The equivalence of the Dekel–Fudenberg iterative procedure and weakly perfect rationalizability*

P. Jean-Jacques Herings[1] and Vincent J. Vannetelbosch[2,3]

[1] Department of Economics, University of Maastricht, Postbus 616, NL-6200 Maastricht,
THE NETHERLANDS (e-mail: P.Herings@algec.unimaas.nl)
[2] CORE, University of Louvain, voie du Roman Pays 34, B-1348 Louvain-la-Neuve, BELGIUM
[3] Institute of Public Economics, Universidad del Pais Vasco, Avda. Lehendakari Aguirre 83,
E-48015 Bilbao, SPAIN (e-mail: eepvaxxv@bs.ehu.es)

Summary. Two approaches have been proposed in the literature to refine the rationalizability solution concept: either assuming that a player believes that with small probability her opponents choose strategies that are irrational, or assuming that their is a small amount of payoff uncertainty. We show that both approaches lead to the same refinement if strategy perturbations are made according to the concept of weakly perfect rationalizability, and if there is payoff uncertainty as in Dekel and Fudenberg [*J. of Econ. Theory* 52 (1990), 243-267]. For both cases, the strategies that survive are obtained by starting with one round of elimination of weakly dominated strategies followed by many rounds of elimination of strictly dominated strategies.

Keywords and Phrases: Rationalizability, Refinements.

JEL Classification Number: C72.

1 Introduction

The solution concept of rationalizability has been introduced independently by Bernheim [2] and Pearce [10]. In some games it fails to eliminate all intuitively unreasonable outcomes, for instance in games with weakly dominated strategies

---

(see e.g. Pearce [10], Herings and Vannetelbosch [9]). Therefore, one has looked for refinements that strengthen the rationalizability concept but still do not assume common expectations of the behavior of the players. Two different approaches have mainly been investigated in the literature. Both approaches propose to select outcomes which are robust to the introduction of small perturbations. One approach is to assume that a player believes that with small probability her opponents choose irrational strategies, or, similarly, her opponents make errors, while the other one consists of assuming a small amount of payoff uncertainty.

The approach of strategy perturbations has produced many refinements like perfect rationalizability (see Bernheim [2]), cautious rationalizability (see Pearce [10]), proper rationalizability (see Schuhmacher [11]), weakly perfect rationalizability (see Herings and Vannetelbosch [9]), and trembling-hand perfect rationalizability (see Herings and Vannetelbosch [9]). An analysis as well as a detailed description of these refinements can be found in Herings and Vannetelbosch [9]. Recently, a general framework for studying refinements of rationalizability has been introduced in Gul [7], who develops a solution concept called τ-theory. In this theory it is modelled in a coherent way that players may behave irrationally with a small probability, which is related to the assumption that players make mistakes with a small probability.

The approach of payoff perturbations, i.e. the assumption of a small amount of payoff uncertainty, has been studied by Dekel and Fudenberg [6], who obtained the following substantial result. Under the assumption that there is a little bit of uncertainty about the payoffs, rationalizability is equivalent to one round of deletion of weakly dominated strategies, followed by iterated deletion of strategies that are strictly dominated. In what follows, this rule for deleting dominated strategies will be referred to as the Dekel–Fudenberg iterative procedure (DF procedure).

There are also other approaches that lead to the DF procedure. Börgers [3] has shown that if it is approximate common knowledge that players maximize expected utility using full support conjectures, then the players choose strategies that correspond to the DF procedure. Brandenburger [4] has obtained a similar result to Börgers [3]. But, instead of approximate common knowledge, Brandenburger uses a lexicographic analogue, called common first-order knowledge. Gul [7] shows that the DF procedure is the weakest perfect τ-theory. For the class of generic extensive-form games with perfect information, Ben-Porath [1] shows that the set of outcomes that are consistent with common certainty of rationality at the beginning of the game coincides with the set of outcomes that survive the DF procedure.

These results suggest that the DF procedure is a well motivated strengthening of rationalizability. To reinforce this claim, we show that the DF procedure receives also support from the most common approach to refine rationalizability, namely by assuming that players believe that with a small probability errors are made by their opponents. In this paper we show that the concept of weakly perfect rationalizability coincides with the DF procedure. For such an equivalence result to hold, however, it is necessary that players believe that their opponents might

choose irrational strategies in a correlated way. We provide a counterexample to equivalence if players believe that their opponents make uncorrelated errors instead.

2 Definitions and notations

We consider a normal-form game $\Gamma(I, S, U)$. The set I is a finite set of players. Each player i has a finite pure-strategy set S_i and a payoff function $U_i : S \to \mathbb{R}$, where $S = \prod_{i \in I} S_i$ and $U = (U_i)_{i \in I}$.

As general notation, we denote by $\Delta(X)$ the set of all Borel probability measures on X. For finite X, we denote by $\Delta^0(X)$ the set of all Borel probability measures giving positive probability to each member of X.

Given $c_i \in \Delta(S_i)$, we denote by $c_i(s_i)$ the probability that c_i assigns to pure strategy s_i. Player i's opponents in the game $\Gamma(I, S, U)$ are denoted by $-i$.

Given a product set T, which is the Cartesian product of individual strategy sets, T_i denotes the strategy set of player i. The Cartesian product $\prod_{j \neq i} T_j$ is denoted by T_{-i}. For $c_{-i} \in \Delta(S_{-i})$, $c_{-i}(s_{-i})$ denotes the probability that c_{-i} assigns to the pure strategy profile s_{-i}.

2.1 The Dekel–Fudenberg iterative procedure

To define the DF procedure we need to define the notions of strict and weak dominance first.

Definition 1 (strict dominance) Let a product set $T \subseteq S$ of pure strategy profiles in the game $\Gamma(I, S, U)$ be given. A pure strategy $s_i \in T_i$ of player i is strictly dominated in T if there exists $c_i \in \Delta(T_i)$ such that $U_i\left(c_i, s_{-i}\right) > U_i\left(s_i, s_{-i}\right)$ for all $s_{-i} \in T_{-i}$.

Given a product set T of pure strategy profiles, the pure strategies of player i that are not strictly dominated in T are denoted by $B_i(T)$. The pure strategy profiles that are not strictly dominated are denoted by $B(T) = \prod_{i \in I} B_i(T)$.

Definition 2 (weak dominance) Let a product set $T \subseteq S$ of pure strategy profiles in the game $\Gamma(I, S, U)$ be given. A pure strategy $s_i \in T_i$ of player i is weakly dominated in T if there exists $c_i \in \Delta(T_i)$ such that $U_i\left(c_i, s_{-i}\right) \geq U_i\left(s_i, s_{-i}\right)$ for all $s_{-i} \in T_{-i}$, and $U_i\left(c_i, s_{-i}\right) > U_i\left(s_i, s_{-i}\right)$ for some $s_{-i} \in T_{-i}$.

Given a product set T of pure strategy profiles, the pure strategies of player i that are not weakly dominated in T are denoted by $W_i(T)$. The pure strategy profiles that are not weakly dominated are denoted $W(T) = \prod_{i \in I} W_i(T)$.

The DF procedure for removing dominated strategies consists of one round of deletion of weakly dominated strategies, followed by an arbitrarily large number of rounds of deletion of strictly dominated strategies. This procedure can be motivated by assuming small payoff uncertainty, see Dekel and Fudenberg [6],

who give the following intuition for this result: "Each player knows his/her own payoffs, and so by our rationality postulate will not choose a weakly dominated strategy. In order to do a second round of deletion players must know that all the others will not choose certain strategies. A small amount of payoff uncertainty cannot alter strong dominance relationships, but can break weak ones, so that after the first round we can only proceed with the iterated deletion of strongly dominated strategies" (Dekel and Fudenberg [6, p.245]).

Definition 3 (DF procedure) Let $P^1 = W(S)$. For $k \geq 2$, $P^k = B\left(P^{k-1}\right)$. The set $P^\infty = \lim_{k \to \infty} P^k$ is the set of pure strategy profiles generated by the DF procedure.

Evidently, $\emptyset \neq P^k \subseteq P^{k-1} \subseteq \ldots \subseteq P^1$. Since the set S_i is finite for each player i, there exists some integer n such that $P^k = P^n$ for all $k \geq n$. Therefore, the limit set P^∞ is well-defined and non-empty.

2.2 Weakly perfect rationalizability

Weakly perfect rationalizability has been introduced by Herings and Vannetelbosch [9]. Here, we adapt our original definition such that it allows the players to hold correlated conjectures. Correlated weakly perfect rationalizability weakens weakly perfect rationalizability because allowing correlated conjectures about the strategies of the opponents makes more strategies rationalizable. Correlated strategies or conjectures appear to make more sense in the context of the non-equilibrium approach than in the equilibrium approach (see e.g. Brandenburger and Dekel [5] or Hammond [8]). The motivation for weakly perfect rationalizability is that a player believes that the others may choose strategies that are irrational. The probability that an opponent chooses such a strategy is small, and subject to an explicit constraint. A special case is that there are rational expectations that other players actually make errors.

For our results it is crucial whether a player believes that her opponents choose irrational strategies in a correlated way or not. To obtain the equivalence with the DF procedure, correlation is essential; we give a counterexample in Section 4. A clean example to obtain this correlation is the case where player i conjectures that her opponents make correlated mistakes. Such conjectures are rational if, for instance, her opponents implement a correlated strategy $c_{-i} \in \Delta(S_{-i})$ by means of a mediator. The mediator randomly selects a pure strategy profile $s_{-i} \in S_{-i}$ with probability $c_{-i}(s_{-i})$. Then the mediator recommends a player j, $j \neq i$, confidentially to use strategy s_j if s_{-i} is the pure strategy profile selected. If the mediator makes errors and chooses with positive probability not exceeding ε any pure strategy profile $\bar{s}_{-i} \in S_{-i}$ by mistake, then this leads the opponents of player i to make correlated mistakes, even if the mediator makes no errors when making recommendations and the players make no errors in playing the recommended strategy. The example above is just one way to generate the kind of correlation that we need for our equivalence result. In the context of

rationalizability, there is no need to assume either that players make mistakes or that there is a correlation device.

Weakly perfect rationalizability will be defined as an iterative procedure. At any stage k, player i has a set $S_i^k(\varepsilon) \subset S_i$ of strategies that are still rational for her to play at stage k. Here $\varepsilon > 0$ is related to the probability that a player conjectures that her opponents choose an irrational strategy[1]. A conjecture of a player is a Borel probability measure over mixed strategy profiles in $\Delta(\prod_{j \neq i} S_j^k(\varepsilon))$, that are subject to correlated errors. A strategy is rational for player i at stage $k + 1$, and belongs to $S_i^{k+1}(\varepsilon)$, if it is a best response against such conjecture. Since such a conjecture is quite hard to work with, we show that for the purposes of expected utility maximization, we can replace it by an element of the set $\Delta^\varepsilon(\prod_{j \neq i} S_j^k(\varepsilon))$, a subset of $\Delta(S_{-i})$ to be defined more precisely below.

Suppose player i beliefs that her opponents make correlated errors with positive probability not exceeding $\varepsilon > 0$.[2] Let e_{-i} be a measure on S_{-i} describing the probability by which irrational strategies are believed to be chosen. For $s_{-i} \in S_{-i}$, $e_{-i}(s_{-i})$ is the probability player i believes that her opponents irrationally choose the pure strategy profile s_{-i}. It holds that $0 < e_{-i}(s_{-i}) \leq \varepsilon$, $s_{-i} \in S_{-i}$. Now, if player i conjectures that her opponents coordinate on a correlated mixed strategy profile $c_{-i} \in \Delta(S_{-i})$ and expects them to make errors according to e_{-i}, then player i should optimize against the probability measure $\bar{c}_{-i} \in \Delta^0(S_{-i})$ satisfying

$$\bar{c}_{-i}(\bar{s}_{-i}) = \left(1 - \sum_{s_{-i} \in S_{-i}} e_{-i}(s_{-i}) \right) c_{-i}(\bar{s}_{-i}) + e_{-i}(\bar{s}_{-i}), \quad \bar{s}_{-i} \in S_{-i}.$$

We show that this implies that for the purposes of expected utility maximization, player i should maximize expected utility against an element of

$$\Delta^\varepsilon(\prod_{j \neq i} S_j^k(\varepsilon)) = \{ c_{-i} \in \Delta^0(S_{-i}) \mid c_{-i}(s_{-i}) \leq \varepsilon \text{ if } s_j \notin S_j^k(\varepsilon) \text{ for some } j \neq i \}.$$

The set $\Delta^\varepsilon(\prod_{j \neq i} S_j^k(\varepsilon))$ contains all correlated, completely mixed strategy profiles that put weight less than or equal to ε on any pure strategy profile containing a pure strategy not in $S_j^k(\varepsilon)$ for some player j.

Clearly, any probability measure on beliefs of player i on correlated strategy profiles subject to correlated errors, will never assign weight exceeding ε to a pure strategy profile $s_{-i} \in S_{-i}$ such that $s_j \notin S_j^k(\varepsilon)$ for some j. Moreover, any strategy profile $\bar{c}_{-i} \in \Delta^\varepsilon(\prod_{j \neq i} S_j^k(\varepsilon))$ can indeed be conjectured. Define

---

[1] This mistake technology is different from the one used in the perfect rationalizability concept due to Bernheim [2], where each player believes that her opponents choose each of their pure strategies with a certain strictly positive minimum probability. It is shown in Herings and Vannetelbosch [9] that for the case of uncorrelated beliefs perfect rationalizability is a refinement of weakly perfect rationalizability.

[2] It is always assumed that ε is a positive number smaller than the reciprocal of the total number of pure strategy profiles in S.

$\overline{S}_{-i} = \{s_{-i} \in S_{-i} \mid \overline{c}_{-i}(s_{-i}) > \varepsilon\}$. Then, $\overline{c}_{-i}$ results, for instance, from a conjecture $c_{-i} \in \Delta(\prod_{j \neq i} S_j^k(\varepsilon))$ subject to an error e_{-i} as defined below.

$$c_{-i}(\overline{s}_{-i}) = \frac{\overline{c}_{-i}(\overline{s}_{-i})}{\sum_{s_{-i} \in \overline{s}_{-i}} \overline{c}_{-i}(s_{-i})} \qquad \text{if } \overline{s}_{-i} \in \overline{S}_{-i},$$

$$c_{-i}(\overline{s}_{-i}) = 0 \qquad \text{if } \overline{s}_{-i} \in S_{-i} \setminus \overline{S}_{-i},$$

$$e_{-i}(\overline{s}_{-i}) = \overline{c}_{-i}(\overline{s}_{-i}) \qquad \text{if } \overline{s}_{-i} \in S_{-i} \setminus \overline{S}_{-i},$$

$$e_{-i}(\overline{s}_{-i}) = \varepsilon\, \overline{c}_{-i}(\overline{s}_{-i}) \qquad \text{if } \overline{s}_{-i} \in \overline{S}_{-i}.$$

Indeed, it is obvious that $(1 - \sum_{s_{-i} \in S_{-i}} e_{-i}(s_{-i})) \cdot c_{-i}(\overline{s}_{-i}) + e_{-i}(\overline{s}_{-i}) = e_{-i}(\overline{s}_{-i}) = \overline{c}_{-i}(\overline{s}_{-i})$ if $\overline{s}_{-i} \in S_{-i} \setminus \overline{S}_{-i}$. Next, consider some $\overline{s}_{-i} \in \overline{S}_{-i}$. Then,

$$\left(1 - \sum_{s_{-i} \in S_{-i}} e_{-i}(s_{-i})\right) \cdot c_{-i}(\overline{s}_{-i}) + e_{-i}(\overline{s}_{-i})$$

$$= \left[1 - \sum_{s_{-i} \in \overline{S}_{-i}} \varepsilon\, \overline{c}_{-i}(s_{-i}) - \sum_{s_{-i} \in S_{-i} \setminus \overline{S}_{-i}} \overline{c}_{-i}(s_{-i})\right]$$

$$\cdot \frac{\overline{c}_{-i}(s_{-i})}{\sum_{s_{-i} \in \overline{S}_{-i}} \overline{c}_{-i}(s_{-i})} + \varepsilon\, \overline{c}_{-i}(\overline{s}_{-i})$$

$$= \left[\overline{c}_{-i}(\overline{s}_{-i}) - \varepsilon\, \overline{c}_{-i}(\overline{s}_{-i}) \sum_{s_{-i} \in \overline{S}_{-i}} \overline{c}_{-i}(s_{-i}) - \overline{c}_{-i}(\overline{s}_{-i}) \sum_{s_{-i} \in S_{-i} \setminus \overline{S}_{-i}} \overline{c}_{-i}(s_{-i})\right.$$

$$\left. + \varepsilon\, \overline{c}_{-i}(\overline{s}_{-i}) \sum_{s_{-i} \in \overline{S}_{-i}} \overline{c}_{-i}(s_{-i})\right] \cdot \frac{1}{\sum_{s_{-i} \in \overline{S}_{-i}} \overline{c}_{-i}(s_{-i})}$$

$$= \left[\overline{c}_{-i}(\overline{s}_{-i}) - \overline{c}_{-i}(\overline{s}_{-i}) \cdot (1 - \sum_{s_{-i} \in \overline{S}_{-i}} \overline{c}_{-i}(s_{-i}))\right] \cdot \frac{1}{\sum_{s_{-i} \in \overline{S}_{-i}} \overline{c}_{-i}(s_{-i})}$$

$$= \overline{c}_{-i}(\overline{s}_{-i}).$$

Now we can define weakly perfect rationalizability by the following iterative procedure.

Definition 4 (weakly perfect rationalizability) Let $\varepsilon > 0$ be given. Let $S^0(\varepsilon) = S$. For $k \geq 1$, $S^k(\varepsilon) = \prod_{i \in I} S_i^k(\varepsilon)$ is inductively defined as follows: s_i belongs to $S_i^k(\varepsilon)$ if there is $c_{-i} \in \Delta^\varepsilon(\prod_{j \neq i} S_j^{k-1}(\varepsilon))$ such that s_i is a best response against c_{-i} within S_i. The set $S^\infty(\varepsilon) = \lim_{k \to \infty} S^k(\varepsilon)$ is the set of ε-weakly perfectly rationalizable strategy profiles and $S^\infty = \lim_{\varepsilon \to 0^+} S^\infty(\varepsilon)$ the set of weakly perfectly rationalizable strategy profiles.

In Definition 4 the limit set S^∞ is given by

$$\lim_{\varepsilon \to 0^+} S^\infty(\varepsilon) = \left\{s \in S \mid \exists \{\varepsilon^t\}_{t=0}^\infty \to 0^+, \exists \{s^t\}_{t=0}^\infty \to s, s^t \in S^\infty(\varepsilon^t)\right\}.$$

Alternatively, since $S^\infty(\varepsilon)$ is decreasing in ε and closed, it holds that $S^\infty = \bigcap_{\varepsilon > 0} S^\infty(\varepsilon)$.

Instead of the algorithmic definition given here, it is possible to define ε-weakly perfect rationalizability in an axiomatic way following the seminal contribution of Pearce [10]. The following three axioms characterize ε-weakly perfect rationalizability.

A1. Each player i forms a subjective prior over her opponents choice of strategy, i.e. a prior over mixed strategy profiles $c_{-i} \in \Delta(S_{-i})$ played subject to an error e_{-i} satisfying $0 < e_{-i}(s_{-i}) \leq \varepsilon$, $s_{-i} \in S_{-i}$.
A2. Each player maximizes her utility relative to her prior.
A3. A1 and A2 are common knowledge.

Similarly to Herings and Vannetelbosch [9] it is possible to obtain the following result.

Theorem 1 *For every normal-form game $\Gamma(I, S, U)$, S^∞ is non-empty and contains all perfect Nash equilibria.*

Although Theorem 1 shows that weakly perfect rationalizability is a well-defined solution concept, it does not give us an easy characterization of the strategy profiles that survive.

3 The equivalence theorem

Theorem 2 gives an easy characterization of the set of weakly perfect rationalizable strategies. It states that weakly perfect rationalizability coincides with the DF procedure. In the proof of Theorem 2 we will frequently use the following lemma from Pearce [10].

Lemma 1 *A strategy s_i is strictly dominated in T if and only if it is not a best response against a correlated conjecture on T_{-i}. A strategy s_i is weakly dominated in T if and only if it is not a best response against a completely mixed correlated conjecture on T_{-i}.*

Theorem 2 *For every normal-form game $\Gamma(I, S, U)$, $P^\infty = S^\infty$.*

Proof Let $\bar{\varepsilon} = (\prod_{i \in I} \#S_i)^{-1}$. For every product set T of pure strategy profiles, if a pure strategy $s_i \in T_i$ is strictly dominated in T, then by Lemma 1 it is not a best response against any correlated conjecture on T_{-i}. By continuity of the payoff function, there is $\varepsilon(i, T)$, $0 < \varepsilon(i, T) \leq \bar{\varepsilon}$, such that a pure strategy $s_i \in T_i$ that is strictly dominated in T is not a best response against any conjecture in

$$\{c_{-i} \in \Delta^0(S_{-i}) \mid c_{-i}(s_{-i}) \leq \varepsilon(i, T) \text{ if } s_j \notin T_j \text{ for some } j \neq i\}.$$

We denote the minimum over all players i and all product sets T of $\varepsilon(i, T)$ by $\hat{\varepsilon}$. We show by induction on k that $P_i^k = S_i^k(\varepsilon)$, for $\varepsilon \in (0, \hat{\varepsilon}]$.

Clearly, $P_i^0 = S_i^0(\varepsilon) = S_i$.

Strategy s_i belongs to P_i^1 if and only if it is not weakly dominated in S. By Lemma 1 s_i is not weakly dominated in S if and only if it is a best response against a completely mixed correlated conjecture on S_{-i}. Strategy s_i is a best response against a completely mixed correlated conjecture on S_{-i} if and only if it is a best response against a conjecture in $\Delta^\varepsilon(\prod_{j \neq i} S_j^0(\varepsilon))$ (indeed, $\Delta^\varepsilon(\prod_{j \neq i} S_j^0(\varepsilon)) = \Delta^0(S_{-i})$ for all $\varepsilon > 0$), which is the case if and only if s_i belongs to $S_i^1(\varepsilon)$. So, $P_i^1 = S_i^1(\varepsilon)$.

Now, let $k \geq 2$ and let $P^{k-1} = S^{k-1}(\varepsilon)$.

Consider any $s_i \in P_i^k$. By Lemma 1, s_i is a best response against some correlated conjecture $\hat{c}_{-i}$ on $\prod_{j \neq i} P_i^{k-1} = \prod_{j \neq i} S_j^{k-1}(\varepsilon)$. Clearly, $s_i \in P_i^1$, since $s_i \in P_i^k \subseteq P_i^1$. So by Lemma 1 s_i is a best response against a completely mixed correlated conjecture $\bar{c}_{-i} \in \Delta^0(S_{-i})$. There is a convex combination of $\hat{c}_{-i}$ and $\bar{c}_{-i}$ belonging to $\Delta^\varepsilon(\prod_{j \neq i} S_j^{k-1}(\varepsilon))$. It is sufficient to put a weight low enough on $\bar{c}_{-i}$. It follows that s_i is a best response against this convex combination, so $s_i \in S_i^k(\varepsilon)$.

Consider any $s_i \in S_i^k(\varepsilon)$. Then s_i is a best response against some $c_{-i} \in \Delta^\varepsilon(\prod_{j \neq i} S_j^{k-1}(\varepsilon))$. By the construction of $\hat{\varepsilon}$, s_i is not strictly dominated in $\prod_j S_j^{k-1}(\varepsilon) = P^{k-1}$. So $s_i \in P_i^k$. $\square$

For every normal-form game $\Gamma(I, S, U)$, a pure strategy survives one round of deletion of weakly dominated strategies followed by iterated deletion of strategies that are strictly dominated if and only if it is weakly perfectly rationalizable.

Theorem 2 allows us to advocate DF procedure for deleting strategies since it is obtained both under the assumption that there is some small uncertainty about the payoffs (see Dekel and Fudenberg [6]) and under the assumption that there is some small uncertainty about the strategies (Theorem 2).

In Herings and Vannetelbosch [9] it is shown that, for the case of uncorrelated conjectures, the concepts of perfect rationalizability [2], cautious rationalizability [10], proper rationalizability [11], weakly perfect rationalizability [9], and trembling-hand perfect rationalizability [9] are different in two-person games. For those games there is no distinction between correlated and uncorrelated conjectures. Weakly perfect rationalizability is the only existing refinement of rationalizability based on strategy perturbations that coincides with the DF procedure.

4 Uncorrelated errors

We will show by means of an example that if a player believes that her opponents choose in an uncorrelated way strategies that are irrational, then weakly perfect rationalizability does not coincide with the DF procedure. Such beliefs are rational if, for instance, her opponents implement a correlated strategy $c_{-i} \in \Delta(S_{-i})$ by means of a mediator. The mediator randomly selects a pure strategy profile $s_{-i} \in S_{-i}$ with probability $c_{-i}(s_{-i})$. Then the mediator recommends a player j, $j \neq i$, confidentially to use strategy s_j if s_{-i} is the pure strategy profile selected. If the mediator makes no errors in randomly selecting a pure strategy profile

$s_{-i} \in S_{-i}$ with probability $c_{-i}(s_{-i})$, but makes uncorrelated mistakes when doing his recommendations, or players make uncorrelated mistakes when carrying out the recommendation, then this leads to uncorrelated mistakes of the players.

Consider the following three-player game in Fig. 1. It is rather obvious that only player 1's strategy X_1, player 2's strategy Y_1, and player 3's strategies Z_1 and Z_2 survive the DF procedure. Indeed, for $k \geq 1$, $P_1^k = \{X_1\}$, $P_2^k = \{Y_1\}$, and $P_3^k = \{Z_1, Z_2\}$. By Theorem 2 this coincides with the strategies selected by weakly perfect rationalizability.

	Y_1	Y_2
X_1	1,1,1	1,0,1
X_2	0,1,1	0,0,0

Z_1

	Y_1	Y_2
X_1	1,1,1	1,0,0
X_2	0,1,0	0,0,1

Z_2

Figure 1. Correlation of mistakes matters

Now consider a version of weakly perfect rationalizability, where players are sure that their opponents make uncorrelated errors. Although this goes against the intuition underlying the conjectured use of correlated strategy profiles, it is a possibility we want to examine. Since the possible conjectures of players are now more restricted, it is obvious that $\widetilde{S}_1^k(\varepsilon) = \{X_1\}$ and $\widetilde{S}_2^k(\varepsilon) = \{Y_1\}$, $k \geq 1$, where a tilde is used to indicate that we are considering the case with uncorrelated errors. If player 3 conjectures that players 1 and 2 are going to play the strategy profile $\{X_2, Y_1\}$, then, if the probability ε by which an irrational strategy is conjectured is sufficiently small, player 3 chooses Z_1. Similarly, if player 3 conjectures that players 1 and 2 are going to coordinate on the strategy profile $\{X_2, Y_2\}$, then player 3 chooses Z_2. Consequently, $\widetilde{S}_3^1(\varepsilon) = \{Z_1, Z_2\}$.

At stage 2, player 3 knows that players 1 and 2 will coordinate on the strategy profile $\{X_1, Y_1\}$. But if player 3 believes that players 1 and 2 make uncorrelated errors, then player 3 will optimize against a conjecture $c_{-3} \in \Delta^0(S_{-3})$ for which $c_{-3}(X_2, Y_2) < c_{-3}(X_2, Y_1)$ and $c_{-3}(X_2, Y_2) < c_{-3}(X_1, Y_2)$. Against such a conjecture it is always optimal for player 3 to use strategy Z_1. It follows that $\widetilde{S}_3^k(\varepsilon) = \{Z_1\}$, $k \geq 2$.

Weakly perfect rationalizability with uncorrelated errors does not coincide with the DF procedure.

5 Two examples

We analyze two examples to conclude. The first example in Fig. 2 is due to Börgers [3]. Börgers' example is a counterexample to Dekel and Fudenberg's [6, Footnote 4] assertion that in two-player normal-form games perfect rationalizability coincides with the DF procedure. Indeed, it can be shown that

	Y_1	Y_2	Y_3
X_1	3 , 0	1 , 0	0 , 0
X_2	0 , 0	1 , 0	3 , 0
X_3	2 , 0	0 , 0	2 , 0
X_4	0 , 3	0 , 2	0 , 0
X_5	0 , 0	0 , 2	0 , 3

Figure 2. Börgers' example

only player 1's strategies X_1, X_2 and player 2's strategies Y_1, Y_2, Y_3 are perfectly rationalizable (see Börgers [3, pp. 274-275] or Herings and Vannetelbosch [9, p.65]). Meanwhile, player 1's strategies X_1, X_2, X_3 and player 2's strategies Y_1, Y_2, Y_3 survive the DF procedure and are weakly perfectly rationalizable. Indeed, $P_1^k = \{X_1, X_2, X_3\}$ and $P_2^k = \{Y_1, Y_2, Y_3\}$, and, it holds that $S_1^k(\varepsilon) = \{X_1, X_2, X_3\}$ and $S_2^k(\varepsilon) = \{Y_1, Y_2, Y_3\}$, $k \geq 1$.

	Y_1	Y_2	Y_3
X_1	2, 1, 1	1, 1, 1	0, 0, 1
X_2	0, 1, 1	1, 1, 1	0, 0, 1
X_3	2, 1, 1	0, 1, 1	0, 0, 1

Z_1

	Y_1	Y_2	Y_3
X_1	1, 1, 1	0, 1, 1	0, 0, 1
X_2	1, 1, 1	2, 1, 1	0, 0, 1
X_3	0, 1, 1	2, 1, 1	0, 0, 1

Z_2

	Y_1	Y_2	Y_3
X_1	1, 1, 0	1, 1, 0	0, 0, 0
X_2	1, 1, 0	1, 1, 0	0, 0, 0
X_3	0, 1, 0	0, 1, 0	2, 0, 0

Z_3

Figure 3. A three-player game

The second example shows the importance of allowing the players to have correlated conjectures in order to derive our equivalence result. The example is a three-player game (see Fig. 3) and is taken from Herings and Vannetelbosch [9], where we have shown that weakly perfect rationalizability without allowing correlated conjectures supports the following strategies: $\{X_1, X_2\}$ for player 1, $\{Y_1, Y_2\}$ for player 2, and $\{Z_1, Z_2\}$ for player 3. Next, consider the DF procedure. It is easily seen that $P_1^1 = \{X_1, X_2, X_3\}$, $P_2^1 = \{Y_1, Y_2\}$, and $P_3^1 = \{Z_1, Z_2\}$. It is not possible in the first iteration to eliminate any strategy of player 1, since all strategies of player 1 are equally good against $(c_2, c_3) = ((1/3, 1/3, 1/3), (1/3, 1/3, 1/3))$. In the second iteration of DF procedure it is impossible to eliminate any other strategy of player 2 or 3. Conse-

quently, for $k \geq 1$, $P_1^k = \{X_1, X_2, X_3\}$, $P_2^k = \{Y_1, Y_2\}$, and $P_3^k = \{Z_1, Z_2\}$. Given Theorem 2, we have $S_1^\infty = \{X_1, X_2, X_3\}$, $S_2^\infty = \{Y_1, Y_2\}$, and $S_3^\infty = \{Z_1, Z_2\}$. That is, X_3 is correlated weakly perfectly rationalizable but not uncorrelated weakly perfectly rationalizable. Intuitively, compared to strategies X_1 and X_2, strategy X_3 is good against the conjectures (Y_1, Z_1), (Y_2, Z_2), and (Y_3, Z_3), but bad against all other strategy combinations. Correlation allows any combination of the first three conjectures to arise with very high probability, which is not possible when conjectures are uncorrelated.

References

1. Ben-Porath, E.: Rationality, Nash equilibrium and backwards induction in perfect-information games. Review of Economic Studies **64**, 23–46 (1997)
2. Bernheim, D.: Rationalizable strategic behavior. Econometrica **52**, 1007–1028 (1984)
3. Börgers, T.: Weak dominance and approximate common knowledge. Journal of Economic Theory **64**, 265–276 (1994)
4. Brandenburger, A.: Lexicographic probabilities and iterated admissibility. In: Dasgupta, P., et al. (eds.) Economic analysis of markets and games, pp. 282–290. Cambridge, MA: MIT Press 1992
5. Brandenburger, A., Dekel, E.: Rationalizability and correlated equilibria. Econometrica **55**, 1391–1402 (1987)
6. Dekel, E., Fudenberg, D.: Rational behavior with payoff uncertainty. Journal of Economic Theory **52**, 243–267 (1990)
7. Gul, F.: Rationality and coherent theories of strategic behavior. Journal of Economic Theory **70**, 1–31 (1996)
8. Hammond, P.: Aspects of rationalizable behavior. In: Binmore, K., Kirman, A., Tani, P. (eds.) Frontiers of game theory , pp. 277–305. Cambridge, MA: MIT Press, 1993
9. Herings, J.J. and Vannetelbosch, V.J.: Refinements of rationalizability for normal-form games. International Journal of Game Theory **28**, 53–68 (1999)
10. Pearce, D.G.: Rationalizable strategic behavior and the problem of perfection. Econometrica **52**, 1029–1050 (1984)
11. Schuhmacher, F.: Proper rationalizability and backward induction. Mimeo, University of Bonn, (1995)