ . In such a game player 3 will always play  $x_t^3 = (1 - \delta_t^3)e_1^3 + \delta_t^3 e_2^3$ . Player 1, however, plays his third row with zero probability for any strategy combination close to  $x$ . Knowing this, player 2 always prefers  $e_3^2$  to  $e_2^2$ , due to the payoff of 10 in the second row of the right matrix. As a consequence, player 1 prefers  $e_2^1$  to  $e_1^1$ , which implies that for any sequence of blocking vectors  $\{\delta_t\}_{t \in \mathbb{N}}$  converging to zero there does not exist a sequence  $\{x_t\}_{t \in \mathbb{N}}$  converging to  $x$  such that  $x_t \in NE(G(\delta_t))$  for all  $t \in \mathbb{N}$ . Therefore,  $x$  is not a fall back equilibrium.

One of the fall back equilibria of this game is  $x' = (e_3^1, e_2^2, e_1^3)$ , which requires player 1 to play a weakly dominated strategy. So clearly,  $x'$  is not a proper (or perfect) equilibrium, as in the corresponding thought experiment all strategies are played with strictly positive probability. In any corresponding blocking game however, player 2 plays  $e_1^2$  with zero probability for any strategy profile close to  $x'$ , and consequently player 1 can maximise his profit by playing  $(1 - \delta_t^1)e_3^1 + \delta_t^1 e_1^1$  for all  $t \in \mathbb{N}$ .

In this game there are also some equilibria that are both proper and fall back, like e.g.  $(e_2^1, e_1^2, e_1^3)$ . The question whether in general the intersection between the sets of fall back and proper equilibria can be empty is still open.  $\triangleleft$

We now focus on the relation between fall back and robust equilibrium (cf. Okada (1983)).

**Definition** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. A strategy profile  $\hat{x} \in \Delta$  is called a *robust equilibrium* of  $G$  if for all  $j \in N$  there exists an open neighbourhood  $U^j(\hat{x}^j)$  of  $\hat{x}^j \in \Delta_{M^j}$  such that for all  $i \in N$

$$\pi^i(\hat{x}^i, \check{x}^{-i}) \geq \pi^i(x^i, \check{x}^{-i})$$

for all  $x^i \in \Delta_{M^i}$  and all  $\check{x}^{-i} \in \prod_{r \in N \setminus \{i\}} U^r(\hat{x}^r)$ .

**Theorem 4.3** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. Then, every robust equilibrium of  $G$  is a strict fall back equilibrium of  $G$ .

**Proof:** Let  $\hat{x} \in \Delta$  be a robust equilibrium. Then by definition for all  $j \in N$  there exists an open neighbourhood  $U^j(\hat{x}^j)$  of  $\hat{x}^j \in \Delta_{M^j}$  such that for all  $i \in N$

$$\pi^i(\hat{x}^i, \check{x}^{-i}) \geq \pi^i(x^i, \check{x}^{-i}) \quad (7)$$

for all  $x^i \in \Delta_{M^i}$  and all  $\check{x}^{-i} \in \prod_{r \in N \setminus \{i\}} U^r(\hat{x}^r)$ .

Let  $\{\delta_t\}_{t \in \mathbb{N}}$  be a sequence of blocking vectors of positive real numbers converging to zero, and let for all  $t \in \mathbb{N}$  the blocking game be given by  $G(\delta_t) = (N, \{\Delta_{M^i}(\delta_t)\}_{i \in N}, \{\pi^i\}_{i \in N})$ . Then we construct a sequence  $\{\hat{x}_t\}_{t \in \mathbb{N}}$  converging to  $\hat{x}$  such that  $\hat{x}_t \in NE(G(\delta_t))$  for all  $t \in \mathbb{N}$ . This shows by Theorem 2.3 that  $\hat{x}$  is a fall back equilibrium.

Define  $N^* = \{i \in N \mid |C(\hat{x}^i)| > 1\}$  and  $N' = N \setminus N^*$ . Assume without loss of generality that for each  $i \in N'$ ,  $\hat{x}^i = e_1^i$ . We introduce for all  $t \in \mathbb{N}$  the game  $G_t(\hat{x}) = (N, \{\Delta_{\hat{M}^i}\}_{i \in N}, \{\hat{\pi}_t^i\}_{i \in N})$ . For all  $i \in N^*$ ,  $\hat{M}^i = \{f_1^i\}$  and for all  $i \in N'$ ,  $\hat{M}^i = \{f_2^i, \dots, f_{m^i}^i\}$ . For all  $i \in N'$  the payoff function  $\hat{\pi}_t^i$  is the mixed extension of

$$\hat{\pi}_t^i(f_{k^i}^i, (f_{kj}^j)_{j \in N \setminus \{i\}}) = \pi^i(e_{k^i}^i, (\hat{x}^j)_{j \in N^*}, ((1 - \delta_t^j)e_1^j + \delta_t^j e_{k^j}^j)_{j \in N' \setminus \{i\}})$$

for all  $(f_{k^i}^i)_{i \in N} \in \prod_{j \in N} \hat{M}^j$ ,  $t \in \mathbb{N}$ . Since each  $i \in N^*$  is a dummy player in  $G_t(\hat{x})$ ,  $t \in \mathbb{N}$ , we do not need to specify their payoff functions explicitly. Then, let  $\tilde{x}_t \in NE(G_t(\hat{x}))$  for all  $t \in \mathbb{N}$ . For all  $t \in \mathbb{N}$  and all  $i \in N'$  we define  $\bar{x}_t^i \in \Delta_{M^i}$  to be the extension of  $\tilde{x}_t^i \in \Delta_{\hat{M}^i}$  to  $\Delta_{M^i}$ , in the sense that  $\bar{x}_{t,k}^i = \tilde{x}_{t,k}^i$  for all  $k \in \hat{M}^i$ ,  $\bar{x}_{t,1}^i = 0$ . Further, for all  $t \in \mathbb{N}$  and  $i \in N^*$ ,  $\bar{x}_t^i = \hat{x}^i$ . Next define the sequence  $\{\hat{x}_t\}_{t \in \mathbb{N}}$  such that

$$\hat{x}_t^i = (1 - \delta_t^i)\hat{x}^i + \delta_t^i \bar{x}_t^i$$

for all  $i \in N$  and all  $t \in \mathbb{N}$ . Note that since  $\bar{x}_{t,1}^i = 0$  for all  $i \in N'$  and all  $t \in \mathbb{N}$  we obtain that  $\hat{x}_t \in \Delta(\delta_t)$  for sufficiently large  $t \in \mathbb{N}$ . Also note that since  $\bar{x}_t^i = \hat{x}^i$  for all  $i \in N^*$  and all  $t \in \mathbb{N}$ , we obtain that  $\hat{x}_t^i = \hat{x}^i$  for all  $t \in \mathbb{N}$ .

Hence,  $\hat{x}_t \in \Delta(\delta_t)$  for sufficiently large  $t \in \mathbb{N}$ , and  $\{\hat{x}_t\}_{t \in \mathbb{N}}$  converges to  $\hat{x}$ . Take  $\hat{t} \in \mathbb{N}$  such

that for all  $i \in N$ ,  $\hat{x}_t^i \in \Delta_{M^i}(\delta_t^i) \cap U^i(\hat{x}^i)$  for all  $t \geq \hat{t}$ . Then, we complete the proof by showing that  $\hat{x}_t^i$  is a best reply against  $\hat{x}_t^{-i}$  in  $G(\delta_t)$  for all  $i \in N$  and for all  $t \geq \hat{t}$ . Let  $i \in N$  and  $t \geq \hat{t}$ . First of all, from (7) it follows that

$$\pi^i(\hat{x}_t^i, \hat{x}_t^{-i}) \geq \pi^i(x_t^i, \hat{x}_t^{-i}) \quad (8)$$

for all  $x_t^i \in \Delta_{M^i}$ . If  $i \in N^*$ , then  $\hat{x}_t^i = \hat{x}^i \in \Delta_{M^i}(\delta_t)$  and  $\hat{x}_t^i$  is a best reply against  $\hat{x}_t^{-i}$  in  $G(\delta_t)$ . So, assume  $i \in N'$ . Then it remains to be shown that  $\bar{x}_t^i \in PSB^i(\hat{x}_t^{-i})$ . Since  $\tilde{x}_t \in NE(G_t(\hat{x}))$ ,

$$\hat{\pi}_t^i(\tilde{x}_t^i, \tilde{x}_t^{-i}) \geq \hat{\pi}_t^i(\hat{x}_t^i, \tilde{x}_t^{-i})$$

for all  $\hat{x}_t^i \in \Delta_{M^i}$ . As a result, we obtain by the definition of  $\hat{\pi}_t^i$  that

$$\pi^i(\bar{x}_t^i, \hat{x}_t^{-i}) \geq \pi^i(x_t^i, \hat{x}_t^{-i}) \quad (9)$$

for all  $x_t^i \in \Delta_{M^i \setminus \{1\}}$ . Hence,  $\bar{x}_t^i \in PSB^i(\hat{x}_t^{-i})$ . Combining (8) and (9) results in

$$\pi^i(\hat{x}_t^i, \hat{x}_t^{-i}) \geq \pi^i(x_t^i, \hat{x}_t^{-i})$$

for all  $x_t^i \in \Delta_{M^i}(\delta_t^i)$ , which implies that  $\hat{x}_t^i$  is a best reply against  $\hat{x}_t^{-i}$  in  $G(\delta_t)$ .

Since the sequence  $\{\delta_t\}_{t \in \mathbb{N}}$  was arbitrary chosen this implies that each robust equilibrium is a strict fall back equilibrium.  $\square$

## 5 Structure of the set of fall back equilibria

For bimatrix games the set of Nash equilibria is the union of finitely many polytopes (cf. Jansen (1981)). The main result provided in this section is that this is also true for the set of fall back equilibria. In order to obtain this result we need several preliminary lemmas.

Let us introduce some notation. For a set  $A$  we denote by  $\text{cl}(A)$  the closure of  $A$  and by  $\text{relint}(A)$  its relative interior. Further, for a bimatrix game  $G = (N, \{\Delta_{M^i}\}_{i \in \{1,2\}}, \{\pi^i\}_{i \in \{1,2\}})$  the strategies  $x^1, \tilde{x}^1 \in \Delta_{M^1}$  are *reply-equivalent* if the following two statements hold:

$$\begin{aligned} PB^2(x^1) &= PB^2(\tilde{x}^1), \\ PSB^2(x^1) &= PSB^2(\tilde{x}^1). \end{aligned}$$

By  $\mathcal{V}_1, \dots, \mathcal{V}_{r_1}$  we denote the finitely many reply-equivalence classes in  $\Delta_{M^1}$ . In a similar way a reply-equivalence relation can be defined for the strategies of player 2. The reply-equivalence classes in  $\Delta_{M^2}$  are denoted by  $\mathcal{W}_1, \dots, \mathcal{W}_{r_2}$ . Note that since the sets of pure best and pure second best replies are determined by linear inequalities, the closure of each reply-equivalence class is a polytope.

By the use of Jansen (1993) we obtain the following two lemmas.

**Lemma 5.1** *Let  $H$  be a face of  $\text{cl}(\mathcal{V}_s)$ ,  $s \in \{1, \dots, r^1\}$ . Then all the elements in  $\text{relint}(H)$  are reply-equivalent.*

**Lemma 5.2** *If the intersection of the closure of two reply-equivalence classes is non-empty, then this intersection is a face of both polytopes.*

Given a bimatrix game  $G = (N, \{\Delta_{M^i}\}_{i \in \{1,2\}}, \{\pi^i\}_{i \in \{1,2\}})$  and reply-equivalence classes  $\mathcal{V}_s$ ,  $s \in \{1, \dots, r^1\}$ , and  $\mathcal{W}_t$ ,  $t \in \{1, \dots, r^2\}$ , the  $(s, t)$ -fall back component is defined by  $FB_{st}(G) = \{x \in FBE(G) \mid x^1 \in \text{cl}(\mathcal{V}_s), x^2 \in \text{cl}(\mathcal{W}_t)\}$ .

**Lemma 5.3** *For every  $s \in \{1, \dots, r^1\}$ ,  $t \in \{1, \dots, r^2\}$ ,  $FB_{st}(G)$  is the cartesian product of two polytopes, which are faces of  $\text{cl}(\mathcal{V}_s)$  and  $\text{cl}(\mathcal{W}_t)$ , respectively.*

**Proof:** Consider the reply-equivalence class  $\mathcal{V}_s$ ,  $s \in \{1, \dots, r^1\}$ , and let  $H$  be a face of  $\text{cl}(\mathcal{V}_s)$ . Take  $x^1 \in \text{relint}(H)$  and  $x^2 \in \text{cl}(\mathcal{W}_t)$ ,  $t \in \{1, \dots, r^2\}$ . We show that whenever  $(x^1, x^2) \in FB_{st}(G)$  it holds that  $(\tilde{x}^1, x^2) \in FB_{st}(G)$  for all  $\tilde{x}^1 \in \text{relint}(H)$ . The fact that the set of fall back equilibria is closed then shows that  $(\tilde{x}^1, x^2) \in FB_{st}(G)$  for all  $\tilde{x}^1 \in H$ , which completes the proof.

Let  $(x^1, x^2) \in FB_{st}(G)$ . If  $|C(x^1)| = 1$ , then  $|H| = 1$  and the statement follows immediately. So, assume  $|C(x^1)| > 1$ . We distinguish between two cases. We first assume that  $|C(x^2)| > 1$ . Then by Proposition 2.4  $x \in NE(G)$ , which implies that  $C(x^1) \subseteq PB^1(x^2)$  and  $C(x^2) \subseteq PB^2(x^1)$ . Let  $\tilde{x}^1 \in \text{relint}(H)$ . Then  $C(\tilde{x}^1) = C(x^1)$ , and hence  $C(\tilde{x}^1) \subseteq PB^1(x^2)$ . Furthermore, by Lemma 5.1 we obtain  $PB^2(\tilde{x}^1) = PB^2(x^1)$  and therefore,  $C(x^2) \subseteq PB^2(\tilde{x}^1)$ . Consequently,  $(\tilde{x}^1, x^2) \in FB_{st}(G)$ .

Next we assume that  $|C(x^2)| = 1$ . Let  $\bar{x}^2 \in \Delta_{M^2}$  with  $C(\bar{x}^2) \cap C(x^2) = \emptyset$  and a  $\bar{\delta}^2 > 0$  be such that for all  $\delta^2 \in (0, \bar{\delta}^2]$  the strategy profile  $(x^1, \hat{x}^2) \in \Delta$ , with  $\hat{x}^2 = (1 - \delta^2)x^2 + \delta^2\bar{x}^2$  satisfies  $C(x^1) \subseteq PB^1(\hat{x}^2)$ ,  $C(x^2) \subseteq PB^2(x^1)$  and  $C(\bar{x}^2) \subseteq PSB^2(x^1)$ .

Note that by Proposition 3.1 this is possible. We show that these conditions are also satisfied for  $(\tilde{x}^1, x^2)$ . Since  $x^1, \tilde{x}^1 \in \text{relint}(H)$ , by Lemma 5.1 we conclude  $PB^2(x^1) = PB^2(\tilde{x}^1)$  and  $PSB^2(x^1) = PSB^2(\tilde{x}^1)$ , and furthermore  $C(x^1) = C(\tilde{x}^1)$ . Consequently,  $C(\tilde{x}^1) = C(x^1) \subseteq PB^1(\hat{x}^2)$ ,  $C(x^2) \subseteq PB^2(x^1) = PB^2(\tilde{x}^1)$  and  $C(\bar{x}^2) \subseteq PSB^2(x^1) = PSB^2(\tilde{x}^1)$ .

Hence, by Proposition 3.1 we obtain that  $(\tilde{x}^1, x^2) \in FBE(G)$ , and as a consequence  $(\tilde{x}^1, x^2) \in FB_{st}(G)$ .  $\square$

Since there are only finitely many combinations of reply-equivalence classes we obtain the following theorem.

**Theorem 5.4** *Let  $G$  be a bimatrix game. Then the set of fall back equilibria of  $G$  is the union of finitely many polytopes.*

A *maximal fall back component* is a fall back component not properly contained in another fall back component. From the first example of Section 1 it follows that a maximal Nash subset may contain more than one maximal fall back component and that a maximal fall back component need not be the face of a maximal Nash subset. The latter result in particular implies that an extreme element of a maximal fall back component need not be an extreme element of a maximal Nash subset. Furthermore, Lemma 5.2 and 5.3 imply that the intersection of two maximal fall back components is either empty or a face of both maximal fall back components.

## 6 Complete fall back equilibrium

In this final section we discuss a modification of the concept of fall back equilibrium. In the thought experiment underlying fall back equilibrium each player faces the possibility that, after all players decided on their action, the action chosen by him is blocked. In that case the player plays a back-up action, chosen by him beforehand. Moreover, we assumed that a back-up action is never blocked.

In this section we consider the possibility that any number of actions of each player is blocked. Consequently, players have to decide beforehand on a second back-up action in case the first back-up action is blocked and a third back-up action in case the second back-up cannot be played either, etc. Hence, each player must decide on a complete ordering of his actions. If all actions of a player turn out to be blocked the game is not played and all players receive zero payoff. This thought experiment is modelled by a corresponding complete fall back game. The equilibrium concept that is based on this thought experiment is called complete fall back equilibrium.

Note that not playing the game is not an option a player can choose, but that this can only be the result of a player not being able to play any of his actions. Therefore, the zero payoff to each player if this situation occurs is arbitrary, as any fixed amount would result in the same set of equilibria. In order to avoid the possibility that the game is not played we could also have chosen for a setup in which at most all but one actions of each player are blocked. This setup leads to a similar, but different equilibrium concept, which we will not discuss.

Let us formalise the concept introduced above. For this we first recall the notation of the second section. A non-cooperative game in strategic form is given by  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$ , with  $N = \{1, \dots, n\}$  the player set,  $\Delta_{M^i}$  the mixed strategy space of player  $i \in N$ , with  $M^i = \{1, \dots, m^i\}$  the set of pure strategies, and  $\pi^i : \prod_{j \in N} \Delta_{M^j} \rightarrow \mathbb{R}$  the von Neumann Morgenstern expected payoff function of player  $i$ . A pure strategy  $k \in M^i$  of player  $i$  is alternatively denoted by  $e_k^i$ , a typical element of  $\Delta_{M^i}$  by  $x^i$ . We denote the probability which  $x^i$  assigns to pure strategy  $k$  by  $x_k^i$ . The set of all strategy profiles is given by  $\Delta = \prod_{i \in N} \Delta_{M^i}$ , a typical element of  $\Delta$  by  $x$ .

The action set in the complete fall back game for player  $i$  equals the set of all orderings on the action set  $M^i$ , and is given by  $\Omega^i$ . Hence, the total number of actions in the

complete fall back game for player  $i$  equals  $\tilde{m}^i = m^i!$ . A typical element of  $\Omega^i$  is denoted by  $\sigma^i$ , where the action on position  $s$  of  $\sigma^i$  is given by  $\sigma^i(s) \in M^i$ . By  $\Omega_k^i \subseteq \Omega^i$ ,  $k \in M^i$ , we denote the set of orders of  $M^i$  such that for all  $\sigma^i \in \Omega_k^i$  it holds that  $\sigma^i(1) = k$ . Similar to the concept of fall back equilibrium we assume that each action of player  $i$  is blocked with the same probability, denoted by  $\varepsilon^i$ , but we allow for different probabilities among players. Hence, let  $\varepsilon = (\varepsilon^1, \dots, \varepsilon^n)$  be an  $n$ -tuple of (small) non-negative probabilities. If player  $i$  plays action  $\sigma^i \in \Omega^i$  in the complete fall back game he plays with probability  $(1 - \varepsilon^i)(\varepsilon^i)^{s-1}$  action  $\sigma^i(s)$  of the game  $G$  for  $s \in \{1, \dots, m^i\}$ . With probability  $(\varepsilon^i)^{m^i}$  all actions of player  $i$  are blocked and the payoff to all players is defined to be zero.

The complete fall back game  $\tilde{G}^C(\varepsilon)$  is given by  $\tilde{G}^C(\varepsilon) = (N, \{\Delta_{\Omega^i}\}_{i \in N}, \{\pi_\varepsilon^i\}_{i \in N})$ , with  $\pi_\varepsilon^i : \prod_{j \in N} \Delta_{\Omega^j} \rightarrow \mathbb{R}$  the extended expected payoff function to player  $i$ . A pure strategy  $\sigma^i \in \Omega^i$  will be alternatively denoted by  $e_\sigma^i$ . Then,  $\pi_\varepsilon^i$  is formally given by

$$\pi_\varepsilon^i((e_{\sigma^j}^j)_{j \in N}) = \sum_{(k^1, \dots, k^n) \in \prod_{r \in N} M^r} \left( \prod_{j \in N} (1 - \varepsilon^j)(\varepsilon^j)^{\sigma^j(k^j)-1} \right) \pi^i((e_{k^j}^j)_{j \in N}).$$

A typical element of  $\Delta_{\Omega^i}$  will be denoted by  $\rho^i$ , the probability which  $\rho^i$  assigns to pure strategy  $\sigma^i$  is given by  $\rho_\sigma^i$ . The set of all strategy profiles is given by  $\tilde{\Delta}^C = \prod_{i \in N} \Delta_{\Omega^i}$ , an element of  $\tilde{\Delta}^C$  by  $\rho$ .

**Definition** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game. A strategy profile  $x \in \Delta$  is called a *complete fall back equilibrium* of  $G$  if there exists a sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  of  $n$ -tuples of positive real numbers converging to zero, and a sequence  $\{\rho_t\}_{t \in \mathbb{N}}$  such that  $\rho_t \in NE(\tilde{G}^C(\varepsilon_t))$  for all  $t \in \mathbb{N}$ , converging to  $\rho \in \tilde{\Delta}^R$ , with  $x_k^i = \sum_{\sigma^i \in \Omega_k^i} \rho_\sigma^i$  for all  $k \in M^i$  and all  $i \in N$ . The set of complete fall back equilibria of a game  $G$  is denoted by  $CFBE(G)$ .

**Theorem 6.1** Let  $G$  be an  $n$ -player strategic game. Then  $CFBE(G)$  is a non-empty and closed subset of  $NE(G)$ .

The proof of this theorem is analogous to the proof of Theorem 2.1 for fall back equilibria.

As the thought experiment underlying the concept of complete fall back equilibrium takes into account all levels of best replies and the thought experiment underlying fall back equilibrium only the first and second, one might expect that the set of complete fall back equilibria refines the set of fall back equilibria. This is however not the case, as can be seen in Example 4.2. In this example the strategy profile  $(e_1^1, e_1^2, e_1^3)$ , which is not a fall back equilibrium, is a complete fall back equilibrium. This is due to the fact that each action in the complete fall back game puts positive probability on all actions of the original game  $G$ . Therefore, player 1 is unable to play  $e_1^3$  with zero probability, which is the main reason for  $(e_1^1, e_1^2, e_1^3)$  not being a fall back equilibrium.

The setup of complete fall back equilibrium is closely related to that of proper equilibrium.

In both concepts the replies of each player are ordered in such a way that a complete set of levels of best replies is obtained. The properness concept then requires that replies of a lower level are played with *some* significant smaller probability than replies from a higher level. The concept of complete fall back equilibrium is however more restrictive, as for each action in the complete fall back game the probability on the actions of the original game are given. Hence, by requiring that players play a best reply in the complete fall back game the probability on each best reply level of the original game is fixed. This is the reason why the set of complete fall back equilibria refines the set of proper equilibria.

**Theorem 6.2** Let  $G$  be an  $n$ -player strategic game. Then, each complete fall back equilibrium of  $G$  is a proper equilibrium of  $G$ .

**Proof:** Let  $G = (N, \{\Delta_{M^i}\}_{i \in N}, \{\pi^i\}_{i \in N})$  be an  $n$ -player strategic game and let  $x \in CFBE(G)$ . Then by definition there exists a sequence  $\{\varepsilon_t\}_{t \in \mathbb{N}}$  of  $n$ -tuples of positive real numbers converging to zero, and a sequence  $\{\rho_t\}_{t \in \mathbb{N}}$  such that  $\rho_t \in NE(\tilde{G}^C(\varepsilon_t))$  for all  $t \in \mathbb{N}$ , converging to  $\rho \in \tilde{\Delta}^C$ , with  $x_k^i = \sum_{\sigma^i \in \Omega_k^i} \rho_{\sigma^i}^i$  for all  $k \in M^i$  and all  $i \in N$ .

We define the sequence  $\{x_t\}_{t \in \mathbb{N}}$  such that for all  $t \in \mathbb{N}$

$$x_{t,k}^i = \sum_{\sigma^i \in \Omega_k^i} (1 - \varepsilon_t^i)(\varepsilon_t^i)^{\sigma^i(k)-1-1} \rho_{t,\sigma}^i$$

for all  $k \in M^i$  and all  $i \in N$ . Note that  $x_t^i$  puts the same probability on the actions of the game  $G$  as  $\rho_t^i$  for all  $i \in N$  and all  $t \in \mathbb{N}$ , and that the sequence  $\{x_t\}_{t \in \mathbb{N}}$  converges to  $x$ . Let the sequence  $\{\hat{\varepsilon}_t\}_{t \in \mathbb{N}}$  be given by  $\hat{\varepsilon}_t = \max_{i \in N} \varepsilon_t^i$  for all  $t \in \mathbb{N}$ .

Let  $i \in N$  and let  $\pi^i(e_k^i, x_t^{-i}) < \pi^i(e_\ell^i, x_t^{-i})$  for some  $k, \ell \in M^i$  and some  $t \in \mathbb{N}$ . Since  $\rho_t \in NE(G^C(\varepsilon_t))$  for all  $t \in \mathbb{N}$  it holds that  $x_{t,k}^i \leq \varepsilon_t^i x_{t,\ell}^i$ . Hence,

$$\begin{aligned} x_{t,k}^i &\leq \varepsilon_t^i x_{t,\ell}^i \\ &\leq \max_{i \in N} \varepsilon_t^i x_{t,\ell}^i \\ &= \hat{\varepsilon}_t x_{t,\ell}^i. \end{aligned}$$

Consequently,  $\{\hat{\varepsilon}_t\}_{t \in \mathbb{N}}$  is a sequence of positive real numbers converging to zero and  $\{x_t\}_{t \in \mathbb{N}}$  is a sequence of completely mixed strategy profiles converging to  $x$  such that

$$\pi^i(e_k^i, x_t^{-i}) < \pi^i(e_\ell^i, x_t^{-i}) \Rightarrow x_{t,k}^i \leq \hat{\varepsilon}_t x_{t,\ell}^i$$

for all  $k, \ell \in M^i$  and all  $i \in N$ . Hence,  $x$  is a proper equilibrium.  $\square$

The following example shows that the set of complete fall back equilibria can be a strict subset of the set of proper equilibria.

**Example 6.3** Consider the following three-player game in which the third player chooses the left ( $e_1^3$ ) or the right ( $e_3^3$ ) matrix:

$$\begin{matrix}
e_1^1 \\
e_2^1 \\
e_3^1
\end{matrix}
\begin{bmatrix}
e_1^2 & e_2^2 & e_3^2 \\
10, 10, 10 & 0, 10, 0 & 0, 0, 1 \\
10, 1, 0 & 2, 0, 0 & 0, 0, 0 \\
0, 0, 0 & 0, 0, 0 & 0, 0, 0
\end{bmatrix}
\begin{bmatrix}
e_1^2 & e_2^2 & e_3^2 \\
1, 0, 10 & 0, 1, 0 & 0, 0, 0 \\
0, 0, 0 & 0, 0, 0 & 0, 0, 0 \\
0, 0, 5 & 0, 0, 0 & 0, 0, 0
\end{bmatrix}.$$

The strategy profile  $x = (e_1^1, e_1^2, e_1^3)$  is a proper equilibrium, but not a complete fall back equilibrium.  $\triangleleft$

## Acknowledgements

The authors would like to thank Marieke Quant and Dries Vermeulen for their helpful comments.

Ruud Hendrickx acknowledges financial support from the Netherlands Organisation for Scientific Research (NWO).

## References

- Borm, P. (1992). On perfectness concepts for bimatrix games. *OR Spektrum*, **14**, 33–42.
- Damme, E. van (1991). *Stability and perfection of Nash Equilibria*. Berlin: Springer Verlag.
- Jansen, M. (1981). Maximal Nash subsets for bimatrix games. *Naval Research Quarterly*, **28**, 147–152.
- Jansen, M. (1993). On the set of proper equilibria of a bimatrix game. *International Journal of Game Theory*, **22**, 97–106.
- Kohlberg, E. and J. Mertens (1986). On strategic stability of equilibria. *Econometrica*, **54**, 1003–1037.
- Kuhn, H. (1953). Extensive games and the problem of information. *Annual Mathematics Studies*, **28**, 193–216.
- Myerson, R. (1978). Refinements of the Nash equilibrium point concept. *International Journal of Game Theory*, **7**, 73–80.
- Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, **54**, 286–295.
- Okada, A. (1983). Robustness of equilibrium points in strategic games. Discussion Paper B-137, department of systems science, Tokyo Institute of Technology, Tokyo, Japan.
- Okada, A. (1984). Strictly perfect equilibrium points of bimatrix games. *International Journal of Game Theory*, **13**, 145–154.
- Selten, R. (1975). Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, **4**, 25–55.
- Vermeulen, A. and M. Jansen (1996). Are strictly perfect equilibria proper? A counterexample. *Journal of Optimization Theory and Applications*, **90**, 225–230.