JOURNAL OF ECONOMIC THEORY **74**, 40–61 (1997)
ARTICLE NO. ET962252

On Rationalizability in Extensive Games

Pierpaolo Battigalli*

Department of Economics, Princeton University, Princeton, New Jersey 08544-1021

Received May 2, 1991; revised July 26, 1996

This paper analyzes Pearce's notion of extensive form rationalizability (EFR). Although EFR was originally defined as a reduction procedure, this paper shows that it can be characterized in terms of restrictions on players' updating systems. These restrictions correspond to a common hierarchy of nested hypotheses. Next the relationship of EFR to more familiar reduction procedures is examined. In generic games of perfect information, EFR is realization-equivalent to iterated weak dominance and backward induction. Equivalence with iterated weak dominance is complete in the subset of games with "iterated perfect information." *Journal of Economic Literature* Classification Number: C72. © 1997 Academic Press

1. INTRODUCTION

In the last few years the Nash equilibrium concept and its most popular refinements[1] (in particular sequential and perfect equilibria) have been seriously criticized along two very different but complementary lines.

On one hand, Bernheim and Pearce have convincingly argued that if a game is played "one-shot" and without preplay communication, there is no *a priori* reason to expect that players will be able to predict opponents' behavior. Therefore there is no reason to expect a Nash equilibrium outcome. But this does not mean that strategic rationality cannot be properly formalized. Bernheim and Pearce defined the concept of *rationalizability* in order to characterize those strategy profiles that could conceivably be played when (a) there is complete information concerning the game to be played, (b) each player maximizes her expected payoff, given some conjecture on her opponents' behavior, (c) it is common knowledge that (a) and (b) hold. In two-person finite normal-form games, rationalizability is equivalent to

* Financial support from the Departments of Economics of Tel Aviv University and Bocconi University is gratefully acknowledged. I thank Ariel Rubinstein for his encouragement and his helpful suggestions. Faruk Gul and Phil Reny found a serious mistake in the previous version of this paper. I also thank an anonymous referee and an associate editor for helpful comments. Of course, all the remaining errors are mine.

[1] See Van Damme [7] and references therein.

0022-0531/97 $25.00

iterated deletion of strictly dominated strategies. In every normal-form game the set of Nash equilibria is included in the set of rationalizable strategies (see Bernheim [3] and Pearce [19]).

On the other hand, there are many extensive games in which sequential, perfect (and also proper) equilibria are not restrictive enough, because they fail to eliminate intuitively incredible threats, or promises. The criteria used to judge such strategies as incredible attempt to formalize the so-called *forward induction principle*: a player should use all information she acquired about her opponents' past behavior in order to improve her prediction of their future, simultaneous, and past (unobserved) behavior, relying on the assumption that they are rational.

Forward induction criteria are usually embedded in refinements of the Nash equilibrium concept which restrict beliefs off-the-equilibrium-path. One way to do it is to check whether a given "obvious way to play the game," or "candidate solution," is stable against deviations, given that such deviations are interpreted as conscious signals. Note that the interpretation depends on the candidate solution which is being tested.[2] If the candidate solution fails the test, it should not be considered an obvious way to play the game, and this may undermine the interpretation of deviations which caused the solution to fail the test. Another approach is to interpret deviations as the result of a mistaken theory: the deviator is supposed to believe in a "wrong equilibrium." But Bernheim's and Pearce's critique applies even more forcefully to these situations: why should one restrict attention to alternative hypotheses given by Nash equilibria, when something unpredictable has happened? and why should the deviator and not the observer be wrong? However, in many interesting cases it is possible to provide very compelling forward induction arguments, which neither depend on a candidate solution, nor rely on equilibrium assumptions (see, e.g., Pearce [19] and Hammond [11]). Therefore, it would be interesting to obtain these restrictions, as well as more conventional backward induction arguments, as the by-product of a formal, non-equilibrium theory of strategic rationality in extensive games. This is exactly what Pearce tried to accomplish with his notion of *extensive-form rationalizability* (EFR). According to EFR, it is common knowledge that no player, at an information set h, would choose an action which is not optimal with respect to any "rational" conjecture consistent with h.

Pearce's solution is quite successful in formalizing many intuitive forward and backward induction arguments, which have been used to solve interesting

[2] This is made transparent in Sobel *et al.* [27] where forward induction equilibria of signaling games are characterized by applying extensive form rationalizability to a modified game where the proposed equilibrium path is replaced with a message yielding the proposed-equilibrium expected payoffs.

games. The theory's unintuitive predictions (such as the lack of any cooperation in the finitely repeated Prisoner Dilemma or of reputation effects in Selten's [26] Chainstore game) coincide with those of more conventional equilibrium theories, suggesting that it would be instructive to understand why EFR fails to match our intuition. Moreover, EFR subsumes normal-form rationalizability, since they coincide for simultaneous games. However, EFR has not received as much attention as normal-form rationalizability.[3] There are at least two reasons for this neglect. One is simply that Pearce's definition of EFR is quite difficult to understand and tricky to handle in formal analysis. The other is that EFR has been seriously criticized for being inadequate in dealing with "counterfactuals" (e.g., Reny [20] and Basu [1]). Many games possess information sets that cannot be reached by rationalizable strategies, yet the players are assumed to optimize against "rational" conjectures also at these information sets. Hence it seems that the theory assumes that "rationality" is common knowledge at "irrational" information sets. This sounds like a contradiction! Note that the more conventional notion of subgame perfection has been similarly criticized (see, e.g., Rosenthal [23]). In fact it has been noted that EFR tends to be equivalent to subgame perfection in games which can be solved by backward induction.

The aim of this paper is to provide additional insights about EFR, attempting to overcome the two drawbacks mentioned above and comparing EFR with more conventional solution concepts. Section 2 briefly describes the game-theoretic set-up. Section 3 provides an alternative characterization of EFR, originally defined as an iterative reduction procedure on the strategies. This new definition is easier to understand and can be *consistently* interpreted as a theory of strategic rationality. Such a theory has to assume (1) that players are rational and endowed with a *hierarchy of hypotheses*, and (2) that this is common knowledge (or, more precisely, "common belief") *at the beginning of the game*. Therefore, one has to assume both "more" and "less" than common knowledge of rationality to obtain EFR: "more," because there must be a hierarchy of hypotheses, and "less," because the theory is not necessarily common knowledge at each information set. Section 4 analyzes the relationship between EFR, iterated weak dominance and subgame perfection. It is shown that in generic games with "iterated perfect information," EFR is equivalent to the former and realization-equivalent to the latter. A realization-equivalence result also holds for general perfect information games. This illustrates the similarity between rationalizability and backward induction.

[3] Most recent textbooks of game theory (e.g., [9, 18, 25]) devote a chapter or a section to normal-form rationalizability, but extensive-form rationalizability is not discussed.

2. THE GAME-THEORETIC SET-UP

The analysis is restricted to finite extensive games with perfect recall and *without chance moves*. Since the formal description of an extensive game is by now standard (see, e.g., Kuhn [14] or van Damme [7]), only the necessary symbols with very terse explanations are given below.

Notation	Terminology
I	the set of players
$-i = I\backslash\{i\}$	the opponents of player i
X^i	the set of decision nodes of i
H^i	the collection of information sets of i (a partition of X^i)
$S = S^i \times S^{-i}$	the set of (pure) strategy profiles
Z	the set of terminal nodes
$u^i\colon Z \to \mathbb{R}$	player i's utility function
$\zeta\colon S \to Z$	the outcome function: $\zeta(s)$ is the terminal node induced by the strategy profile s
$U^i = u^i \circ \zeta$	player i's normal-form utility function

By perfect recall, a given player's information sets can be partially ordered without ambiguities according to the precedence relation between the respective nodes. Let $g, h \in H^i$; *g comes before h* (or *h follows g*) if and only if every node of h follows a node of g.

In the following analysis there is no need to assume that players actually choose mixed strategies. Mixed strategies are interpreted as beliefs. Let $\Delta(Y)$ denote the space of probability measures on a given set Y. A mixed strategy $c^i \in \Delta(S^i)$ represents a probabilistic conjecture concerning player i. The letter "c" is used to stress this interpretation. A *conjecture* of player i is a probability measure $c^{-i} \in \Delta(S^{-i})$. If i regards her opponents' strategies as independent random elements, then her conjecture space is restricted to the set of *uncorrelated conjectures*[4]

$$\Delta^*(S^{-i}) := \left\{ c^{-i} \in \Delta(S^{-i}) \mid \exists (c^j)_{j \neq i} \in \left(\mathop{\times}_{j \neq i} \Delta(S^j) \right), \forall s^{-i} \in S^{-i}, c^{-i}(s^{-i}) = \prod_{j \neq i} c^j(s^j) \right\}.$$

[4] The justifications for having uncorrelated conjectures at unreached information sets are discussed critically in Kreps and Ramey [13] and Fudenberg *et al.* [8]. Battigalli [2] shows that having an uncorrelated conjecture at each information set does not properly formalize the above mentioned notion of stochastic independence in the context of extensive games. Lemma 1 in the Appendix shows that the difference between correlated and uncorrelated conjectures is immaterial is games with perfect information.

Player i's expected payoff given (s^i, c^{-i}) is denoted by $U^i(s^i, c^{-i})$.

A conjecture c^{-i} *reaches* an information set $h \in H^i$ if there are a strategy s^i and a profile s^{-i} in the support of c^{-i}, such that the path induced by (s^i, s^{-i}) reaches h. This means that conjecture c^{-i} is not falsified if h is reached. Similarly s^i reaches h if there is an s^{-i} such that the path induced by (s^i, s^{-i}) reaches h. One may check that the probability of reaching an information set $h \in H^i$ given (s^i, c^{-i}) is positive if and only if both s^i and c^{-i} reach h. A set $B^{-i} \subseteq S^{-i}$ reaches h if there is a profile $s^{-i} \in B^{-i}$ reaching h.

3. CHARACTERIZATION AND INTERPRETATION OF EXTENSIVE FORM RATIONALIZABILITY

Pearce's original definition of EFR may be difficult to understand, not only because it is quite complex, but also because it is not completely clear how to interpret it. Moreover, according to some critics, the informal theory which EFR is supposed to represent is contradictory (see Reny [20] and Binmore [5]).

Consider the game of Fig. 1. EFR implies that player 1 chooses A. In fact, A strictly dominates R and at information set h player 2 should believe that 1 has played L, otherwise 1 would be irrational, an instance of forward induction. But this induces 2 to play l at h, which in turn induces 1 to choose A. Therefore h is inconsistent with common knowledge of rationality. Why should 2 believe at h that 1 is not irrational? Might not player 1 recognize that she may be able to trick player 2 into playing r, just by pretending to be irrational? If she does, she may hope to grab \$2, her best payoff, playing L.

These problems concerning counterfactuals and the strategic manipulation of beliefs are important. EFR, like most extensive game theories,

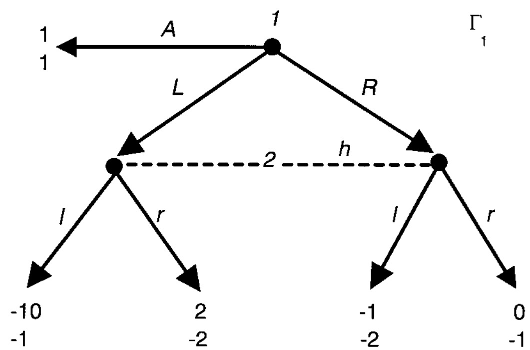

FIG. 1. A game in extensive form.

ignores or, at best, "sterilizes" them. But this does not mean that EFR is inherently contradictory. Note that game Γ_1 has been solved by a *reduction procedure*, which progressively eliminates strategies. In the alternative characterization of EFR given in this section there is no iterated deletion of *strategies*. Only players' *beliefs* are restricted. Since conjectures in an extensive game may be falsified and then have to be replaced, EFR restrictions will not concern isolated conjectures, but updating systems of conjectures. Such restrictions will be modeled as a hierarchy of hypotheses, ruling out strategic manipulation. It will be shown that this hierarchy corresponds to the sequence of strategy sets given by Pearce's iterative deletion procedure. For example, in game Γ_1, it is plausible to assume that player 2 at h believes that player 1 does not choose strictly dominated strategies because this belief is a minimal rationality hypothesis consistent with information h. According to EFR, player 1 believes that player 2 would hold this minimal hypothesis if h were reached.

The building block of the analysis is a theory of "naive" individual rationality, which does not take into account truly strategic considerations. An individually rational player i maximizes her expected payoff at each information set h reached by the play, given her conjecture at h, $c^{-i}(h)$, where $c^{-i}(h)$ reaches h. But this is not enough. It is necessary to assume that she has a consistent pattern of expectations. She should not change her conjecture unless the play reaches an information set which falsifies it. If i's pattern of expectations fails to meet this consistency condition, she may change her plan of action without a good reason, actually implementing a strictly dominated strategy.

Definition 1. A *consistent updating system* for player i is a mapping $c^{-i}(\cdot)$: $H^i \to \Delta(S^{-i})$ such that for all information sets $g, h \in H^i$:

(i) $c^{-i}(h)$ reaches h,

(ii) if g comes before h and $c^{-i}(g)$ reaches h, then $c^{-i}(g) = c^{-i}(h)$.

A consistent updating system for player i is *uncorrelated* if its range is included in $\Delta^*(S^{-i})$.

Consistency of the updating system requires that the conjecture at h is consistent with h being reached and that no conjecture is abandoned unless falsified. That is, players update according to Bayes rule whenever possible.

A strategy is individually rational if it is a *weakly sequential best response* to some consistent updating system. The phrase "weakly sequential" is used because, unlike much of the refinement literature, here it is not required that a strategy specifies behavior at information sets that can not be reached by that strategy. (For more on this see Rubinstein [24] and Reny [21].) According to the following definition, $R(1)$ is the set of individually rational strategy profiles. Higher degrees of strategic rationality are constructed recursively.

Definition 2. Let $R_c(0) := S$. Then $R_c(n)$ $(n \geqslant 1)$ is inductively defined as follows: for all $i \in I$, $s^i \in R^i_c(n)$ if there exists a consistent updating system $c^{-i}(\cdot)$ such that

(a) for all $h \in H^i$ and all $k \in \{0, ..., n-1\}$, if $R^{-i}_c(k)$ reaches h, then $c^{-i}(h) \in \Delta(R^{-i}_c(k))$,

(b) for all $h \in H^i$, if s^i reaches h, then s^i is a best response to $c^{-i}(h)$ at h, that is,

$$\text{for all} \quad t^i \in S^i, \; U^i(s^i, c^{-i}(h)) \geqslant U^i(s^i/t^i_h, c^{-i}(h)),$$

where s^i/t^i_h is the strategy which results from s^i when behavior at h and its followers in H^i is specified by t^i.

The subscript c in $R^i_c(n)$ is mnemonic for "correlation." $R(n)$ is defined in the same way with the additional condition that $c^{-i}(\cdot)$ must be uncorrelated.

Remark 1. $\{R(n);\ n \geqslant 0\}$ is a weakly decreasing sequence, i.e., $R(n+1) \subseteq R(n)$ for all n. Since S is finite, the sequence converges in finitely many steps. The limit set is given by the first integer N such that $R(N) = R(N+1)$. The same holds for $R_c(\cdot)$.

Definition 2 and Remark 1 can be interpreted as follows. Consider the limit set $R(N)$. The sequence $R^{-i}(0), R^{-i}(1), ..., R^{-i}(N)$ represents a hierarchy of increasingly strong hypotheses about the behavior of player i's opponents. When player i implements a strategy $s^i \in R^i(N)$, she always holds the strongest hypothesis which is consistent with her information and optimizes accordingly. At the beginning of the game, it is a common belief that all players update and behave in this way. A similar iterative procedure is incorporated in Reny's [21] notion of "explicability." Yet he offers no explicit characterization of Pearce's notion of EFR.[5]

Let $P(n)$ denote the set of pure strategies which survive the nth iteration of Pearce's procedure.

Definition 3. Let $P_c(0) := S$. Then $P_c(n)$ $(n \geqslant 1)$ is inductively defined as follows: for all $i \in I$, $s^i \in P^i_c(n)$ if

(a) $s^i \in P^i_c(n-1)$,

(b) there exists a consistent updating system $c^{-i}(\cdot)$ such that for all $h \in H^i$, if s^i and $P^{-i}_c(n-1)$ reach h then

[5] It can be shown that Reny's [21] (weakest) procedure corresponds to EFR in two-player games, but not in n-player games.

(b′) $c^{-i}(h) \in \Delta(P_c^{-i}(n-1))$,

(b″) for all $t^i \in P_c^i(n-1)$, $U^i(s^i, c^{-i}(h)) \geqslant U^i(s^i/t_h^i, c^{-i}(h))$.[6]

$P(n)$ is defined in the same way with the additional condition that $c^{-i}(\cdot)$ must be uncorrelated. The set of *rationalizable* (*correlated-rationalizable*) *strategy profiles* is $\bigcap_{n>0} P(n)$ ($\bigcap_{n>0} P_c(n)$).

There are two differences between Definition 2 and 3: (i) all strategies in S^i are candidates for $R^i(n)$, while only strategies in $P^i(n-1)$ are candidates for $P^i(n)$, thus maximization is performed over a smaller domain of "competitors" in Definition 3; (ii) maximization is performed for a smaller class of information sets in Definition 3.

These differences depend on $R(n)$ being determined by restrictions on players' updating systems and $P(n)$ by a reduction procedure. Nevertheless, Theorem 1 shows that the two definitions are equivalent.

THEOREM 1. *For all* $n \geqslant 0$, $R(n) = P(n)$ *and* $R_c(n) = P_c(n)$.

Proof. See the Appendix.

COROLLARY 1. *Let* $N(N_c)$ *be the smallest integer such that* $R(N) = R(N+1)$ $(R_c(N_c) = (R_c(N_c+1))$. *Then* $R(N) = \bigcap_{n>0} P(n) \neq \varnothing$ $(R_c(N_c) = \bigcap_{n>0} P_c(n) \neq \varnothing)$.

Proof. This follows by inspection of Definition 3 and Theorem 1. At each step it is possible to find a consistent updating system $c^{-i}(\cdot)$ which meets condition (b′) for all h reached by $P(n-1)$. In addition, a strategy s^i satisfying (b″) can be found by dynamic programming, relying on consistency of $c^{-i}(\cdot)$ and finiteness of S^i. ∎

4. RATIONALIZABILITY, ITERATED DOMINANCE, AND BACKWARD INDUCTION

This section explores the differences and similarities between EFR, iterative deletion of weakly dominated strategies and subgame perfection (with the associated backward induction procedure) in games of perfect information. It is shown that in generic perfect information games these solution concepts are realization-equivalent and that EFR is completely equivalent to iterated weak dominance on a subset of such games. Before stating the main results it is useful to discuss some examples.

[6] Condition (b″) differs slightly from the original definition in Pearce [19], as Pearce considers h-replacements $s^i/t_h^i \in P_c^i(n-1)$. The present definition simplifies the analysis and yet captures the spirit of the original one. It is shown in the working paper version of this paper that the difference is immaterial.

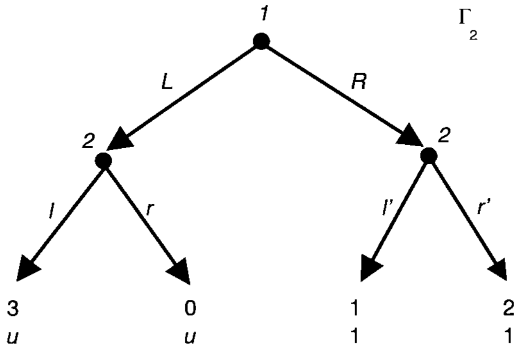

FIG. 2. A perfect information game with relevant ties.

On the Relevance of Ties between Payoffs at Terminal Nodes

The "genericity" property used here says that no player with deterministic conditional expectations is ever indifferent between two actions. Formally, an extensive game is *without relevant ties* if for all $z', z'' \in Z$, all $i \in I$, if $z' \neq z''$ and i is the player moving at the last common predecessor of z' and z'',[7] then $u^i(z') \neq u^i(z'')$.

The proof of the following remark is left to the reader.

Remark 2. An extensive game (without chance moves) has no relevant ties if and only if whenever a player can affect the outcome (given the opponents' strategies) she also affects her own payoff; that is, for all $i \in I$, all $s^i, t^i \in S^i$, all $s^{-i} \in S^{-i}$, $\zeta(s^i, s^{-i}) \neq \zeta(t^i, s^{-i})$ implies $U^i(s^i, s^{-i}) \neq U^i(t^i, s^{-i})$.

The example depicted in Fig. 2 illustrates the meaning and the role of relevant ties. In game Γ_2 there are exactly two relevant ties, which make player 2 always indifferent between "left" and "right." EFR and iterated weak dominance have no bite in this game, whereas the equilibrium condition rules out the pair of strategies (R, ll'). This is true independent of the value of the parameter u and $u = 1$ does not increase the number of relevant ties relative to the case where u is not equal to 1. Note however that the value of u may be relevant with respect to other solution concepts not discussed here: for example, if $u < 1$, player 2 could credibly announce that she is going to play r after L to induce player 1 to choose R.

Backward Iterated Dominance and EFR

For all games where the backward induction procedure can be unambiguously used to compute the set of subgame perfect equilibria, in

[7] The last common predecessor of two nodes t' and t'' in a tree is the maximal element of the set $\{x \mid x < t' \text{ and } x < t''\}$, where $<$ is the precedence relation.

particular for all games without relevant ties, it is possible to define a corresponding order of (non-maximal) elimination of weakly dominated strategies according to the following rule: if backward induction deletes action a at node x, delete all the strategies reaching x and choosing a (see, e.g., Osborne and Rubinstein [25, Sec. 6.6]). EFR and the maximal iterated deletion of weakly dominated strategies are best compared with this elimination procedure, which may be called *backward iterated dominance.*

According to Theorem 1, whenever backward iterated dominance and EFR correspond to the same order of elimination , the subgame perfect strategies at unreached decision nodes can be interpreted as rationalizable updated expectations. This kind of equivalence does not hold in the games without relevant ties discussed below. But Theorem 4 will show that all the considered solution procedures eventually select the subgame perfect terminal node.

In game Γ_3 (see [21, Fig. 3]) backward iterated dominance eliminates $a'a''$, $A'A''$, $a'd''$ and $A'D''$ in the given order and the unique subgame perfect equilibrium is $(D'D'', d'd'')$. EFR coincides with maximal iterated weak dominance and eliminates $\{A'D'', a'a''\}$, $\{A'A'', d'd'', d'a''\}$ in the given order. The rationalizable solution is $(D'., a'd'')$, yielding the same outcome, but prescribing a different strategy for player 2 and different beliefs for both players. Indeed, according to EFR, if player 2 were reached she would assume that player 1 is carrying out the "less irrational" strategy consistent with A', that is the undominated strategy $A'A''$. This yields choice a'.

Iterated Weak Dominance and Iterated Perfect Information

There are two definitions of iterated weak dominance (as well as two definitions of iterated strict dominance). Although they are not equivalent, both are widely used in the literature.

Fix two non-empty subsets $B^i \subseteq S^i$ and $B^{-i} \subseteq S^{-i}$. A mixed strategy $\mu^i \in \Delta(S^i)$ *dominates* $\delta^i \in \Delta(S^i)$ *on* B^{-i} if for all $s^{-i} \in B^{-i}$, $U^i(\mu^i, s^{-i}) \geqslant U^i(\delta^i, s^{-i})$ and for some $t^{-i} \in B^{-i}$ the inequality is strict. A pure strategy $s^i \in B^i$ is *inferior* in $B^i \times B^{-i}$ if there exists a *mixed* strategy $\mu^i \in \Delta(B^i)$ which

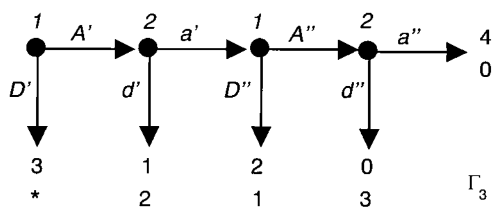

FIG. 3. A perfect information game without relevant ties where the unique subgame perfect strategy for player 2 is not rationalizable. (Irrelevant payoffs are replaced by *.)

dominates s^i on B^{-i}. A strategy $s^i \in B^i$ is *dominated in* $B^i \times B^{-i}$ if there exists a *pure* strategy $t^i \in B^i$ which dominates s^i. A strategy $s^i \in B^i$ is *redundant* on $B^i \times B^{-i}$ if there exists $t^i \in B^i$ such that for all $j \in I$, all $s^{-i} \in B^{-i}$, $U^j(s^i, s^{-i}) = U^j(t^i, s^{-i})$. *Player i* is redundant on $B^i \times B^{-i}$ if every $s^i \in B^i$ is redundant on $B^i \times B^{-i}$.

The sets $W(n)$ and $W_p(n)$ of iteratively non-inferior and iteratively undominated strategies are inductively defined as follows: $W(0) := S =: W_p(0)$; $W(n+1)$ is the set of profiles of strategies which are not inferior in $W(n)$; $W_p(n+1)$ is the set of profiles of strategies which are not dominated in $W_p(n)$. (W is mnemonic for *weak* dominance and subscript p is mnemonic for domination by a *pure* strategy.)

The $W(\cdot)$ procedure informally relies on the assumption that players' attitudes toward risk are common knowledge.[8] On the other hand the $W_p(\cdot)$ procedure only requires common knowledge of players' ordinal preferences over outcomes.[9] It is proved in the Appendix (Lemma 4) that in every game of perfect information a strategy is inferior if and only if it is weakly dominated. This does not seem surprising, since intuition suggests that attitudes toward risk should not be relevant in games of perfect information. However this equivalence does not generally hold for iterations of the two dominance procedures beyond the first step. This is shown by an example due to Faruk Gul and Phil Reny (private communication).[10]

In game Γ_4 (Fig. 4), player 1's strategy $IN.L'L''$ is dominated by OUT and the only undominated strategy for player 3 is $r'r''$. No other strategy is either dominated or inferior on S. Therefore, in the residual game, player 2's strategy *out* is dominated by mixed strategy $1/2[a] + 1/2[b]$, although it is dominated neither by a nor by b (*out* is also iteratively undominated). This can happen because the residual normal form game obtained after deleting all the dominated strategies does not correspond to any game of perfect information: while both actions L' and L'' can be chosen when a profile of undominated strategies is implemented, strategy $IN.L'L''$ does not belong to the residual normal form.

Fix an extensive form with perfect information E and for every extensive game $\Gamma(\mathbf{u}) = (E, \mathbf{u})$ ($\mathbf{u} \in \mathbb{R}^{Z \times I}$ is the vector of payoff functions) let $\{W_p(n, \mathbf{u})\}$ be the iterated dominance procedure on $\Gamma(\mathbf{u})$. For every $n \geqslant 0$,

[8] Stahl [28] offers a Bayesian characterization of the $W(\cdot)$ procedure (see also [2]).

[9] EFR can be modified to incorporate the assumption that only ordinal preferences over outcomes are common knowledge. Then a result due to Börgers [6, Lemma] can be used to show that, in games without relevant ties, the modified EFR is equivalent to the W_p procedure.

[10] This example also shows that Lemma 4 and Theorem 2 of a previous version of this paper are wrong. I am grateful to Faruk Gul and Phil Reny for pointing out this serious mistake.

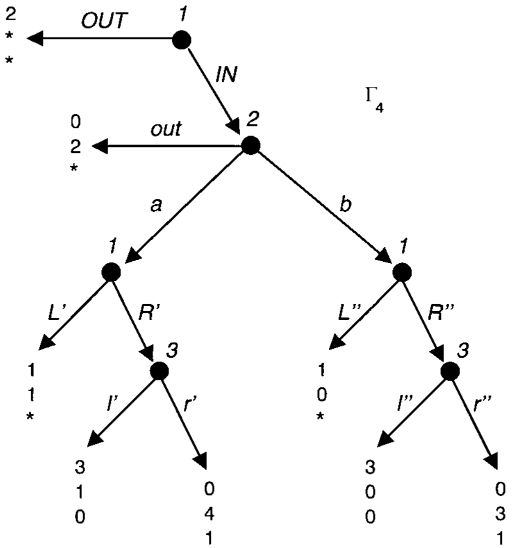

FIG. 4. A perfect information game without relevant ties where iterated deletion of inferior strategies does not coincide with iterated deletion of dominated strategies.

let $\Gamma(n, \mathbf{u})$ be the perfect information game obtained by restricting $\Gamma(\mathbf{u})$ to the set of nodes which are reached by some strategy profile in $W_{\rm p}(n, \mathbf{u})$ (that is the set of terminal nodes $\zeta(W_{\rm p}(n, \mathbf{u}))$ and all the predecessors of these terminal nodes). Extensive form E has *iterated perfect information* if for every payoff vector $\mathbf{u} = (u^i)_{i \in I}$ and every $n \geqslant 0$, the normal form of $\Gamma(n, \mathbf{u})$ is equivalent to $(I, W(n), (u^i \circ \zeta)_{i \in I})$ up to relabeling and deletion of redundant strategies and players.[11] It is easy to check that, for game Γ_4, $\Gamma(1)$ does not have the same reduced normal form as $(I, W(1), (u^i \circ \zeta)_{i \in I})$.

Note that, unlike the other games considered so far, the extensive form of Γ_4 does not satisfy the following property:

(IPI) for every player $i \in I$ and every pair of nodes $x', x'' \in X^i$ on distinct paths, *at least one* of the following conditions holds:

(a) x' and x'' have no common predecessor which is also a decision node of player i (i.e., $X^i \cap \{w \mid w < x' \text{ and } w < x''\} = \varnothing$);

(b) the last common predecessor of x' and x'' is a decision node of player i;

(c) all the non terminal successors of either x' or x'' are decision nodes of player i.

[11] This is essentially the same as saying that $\Gamma(n, \mathbf{u})$ and $(I, W_{\rm p}(n), (\zeta \circ u^i))$ have the same pure strategy reduced normal form (see Thompson [29]).

Several interesting examples in the literature satisfy (IPI): all the "Centipede-like" games, all the games whereby no player moves more than once in any path (cf. Bonanno's [5]), and some simple extensive forms discussed in the literature on subgame perfect implementation (e.g., Moore and Repullo [16, pp. 1119–1197]). It turns out that (IPI) characterizes iterated perfect information.

THEOREM 2. *If an extensive form E with perfect information satisfies* (IPI), *then it has iterated perfect information.*[12]

Proof. See the Appendix.

Equivalence Results

THEOREM 3. *For every game* $\Gamma = (E, \mathbf{u})$ *without relevant ties, if E has iterated perfect information, then iterated deletion of inferior strategies, iterated deletion of dominated strategies, uncorrelated and correlated rationalizability coincide, i.e., for all* $n \geqslant 0$, $W(n) = W_{\mathrm{p}}(n) = P(n) = P_c(n)$.

The proof is in the Appendix, where it is shown how the theorem can be extended to games *with* chance moves. The proof also makes clear that correlated rationalizability and iterated dominance coincide step by step in *all* games without relevant ties such that $W(n) = W_{\mathrm{p}}(n)$ for all n. Since the iterated dominance procedure $W_{\mathrm{p}}(\cdot)$ (without chance moves) only depends on ordinal preferences over terminal nodes, this shows that in all such games knowledge of players' attitudes toward risk does not play any role in the EFR solution and in the reasoning processes represented by EFR. Theorem 3 also shows that in games with iterated perfect information the EFR solution can be computed considering the reduced game trees $\Gamma(n)$ obtained by inductively deleting the nodes which are inconsistent with step $n-1$ (game Γ_4 shows that this is not true in games without iterated perfect information).

Theorem 3 can be used to provide a simple proof of the similarity between EFR and subgame-perfection in generic games with iterated perfect information. This is a consequence of the well known fact that in generic games of perfect information iterated dominance is outcome-equivalent to backward induction (Moulin [17] and Gretlein [10] proved this result for games where each player has a strict preference order over consequences, which are a function of the terminal node; for games without relevant ties one can use Marx and Swinkels' [15] result on order independence of

[12] Generalizing example Γ_4 one can show that also the converse is true: if E does not satisfy IPI, there is some $\mathbf{u}$ (actually an open set of payoff vectors) such that $(E, \mathbf{u})$ does not have iterated perfect information.

iterative dominance). This outcome-equivalence result can be extended to generic games of perfect information.

THEOREM 4 (cf. Reny [21, Proposition 3]). *In every perfect information game without relevant ties, all the profiles of un/correlated rationalizable strategies, iteratively undominated strategies and iteratively non inferior strategies yield the same terminal node as the unique subgame perfect equilibrium.*

Proof. See the Appendix.[13]

Theorems 2, 3, and 4 simplify the computation of EFR outcomes as well as establishing a Bayesian foundation for the subgame-perfect outcome in generic games of perfect information.

APPENDIX

Proof of Theorem 1

The proof for the case of uncorrelated conjectures is given, the proof for the other case is virtually identical.

By inspection of Definitions 2 and 3 $R(0) = P(0)$. Assume that $R(n-1) = P(n-1)$.

Let $s^i \in R^i(n)$. Since $R^i(n-1) = P(n-1)$ (induction hypothesis) and $R(n) \subseteq R(n-1)$, $s^i \in P^i(n-1)$. Thus condition (a) in Definition 3 is met. By Definition 2 and the induction hypothesis, there exists an uncorrelated consistent updating system $c^{-i}(\cdot)$ such that for all $h \in H^i$ reached by s^i:

(β') if $P^{-i}(n-1)$ reaches h, then $c^{-i}(h) \in \Delta^*(P^{-i}(n-1)) = \Delta^*(R^{-i}(n-1))$,

(β'') s^i is a best response at h to $c^{-i}(h)$.

(β') and (β'') imply (b′) and (b″) in Definition 3. Hence $s^i \in P^i(n)$.

Let $s^i \in P^i(n)$. Since $P^i(n-1) = R^i(n-1)$ (induction hypothesis) and $P(n) \subseteq P(n-1)$, $s^i \in R^i(n-1)$. Therefore there are two uncorrelated consistent updating systems $c_{n-1}^{-i}(\cdot)$ and $c_n^{-i}(\cdot)$ such that for all $h \in H^i$:

[13] Following Reny [21], the proof uses some properties of Kohlberg and Mertens' [2] "fully stable sets." Unfortunately, I am not aware of more intuitive proofs of this result. Alternatively, one can show that a generalization of the property mentioned in Remark 3 for the set of mixed strategies holds generically and then apply results about order independence of iterative dominance procedures (see Marx and Swinkels [15]). But the set of payoffs for which this extension holds does not satisfy nice structural properties such as no relevant ties.

$(n-1)$ if s^i reaches h, then s^i is a best response at h to $c_{n-1}^{-i}(h)$, and for all $k=0, ..., n-2$, if $R^{-i}(k)$ reaches h, then $c_{n-1}^{-i}(h) \in \Delta^*(R^{-i}(k))$ (see Definition 2 for $s^i \in R^i(n-1)$);

(n) if $R^{-i}(n-1)$ and s^i reach h, then $c_n^{-i}(h) \in \Delta^*(R^{-i}(n-1))$ and for all $t^i \in R^i(n-1)$, $U^i(s^i, c_n^{-i}(h)) \geqslant U^i(s^i/t_h^i, c_n^{-i}(h))$ (see Definition 3 for $s^i \in P^i(n)$, use the induction hypothesis, and recall that s^i/t_h denotes the strategy that selects the actions given by t^i at information set h and its followers, and the actions given by s^i otherwise).

It may be assumed without loss of generality that $c_n^{-i}(h) \in \Delta^*(R^{-i}(n-1))$ for all h reached by $R^{-i}(n-1)$, even if h is not reached by s^i. Now define $c^{-i}(\cdot)$ as follows:

$$c^{-i}(h) = \begin{cases} c_n^{-i}(h), & \text{if } \; R^{-i}(n-1) \text{ reaches } h \\ c_{n-1}^{-i}(h), & \text{otherwise.} \end{cases}$$

It is easy to check that $c^{-i}(\cdot)$ satisfies conditions (i) and (ii) of Definition 1. Hence $c^{-i}(\cdot)$ is a well-defined consistent updating system. By construction, for all $k=0, ..., n-1$, if $R^{-i}(k)$ reaches h, then $c^{-i}(h) \in \Delta^*(R^{-i}(k))$.

Since $c^{-i}(\cdot)$ meets the conditions stated in Definition 2 (a), $s^i \in R^i(n)$ if s^i is a weakly sequential best response to $c^{-i}(\cdot)$. By (n) and $(n-1)$, this is not true only if there exists an information set $h \in H^i$ reached by s^i and $R^{-i}(n-1)$, and a strategy t^i which *does not belong* to $R^i(n-1)$ such that s^i/t_h^i is better than s^i against $c^{-i}(h)$. In this case s_i would be only a *constrained* best response to $c^{-i}(h)$ at h and *no* unconstrained best response would belong to $R^i(n-1)$. Therefore it is sufficient to show that the following is true:

Claim. *There exists a strategy $t^i \in R^i(n-1)$ such that for all $h \in H^i$ reached by s^i, s^i/t_h^i is a best response to $c^{-i}(h)$ at h.*

Proof of the Claim (Sketch). Exploit the arborescence structure and finiteness of H^i and construct t^i by dynamic programming given the probabilities implied by $c^{-i}(\cdot)$ (as i has perfect recall and $c^{-i}(\cdot)$ is consistent, $(H^i, c^{-i}(\cdot))$ corresponds to a decision tree for i). For all $h \in H^i$ (in particular, all the information sets for player i reached either by t^i or by s^i) the continuation strategy t_h^i is a best response against $c^{-i}(h)$. Thus $t^i \in R^i(n) \subseteq R^i(n-1)$ and s^i/t_h^i is a best response to $c^{-i}(h)$ at h for all $h \in H^i$ reached by s^i. ∎

Proof of Theorem 2

Consider a perfect information game Γ with extensive form E satisfying (IPI). Let $\{W_{\mathrm{p}}(n)\}$ be the dominance procedure on Γ. Consider the residual tree $(X(n) \cup Z(n), <)$ of nodes reached by strategy profiles in

$W(n)$ (that is, $Z(n) := \zeta(W_p(n))$ and $X(n)$ is the set of predecessors of $Z(n)$) and let $\Gamma(n)$ be the corresponding *residual game tree* with strategy sets $(S^i(n))$. It must be shown that the residual game tree $\Gamma(n)$ has the same reduced normal form as the *residual strategic game* $(W_p(n), (u^i \circ \zeta))$.

Note that each strategy $\hat{s}^i \in W^i_p(n)$ of the residual strategic game corresponds to a function from X^i to the set of actions, whereas each strategy $s^i \in S^i(n)$ of the residual game tree is a function with domain $X^i(n) := X^i \cap X(n)$, the residual set of i's decision nodes. (If $X^i(n) = \emptyset$, then player i is necessarily redundant in the residual strategic game and can be ignored.) It is sufficient to prove that for each $s^i \in S^i(n)$ there is some $\hat{s}^i \in W^i_p(n)$ such that s^i is the restriction of $\hat{s}^i$ on $X^i(n)$. The proof is constructive. Let $v_1, v_2, ...$ be nodes in X^i without predecessors in X^i. It is shown that for each v_k there is some n-undominated strategy $\hat{s}^i_k \in W^i_p(n)$ which coincides with s^i on the intersection between $X^i(n)$ and the subtree with root v_k (if $v_k \in X^i(n)$). Then $\hat{s}_i$ is obtained as the strategy which coincides with $\hat{s}^i_k$ on the subtree with root v_k for each $k = 1, 2, ...$ The reader can verify that $\hat{s}_i \in W^i_p(n)$.

Recall the property IPI states that for each pair of decision nodes of player i on distinct paths *at least one* out three conditions (a), (b) or (c) holds. But to convey more easily the intuition of the proof, the special case in which condition (a) of IPI always holds is considered first. In this case X^i can be partitioned in m subsets $X^i_1, ..., X^i_m$ such that each X^i_k is totally ordered (i.e., contained in a single path). Let $v_1, ..., v_m$ be the least elements of these m subsets. Since the residual set $X(n)$ is a tree and for this particular player i $X^i(n) \neq \emptyset$, for at least one k, $v_k \in X^i(n)$. Assume w.l.o.g. that $v_1, ..., v_l \in X^i(n)$, $l \leqslant m$. Obviously every node $v_1, ..., v_l$ is reached by any strategy of player i in the reduced game tree $\Gamma(n)$.

Fix one such strategy $s^i \in S^i(n)$. For each $k = 1, ..., l$ there is a terminal node in the residual game tree which follows v_k and is reached by s^i. Let $z_k \in Z(n)$ denote this node. By definition there must be an n-undominated profile $(t^i_k, t^{-i}_k) \in W_p(n)$ reaching z_k. By construction, $t^i_k \in W^i_p(n)$ reaches the same nodes as s^i on the reduced subtree with root v_k and takes the same actions at such nodes. Therefore there is a realization equivalent strategy $\hat{s}^i_k \in W^i_p(n)$ which coincides with s^i on the subtree with root v_k. The desired n-undominated strategy $\hat{s}^i \in W^i_p(n)$ is thus obtained as follows: $\hat{s}^i(x) = \hat{s}^i_k(x)$ for $x \in X^i_k$, otherwise $\hat{s}^i(x) = t^i(x)$, where $t^i \in W^i_p(n)$ is some n-undominated strategy.

This construction can be extended to the general case where either (a) or (b) or (c) hold. Consider the set $X^i(n, s^i)$ of decision nodes for player i in the residual game tree which are reached by s^i and are followed by some opponents' decision nodes, that is, $X^i(n, s^i) := \{x \in X^i(n) \mid (s^i \text{ reaches } x) \text{ and } (\exists y \in X^{-i}, x < y)\}$. Let $\{x_1, ..., x_l\}$ be the set of maximal elements of $X^i(n, s^i)$ (with respect to the precedence relation $<$). For each $k = 1, ..., l$, let v_k be the least element of $\{x \in X^i(n) \mid x \leqslant x_k\}$.

CLAIM. *If $j \neq k$, v_j and v_k are distinct nodes on distinct paths.*

Proof of the Claim. By definition of $\{x_1, ..., x_l\}$, x_j and x_k $(j \neq k)$ are nodes on distinct paths such that (b) and (c) of (IPI) do not hold. Therefore (a) must hold, and this implies that v_j and v_k are distinct nodes on distinct paths. The fact that (c) does not hold is obvious. By way of contradiction, assume that (b) holds; i.e., the last common predecessor of x_j and x_k is a node $x \in X^i$. Let $a_j \in A(x)$ and $a_k \in A(x)$ be the actions preceding x_j and x_k, respectively. Clearly, $a_j \neq a_k$, $x \in X^i(n)$, and s^i is defined at x. Since s^i reaches x_j and x_k, $a_j = s^i(x) = a_k$, which contradicts $a_j \neq a_k$.

Note that—by definition of the residual game tree and $S^i(n)$—starting at every node $x \in X^i(n)$ reached by s^i which is not followed by opponents' decision nodes, s^i must choose an optimal sequence of actions. Taking this into account, for each $k = 1, ..., l$, one can repeat essentially the same construction as above and find some $\hat{s}^i_k \in W^i_{\mathrm{p}}(n)$ choosing the same action as s^i at every node $x \in X^i(n)$ such that $v_k \leqslant x$. The desired strategy $\hat{s}^i \in W^i_{\mathrm{p}}(n)$ is then obtained as follows: $\hat{s}^i(x) = \hat{s}^i_k(x)$ for $x \in X^i$, $v_k \leqslant x$, $k = 1, ..., l$; $\hat{s}^i(x) = s^i(x)$ for $x \in X^i(n) \backslash (\bigcup_{k=1}^{k=l} \{y \mid v_k \leqslant y\})$; otherwise $\hat{s}^i(x) = t^i(x)$, where $t^i \in W^i_{\mathrm{p}}(n)$ is some n-undominated strategy. ∎

Proof of Theorem 3

LEMMA 1. *In every game of perfect information correlated and uncorrelated EFR coincide, i.e., for all n, $P_c(n) = P(n)$.*

Proof. By definition, $P(0) = S = P_c(0)$. Assume that $P(n) = P_c(n)$. By inspection of Definition 3, $P(n+1) \subseteq P_c(n+1)$. Assume that s^i is a best response at h against a correlated conjecture $c^{-i} \in \Delta(P^{-i}(n))$ reaching h. Now coalesce i's opponents into a unique player $(-i)$. By perfect information, $(-i)$ has perfect recall. Therefore there exists a behavioral strategy π^{-i} which is realization-equivalent to c^{-i} (Kuhn [14, Theorem 4]). π^{-i} may be partitioned into its single player components: $\pi^{-i} = (\pi^j)_{j \neq i}$ and each π^j is realization-equivalent to a mixed strategy $\mu^j \in \Delta(P^j(n))$. Let μ^{-i} be the product measure obtained by $(\mu^j)_{j \neq i}$. Then s^i is a best response at h against the uncorrelated conjecture $\mu^{-i} \in \Delta^*(P^{-i}(n))$ reaching h. Hence $P_c(n+1) \subseteq P(n+1)$. ∎

Let $\Delta^o(Y)$ denote the set of full-support probability measures on Y.

RESULT 1. *s^i is inferior in $B^i \times B^{-i}$ if and only if, for all $\mu^{-i} \in \Delta^o(B^{-i})$, there exists $t^i \in B^i$ such that $U^i(t^i, \mu^{-i}) > U^i(s^i, \mu^{-i})$. (See, e.g., Pearce [19, Lemma 4].)*

LEMMA 2. *If s^i is not inferior in $P_c(n)$, then $s^i \in P_c(n+1)$.*

Proof. If s^i is not inferior in $P_c(n)$, then by Result 1, s^i is a best response to some full-support conjecture $\mu^{-i} \in \Delta^o(P_c^{-i}(n))$. Since μ^{-i} and $P_c^{-i}(n)$ reach the same information sets, there exists a consistent updating system $c^{-i}(\cdot)$ such that $c^{-i}(h) = \mu^{-i}$ for all $h \in H^i$ reached by $P_c^{-i}(n)$, and s^i must be a best response at h to $c^{-i}(h)$ for all such h. Therefore $s^i \in P_c^i(n+1)$. ▮

Lemma 3. *In every game without relevant ties, if $s^i \in P_c^i(n+1)$ then s^i is not dominated in $P_c(n)$.*

Proof. Assume s^i is dominated by t^i on $P_c^{-i}(n)$; i.e.,

$$\forall s^{-i} \in P_c^{-i}(n),\ U^i(s^i, s^{-i}) \leqslant U^i(t^i, s^{-i}) \tag{1}$$

$$\exists t^{-i} \in P_c^{-i}(n),\ U^i(s^i, t^{-i}) < U^i(t^i, t^{-i}). \tag{2}$$

By (2) there must be an information set $h \in H^i$ reached by s^i, t^i and $P_c^{-i}(n)$ such that s^i and t^i prescribe different actions at h; otherwise s^i and t^i would be payoff-equivalent on $P_c^{-i}(n)$ (Kuhn [14, Theorem 1]). Then for every s^{-i} reaching h, $\zeta(s^i, s^{-i}) \neq \zeta(t^i, s^{-i})$. By Remark 2, since there are no relevant ties,

$$U^i(s^i, s^{-i}) = u^i[\zeta(s^i, s^{-i})] \neq u^i[\zeta(t^i, s^{-i})] = U^i(t^i, s^{-i}). \tag{3}$$

By (1) and (3), for all $s^{-i} \in P_c^{-i}(n)$ reaching h $U^i(s^i, s^{-i}) < U^i(t^i, s^{-i})$. Therefore s^i cannot be a best response at h against any conjecture $c^{-i} \in \Delta(P_c^{-i}(n))$ reaching h and it does not belong to $P_c^{-i}(n+1)$. ▮

Note that here, it is crucial that there are no chance moves, thus allowing us to write the outcome function ζ as a deterministic function, which is essential to derive (3). However, Remark 2 suggests how to extend the result to games with chance moves. Let $\zeta(\cdot \mid s) \in \Delta(Z)$ be the outcome distribution induced by a strategy profile s and let us say that a game is without relevant ties if, for all $i \in I$, $s^i, t^i \in S^i$, $s^{-i} \in S^{-i}$, $\zeta(\cdot \mid s^i, s^{-i}) \neq \zeta(\cdot \mid t^i, s^{-i})$ implies $U^i(s) \neq U^i(t)$. Then Lemma 3 still holds.

Lemma 4. *In every game of perfect information inferior strategies and dominated strategies coincide, i.e., $W(1) = W_{\mathrm{p}}(1)$.*

Proof. It is trivially true that $W(1) \subseteq W_{\mathrm{p}}(1)$. Therefore it is sufficient to prove that $S \backslash W(1) \subseteq S \backslash W_{\mathrm{p}}(1)$.

Let $s^i \in S^i$ be dominated by some mixed strategy $\mu^i \in \Delta(S^i)$. By perfect information, the conjectures in $\Delta(S^{-i})$ are realization-equivalent to some behavioral profiles in Π^{-i} (cf., Proof of Lemma 1). Let $\Pi^{-i}(k)$ be the compact set of the behavioral profiles assigning to each action a probability of at least $1/k$. $\Pi^{-i}(k)$ is nonempty if the integer k is large enough. Consider the following programs:

$$\max_{t^i \in S^i} \{ \max_{\pi^{-i} \in \Pi^{-i}(k)} [U^i(t^i, \pi^{-i}) - U^i(s^i, \pi^{-i})]\} \tag{4}$$

$$\min_{\pi^{-i} \in \Pi^{-i}(k)} \{\max_{t^i \in S^i} [U^i(t^i, \pi^{-i}) - U^i(s^i, \pi^{-i})\}. \tag{5}$$

By perfect information, dynamic programming can be used to solve (4) and (5), showing that they are equivalent programs (attach to each terminal node $z = \zeta(t^i, t^{-i})$ the payoff $\{u^i(z) - u^i[\zeta(s^i, t^{-i})]\}$ and work backwards). Let $\{(r^i_k, \pi^{-i}_k)\}$ be a sequence of solutions and fix a cluster point (r^i, π^{-i}). By a standard continuity argument, (r^i, π^{-i}) is also a solution of the programs obtained by (4) and (5) replacing $\Pi^{-i}(k)$ with Π^{-i}. Since S^i is a finite set and the second term of the objective function is independent of t^i, r^i must also be a best reply to some strictly positive profiles π^{-i}_k, corresponding to an equivalent conjecture in $\Delta^o(S^{-i})$. By Result 1, $r^i \in W^i(1)$. By assumption, $s^i \notin W^i(1)$. Therefore:

$$\exists s^{-i} \in S^{-i},\ U^i(r^i, s^{-i}) \neq U^i(s^i, s^{-i}). \tag{6}$$

The min-max inequalities and the fact that μ^i dominates s^i yield:

$$\forall t^{-i} \in S^{-i},\ U^i(r^i, t^{-i}) - U^i(s^i, t^{-i}) \geqslant U^i(r^i, \pi^{-i}) - U^i(s^i, \pi^{-i}) \geqslant U^i(\mu^i, \pi^{-i}) - U^i(s^i, \pi^{-i}) \geqslant 0. \tag{7}$$

By (6) and (7), r^i dominates s^i. ▮

Proof of Theorem 3. By definition, $W(0) = P_c(0) = P(0) = W_p(0) = S$. Assume that $W(n) = P_c(n) = P(n) = W_p(n)$. Then Lemmas 1, 2, and 3 yield $W(n+1) \subseteq P_c(n+1) = P(n+1) \subseteq W_p(n+1)$. Since the extensive form has iterated perfect information, there is a perfect information game $\Gamma(n)$ with the same reduced normal form as $(I, W(n), (U^i))$. Applying Lemma 4 to $\Gamma(n)$, one obtains $W(n+1) = W_p(n+1)$. Therefore $W(n+1) = P_c(n+1) = P(n+1) = W_p(n+1)$. ▮

Proof of Theorem 4

LEMMA 5. *In every game of perfect information without relevant ties the sets* $\zeta(\bigcap_{n=0}^{\infty} W(n))$, $\zeta(\bigcap_{n=0}^{\infty} W_p(n))$, $\zeta(\bigcap_{n=0}^{\infty} P_c(n))$ *are singletons.*

Proof. Assume that for some n, $\#\zeta(S(n)) > 1$, where $S(n) = W(n)$ or $S(n) = W_p(n)$ or $S(n) = P_c(n)$ and $\# Y$ is the cardinality of set Y. It will be shown that $S(n+1) \subset S(n)$ (strict containment). Since S is finite and $\bigcap_{n=0}^{\infty} S(n)$ is not empty, this implies that $\zeta(\bigcap_{n=0}^{\infty} S(n))$ must be a singleton.

Let $\hat{X}(n) := \{x \in X \mid \#(Z(x) \cap \zeta(S(n)) > 1\}$. By assumption $\hat{X}(n)$ is not empty (it includes the root). Let $x \in X^i$ be a maximal element of $\hat{X}(n)$ (with respect to the precedence relation) and let z' and z'' be distinct terminal successors of x in $\zeta(S(n))$, i.e., $Z(x) \cap \zeta(S(n)) \supseteq \{z', z''\}$, $z' \neq z''$. Since x is maximal in $\hat{X}(n)$, x must be the last common predecessor of z' and z'' and we may assume w.l.o.g. that $u^i(z') < u^i(z'')$ (there is no relevant tie). Let $s' \in S^i(n)$ reach z' and $s'' \in S^i(n)$ reach z''. Clearly both s' and s'' reach x. Modify s'' so that it coincides with s' outside the subgame with root x, obtaining s'/s''_x. Then s' is weakly dominated by s'/s'' on $S^i \times S^{-i}(n)$. Indeed, for all s^i not reaching x, $\zeta(s', s^{-i}) = \zeta(s'/s''_x, s^{-i})$; for all $s^{-i} \in S^{-i}(n)$ reaching x, since x is maximal in $\hat{X}(n)$ it must be the case that

$$u^i[\zeta(s', s^{-i})] = u^i(z') < u^i(z'') = u^i[\zeta(s'/s''_x, s^{-i})].$$

Note that the dominating strategy s'/s''_x might be outside $S^i(n)$. Therefore one cannot immediately conclude that $s' \in S^i(n) \backslash S^i(n+1)$ for $\{S(k)\} = \{W(k)\}$ or $\{S(k)\} = \{W_{\mathrm{p}}(k)\}$. But it is well known that in every finite game if a strategy s^i is dominated (hence also inferior) on $S^i \times W_{\mathrm{p}}^{-i}(n)$ (or $S^i \times W^{-i}(n)$), then it is also dominated on $W_{\mathrm{p}}(n)$ (or inferior on $W(n)$) (see, e.g., Stahl [28] or Hammond [11]). Analogously, Theorem 1 implies that $s' \in P^i_c(n) \backslash P^i_c(n+1) = R^i_n(n) \backslash R^i_c(n+1)$, because s' is not a best reply at x against any conjecture $c^{-i} \in \Delta(P_c^{-i}(n))$ reaching x. Thus, if $\zeta(S(n))$ has at least two elements, $S(n+1) \subset S(n)$. ∎

LEMMA 6. *In every game of perfect information without relevant ties every mixed equilibrium in the same connected component of Nash equilibria induces the same degenerate probability distribution on terminal nodes.*

Proof. It is sufficient to show that every Nash equilibrium induces a unique terminal node. Let σ be a Nash equilibrium of a game with perfect information and suppose that $\#\zeta(\mathrm{Supp}(\sigma)) \geqslant 2$. It will be shown that the game has a relevant tie. Choose x, z', z'', a' and a'' as in the proof of Lemma 5. That is, x is a maximal element of the non-empty set $\hat{X}(\sigma) := \{x \in X \mid \#(Z(x) \cap \zeta(\mathrm{Supp}(\sigma)) > 1\}$, $z', z'' \in \zeta(\mathrm{Supp}(\sigma))$ are two distinct terminal followers of x (which is their last common predecessor) following actions a' and a'', respectively. Let x be a node of player i. As $\{x\}$ is an information set on the equilibrium path and the continuations after a' and a'' are deterministic (by maximality of x in $\hat{X}(\sigma)$), the equilibrium indifference condition implies that $u^i(z') = u^i(z'')$. Thus the game has a relevant tie. ∎

Proof of Theorem 4. The proof relies on some results about Kohlberg and Mertens' "full stability" [12, p. 1024].

RESULT 2. *Every fully stable set F of a (finite normal form) game G contains a fully stable set of the game G' obtained from G by deletion of a set of pure strategies weakly dominated by mixed strategies for a set of players.* (*This is a generalization of* [12, *Proposition* 2], *which follows from an analogous proof.*)

RESULT 3. *Every game has a fully stable set contained in a single connected component of Nash equilibria.* (*See* [12, pp. 1022–1024].)

RESULT 4. *Every fully stable set contains a proper equilibrium, which induces the same distribution over terminal nodes as some subgame perfect equilibrium.* (*See* [12, *Propositions* 0 *and* 5].)

Consider a perfect information game without relevant ties. Then Lemmas 1, 2, and 5 and Result 2 imply that $\{P(n)\}$, $\{P_c(n)\}$, $\{W_p(n)\}$, and $\{W(n)\}$ are iterative dominance procedures selecting a terminal node consistent with some fully stable set F. Lemma 6 and Results 3 and 4 imply that there is a unique fully stable set F^* such that every (mixed) equilibrium $\sigma \in F$ induces the unique backward induction outcome z^*. ∎

REFERENCES

1. K. Basu, Strategic irrationality in extensive games, *Math. Soc. Sciences* **15** (1988), 247–260.
2. P. Battigalli, Strategic rationality orderings and the best rationalization principle, *Games Econ. Behav.* **13** (1996), 178–200.
3. D. Bernheim, Rationalizable strategic behavior, *Econometrica* **52** (1984), 1007–1028.
4. K. Binmore, Modeling rational players, I, *Econ. Philos.* **3** (1987), 179–214.
5. G. Bonanno, The logic of rational play in games of perfect information, *Econ. Philos.* **7** (1991), 37–65.
6. T. Borgers, Pure strategy dominance, *Econometrica* **61** (1993), 423–430.
7. E. van Damme, "Stability and Perfection of Nash Equilibria," 2nd ed., Springer-Verlag, Berlin, 1991.
8. D. Fudenberg, D. Kreps, and D. Levine, On the robustness of equilibrium refinements, *J. Econ. Theory* **44** (1988), 354–380.
9. D. Fudenberg and J. Tirole, "Game Theory," MIT Press, Cambridge, MA, 1991.
10. R. Gretlein, Dominance elimination procedures on finite alternative games, *Int. J. Game Theory* **12** (1983), 107–113.
11. P. Hammond, Aspects of rationalizable behavior, *in* "Frontiers of Game Theory" (K. Binmore, A. Kirman, and P. Tani, Eds.), MIT Press, Cambridge, MA, 1994.
12. E. Kohlberg and J.-F. Mertens, On the strategic stability of equilibria, *Econometrica* **54** (1986), 1003–1039.
13. D. Kreps and G. Ramey, Structural consistency, consistency and sequential rationality, *Econometrica* **55** (1987), 1331–1348.
14. H. W. Kuhn, Extensive games and the problem of information, *in* "Contribution to the Theory of Games" (Kuhn and Tucker, Eds.), Vol. 2, Princeton Univ. Press, Princeton, NJ, 1953.

15. L. Marx and J. Swinkels, "Order Independence for Iterated Weak Dominance," *Games Econ. Behav.*, in press.
16. J. Moore and R. Repullo, Subgame perfect implementation, *Econometrica* **56** (1988), 1191–1220.
17. H. Moulin, Dominance solvable voting schemes, *Econometrica* **47** (1979), 1351–1373.
18. R. Myerson, "Game Theory: Analysis of Conflict," Harvard Univ. Press, Cambridge, MA, 1991.
19. D. Pearce, Rationalizable strategic behavior and the problem of perfection, *Econometrica* **52** (1984), 1029–1050.
20. P. Reny, "Rationality, Common Knowledge and the Theory of Games," Ph.D thesis, Department of Economics, Princeton University, 1985.
21. P. Reny, Backward induction, normal form perfection and explicable equilibria, *Econometrica* **60** (1992), 627–649.
22. P. Reny, Common belief and the theory of games with perfect information, *J. Econ. Theory* **59** (1993), 257–274.
23. R. Rosenthal, Games of perfect information, predatory pricing and the chain-store paradox, *J. Econ. Theory* **25** (1981), 92–100.
24. A. Rubinstein, Comments on the interpretation of game theory, *Econometrica* **59** (1991), 909–924.
25. A. Rubinstein and M. Osborne, "A Course in Game Theory," MIT Press, Cambridge, MA, 1994.
26. R. Selten, The chainstore paradox, *Theory Decision* **9** (1978), 127–159.
27. J. Sobel, L. Stole, and I. Zapater, Fixed point rationalizability, *J. Econ. Theory* **50** (1990), 304–331.
28. D. Stahl, Lexicographic rationality, common knowledge and iterated admissibility, unpublished paper, Department of Economics, University of Texas, Austin, July 1991.
29. F. Thompson, "Equivalence of Games in Extensive Form," RM 759, The Rand Corporation, 1952.