Theoretical Economics 8 (2013), 125–163 1555-7561/20130125

# The order independence of iterated dominance in extensive games

JING CHEN
Institute for Advanced Study, Princeton and Department of Computer Science, Stony Brook University

SILVIO MICALI
Department of Electrical Engineering and Computer Science, MIT

Shimoji and Watson (1998) prove that a strategy of an extensive game is rationalizable in the sense of Pearce if and only if it survives the maximal elimination of conditionally dominated strategies. Briefly, this process iteratively eliminates conditionally dominated strategies according to a specific order, which is also the start of an order of elimination of weakly dominated strategies. Since the final set of possible payoff profiles, or terminal nodes, surviving iterated elimination of weakly dominated strategies may be order-dependent, one may suspect that the same holds for conditional dominance.

We prove that, although the sets of strategy profiles surviving two arbitrary elimination orders of conditional dominance may be very different from each other, they are equivalent in the following sense: for each player $i$ and each pair of elimination orders, there exists a function $\phi_i$ mapping each strategy of $i$ surviving the first order to a strategy of $i$ surviving the second order, such that, for every strategy profile $s$ surviving the first order, the profile $(\phi_i(s_i))_i$ induces the same *terminal node* as $s$ does.

To prove our results, we put forward a new notion of dominance and an elementary characterization of extensive-form rationalizability (EFR) that may be of independent interest. We also establish connections between EFR and other existing iterated dominance procedures, using our notion of dominance and our characterization of EFR.

KEYWORDS. Extensive-form rationalizability, dominance, iterative elimination, equivalence.

JEL CLASSIFICATION. C72, C73.

---

Jing Chen: jingchen@cs.stonybrook.edu
Silvio Micali: silvio@csail.mit.edu
We would like to thank David Easley and Matt Jackson for helpful discussions in earlier stages of this work, and Robert Aumann and Makoto Shimoji for providing additional references. Finally, we would like to acknowledge the editor and a superb set of anonymous referees for their excellent suggestions. (Any remaining lack of clarity is solely due to the authors!) This work is supported in part by ONR Grant N00014-09-1-0597.


DOI: 10.3982/TE942

## 1. Introduction

The notion of rationalizability was put forward by Bernheim (1984) and Pearce (1984) for normal-form games. The extension of this notion to extensive games, *extensive-form rationalizability* (EFR), was initially proposed by Pearce (1984) and then clarified by Battigalli (1997).

Rationalizable and extensive-form rationalizable strategies (EFR strategies) possess algorithmic characterizations. For normal-form games, if each player is allowed to believe that the other players' strategies are correlated, then a player's strategy is rationalizable if and only if it survives the iterated elimination of *strictly dominated* strategies. It is well known (see, e.g., the proofs in Gilboa et al. 1990 and Osborne and Rubinstein 1994) that the order of elimination is irrelevant: no matter which order is used, the surviving strategies are the same.[1]

For extensive games, the situation is more complex. EFR strategies by definition are strategies surviving the process of *maximal (iterated) elimination*. According to this process, at each step, *all* strategies that are "never a best response" (to the currently surviving ones) are *simultaneously* eliminated. The process stops when no such strategy can be found. Assuming (as we do) perfect recall, Shimoji and Watson (1998) prove that the EFR strategies can be obtained by the maximal elimination of *conditionally dominated strategies*, whose definition is recalled in Section 5.

However, the maximal elimination order is not the only meaningful one,[2] and different elimination orders of conditionally dominated strategies often yield vastly different sets of surviving strategies. Nonetheless, we show that all such sets are equivalent in a very strong sense. We prove this equivalence in two steps. First, we establish a connection between conditional dominance and a new, auxiliary notion, *distinguishable dominance*. Then we prove an order-independence result for distinguishable dominance.

### *A bridge lemma between distinguishable and conditional dominance*

Our notion of distinguishable dominance can be summarized as follows. For every profile $s$ and every subset $I$ of the players, call $(s_i)_{i\in I}$ a *subprofile*, and more simply denote it by $s_I$. Then a (pure) strategy $a$ of player $i$ is *distinguishably dominated* by another (possibly mixed) strategy $b$ of $i$ if the following conditions hold.

(a) There exist strategy subprofiles $s_{-i}$ distinguishing $a$ and $b$, that is, the (distributions of) terminal nodes reached by $(a, s_{-i})$ and $(b, s_{-i})$ do not coincide.

(b) For every subprofile $s_{-i}$ distinguishing $a$ and $b$, $i$'s (expected) payoff is smaller for $(a, s_{-i})$ than for $(b, s_{-i})$.

[1] We always consider finite games in this paper. But it is worth mentioning that for infinite games, the order of iterated elimination of strictly dominated strategies may matter, as shown by Dufwenberg and Stegeman (2002).

[2] For instance, in some extensive games, backward induction may be an elimination order of conditionally dominated strategies that is not maximal, as will be shown in Example 2.

We prove that each elimination order of distinguishable dominance is also an elimination order of conditional dominance and vice versa. This bridge lemma leads to an alternative characterization of EFR and enables us to extend our order-independence theorem to conditional dominance as well.

Distinguishable dominance is formally presented in Section 4 and the bridge lemma is formalized in Section 5.

*Our order-independence theorem*

We denote by $\mathbb{ER}_i$ the set of EFR strategies of player $i$ and denote by $\mathbb{ER}$ the Cartesian product $\times_{i\in N}\mathbb{ER}_i$, where $N = \{1, \dots, n\}$ is the set of players.

In extensive games, whether using conditional or distinguishable dominance, different orders of elimination yield different sets of surviving strategy profiles. We prove, however, that all such sets are equivalent to each other, and thus (via our bridge lemma) to $\mathbb{ER}$, in a very strong sense. This is best explained by considering—for simplicity only—a product set $R$ of surviving strategy profiles such that the cardinality of each $R_i$ equals that of $\mathbb{ER}_i$. In this case there exists a profile $\phi$ (depending on $R$ and $\mathbb{ER}$) of functions such that

1. each $\phi_i$ is a bijection between $\mathbb{ER}_i$ and $R_i$
2. for each profile $s \in \mathbb{ER}$, both $s$ and $\phi(s) \triangleq (\phi_i(s_i))_{i\in N}$ yield the same terminal node (which of course implies that $s$ and $\phi(s)$ are payoff-equivalent).

Accordingly, the players are totally indifferent between an execution of $s$ and an execution of $\phi(s)$. (This implies that if the game is one of imperfect information, then each player sees the same sequence of information sets.) In other words, although the sets $\mathbb{ER}$ and $R$ may consist of very different strategy profiles, when considering the terminal nodes induced by them, it is as if they consisted of the same strategy profiles.

Our order-independence theorem and our bridge lemma together establish that the iterated elimination of conditionally or distinguishably dominated strategies is essentially as order-independent as that of strictly dominated strategies. Not only do these results make finding EFR outcomes easier, but they also show that EFR is actually a tighter and less arbitrary concept than previously thought.

Our main theorem is presented in Section 6. A more general version of it is presented in Section 7.

## 2. Connections with other works

*A new connection between EFR and nice weak dominance*

Our results help establish connections between EFR and other existing solution concepts. For instance, Marx and Swinkels (1997) define *nice weak dominance* and prove that the iterated elimination of nicely weakly dominated strategies is order-independent, up to payoff equivalence. We note that (i) distinguishable dominance and nice weak dominance coincide in games with generic payoffs, and (ii) distinguishable dominance always implies nice weak dominance. Because different orders of iterated elimination of distinguishably dominated strategies yield the same set of histories, they also yield the same set of payoff profiles. Thus, taken together, our bridge lemma and the result of Marx and Swinkels (1997) imply that the set of payoff profiles generated by EFR strategies always contains the set of payoff profiles generated by iterated elimination of nicely weakly dominated strategies. We flesh out this implication in Section 8. It is also easy to see that this containment can be strict for some games.

Marx and Swinkels (1997) also identify a condition—the *transference of decision-maker indifference*$^*$, TDI$^*$, condition—under which nice weak dominance coincides with weak dominance. Therefore, in all games satisfying the TDI$^*$ condition, the set of payoff profiles generated by iterated elimination of weakly dominated strategies is also contained by that generated by EFR strategies.

We note that Brandenburger and Friedenberg (2011) show that in a game satisfying *no relevant convexities*, a condition stronger than TDI$^*$, the set of strategies surviving maximal elimination of weakly dominated strategies coincides with EFR.

*Connection with Apt (2004)*

Apt (2004) provides a unified method for proving order independence for various dominance relations. His approach is clearly related to ours, in the sense that both use basic tools from the literature of abstract reduction systems. The proof of our main order-independence theorem is based on the strong Church–Rosser property, while Apt's main technique is a generalization of Newman's lemma, which relies on the weak Church–Rosser property. We note, however, that Apt did not prove or claim our result, and that our main theorem does not directly follow from his.

*Additional related work*

A lot of previous work is devoted to elimination orders in games with generic payoffs. In particular, Shimoji (2004) provides a proof of order independence for conditional dominance for such games. When the game is, in addition, of perfect information, Gretlein (1983) proves order independence for weak dominance, and Battigalli (1997) proves that EFR and backward induction are history-equivalent. All these results can be viewed as special cases of our work.[3]

Without dealing with different elimination orders, some payoff equivalence is explored by Moulin (1979) for voting games, but, as pointed out by Gretlein (1982), his argument is incomplete. A complete argument is provided by Rochet (1980) and Gretlein (1983).

Also, Robles (2006), using a notion of dominance directly derived from Shimoji and Watson's notion of conditional dominance with strong replacements, explores the same direction we do, but—as he kindly told us—without a satisfactory proof.

[3]In games with generic payoffs, distinguishable dominance and weak dominance coincide, and backward induction is a particular elimination order of distinguishably dominated strategies.

In an expanded version of this paper (Chen and Micali 2012), we further discuss a new connection between EFR and backward induction, and the use of our notion of dominance in mechanism design.

Finally, we wish to acknowledge the epistemic game theory literature on EFR (see, in particular, Battigalli 1997 and Battigalli and Siniscalchi 2002), which provides a conceptual foundation for the solution concepts studied in our work.

## 3. Preliminaries

We consider finite extensive games of complete information with perfect recall and no moves of nature. Such games can be defined via either "collections of terminal histories" or "game trees," and we prefer the latter approach. Recall that a finite directed tree is a connected, directed, acyclic graph where each node has in-degree at most 1. The unique node of in-degree 0 is referred to as the *root* and each node of out-degree 0 as a *leaf*. A node that is not a leaf is referred to as an *internal node*. If there is an edge from node $x$ to node $y$, we refer to $y$ as a *child* of $x$ and to $x$ as the *parent* of $y$.

### *Extensive games*

An extensive game consists of the following components.

- A finite set, $N = \{1, \dots, n\}$, referred to as the *set of players*.
- A finite directed tree, referred to as the *game tree*, with each leaf referred to as a *terminal node* and each internal node as a *decision node*.
- For each decision node $x$,
  - (i) a subset of players, $P(x)$, referred to as the *players (simultaneously) acting*[4] *at x*
  - (ii) for each $i \in P(x)$, a finite set, $A_i(x)$, referred to as the set of *actions available to i at x*
  - (iii) a bijection $\chi_x$ between the set of $x$'s children and the Cartesian product $\times_{i \in P(x)} A_i(x)$.
- For each player $i$, a partition of all decision nodes $x$ for which $i \in P(x)$, $\mathcal{I}_i$, such that if $x, y \in I \in \mathcal{I}_i$, then $A_i(x) = A_i(y)$. If $x \in I \in \mathcal{I}_i$, then we refer to $I$ as an *information set of i* and set $A_i(I) \triangleq A_i(x)$.
- For each player $i$ and each terminal node $z$, a number $u_i(z)$, referred to as *i's payoff at z*.

(Pictorially, a play of an extensive game starts at the root and proceeds in a node-to-child fashion, until a terminal node is reached. Specifically, if, at a decision node $x$,

[4]Traditionally, only one player acts at a decision node, but extensive games with simultaneous moves have also been considered and our results apply to such games as well.

each player $i$ in $P(x)$ chooses an action $a_i$ in $A_i(x)$, then $\chi_x((a_i)_{i\in P(x)})$ is the next node reached.)

*Basic notation*

- The *height* of a node is the number of edges in the longest (directed) path from it to a leaf. (Accordingly, a leaf has height 0.) The height of the game tree is the height of its root.
- A pure strategy $s_i$ of a player $i$ is a function mapping each $I$ in $\mathcal{I}_i$ to an action in $A_i(I)$. If $x \in I \in \mathcal{I}_i$, then we set $s_i(x) \triangleq s_i(I)$. We refer to $s_i(x)$ as the action taken by $i$ at $x$ according to $s_i$.
- We denote the set of all pure strategies of a player $i$ by $S_i$ and set $S \triangleq \times_{i\in N} S_i$.
- If $X$ is a finite set, then $\Delta(X)$ denotes the set of all probability distributions over $X$.
- For each player $i$, a mixed strategy of $i$ is an element in $\Delta(S_i)$. If $\sigma_i \in \Delta(S_i)$ and $s_i \in S_i$, then $\sigma_i(s_i)$ denotes the probability assigned to $s_i$ by $\sigma_i$.
- A strategy or strategy profile is always pure if it is represented by a lowercase Latin letter; it is mixed (maybe degenerated) if it is represented by a lowercase Greek letter.
- Given a pure strategy profile $s$, $u_i(s)$ denotes the payoff of player $i$ at the terminal node determined by $s$. Given a mixed strategy profile $\sigma$, $u_i(\sigma)$ denotes the expected payoff of $i$ induced by $\sigma$.
- For all players $i$ and all (different) information sets $I$ and $I'$ in $\mathcal{I}_i$, $I'$ *follows* $I$ if there exists a decision node $x' \in I'$ such that the path from the root to $x'$ goes through a decision node in $I$.[5]

*Histories*

The *history* of a pure strategy profile $s$ consists of the sequence of nodes in the game tree reached in a play of the game according to $s$. We denote by $H$ the function mapping each pure strategy profile to its history. Thus, following standard conventions, if $X$ is a set of pure strategy profiles, then $H(X) = \{H(s) : s \in X\}$. If $\sigma$ is a mixed strategy profile, then $H(\sigma)$ is the distribution induced by $\sigma$ over the histories of the strategy profiles in the support of $\sigma$.

A pure strategy subprofile $s_P$ *reaches a node* $x$ if there exists a pure strategy subprofile $s_{-P}$ such that $x \in H(s)$, and $s_P$ *reaches an information set* $I$ if there exists a decision node $x \in I$ such that $s_P$ reaches $x$. Letting $I$ be an information set of a player $i$, the set of all pure strategies of $i$ reaching $I$ is denoted by $S_i(I)$, the set of all pure strategy subprofiles of $-i$ reaching $I$ is denoted by $S_{-i}(I)$, and the set of all pure strategy profiles reaching $I$ is denoted by $S(I)$.

[5]Assuming perfect recall (as defined in Osborne and Rubinstein 1994, p. 203), $I'$ follows $I$ implies that for *each* decision node $x' \in I'$, the path from the root to $x'$ goes through a decision node in $I$.

A mixed strategy subprofile $\sigma_P$ *reaches a node* $x$ (respectively, *an information set* $I$) if for every pure strategy subprofile $a_P$ in the support of $\sigma_P$, $a_P$ reaches $x$ (respectively, $I$). If a strategy profile $\sigma$ reaches $x$ (respectively, $I$), we may also say that $H(\sigma)$ reaches $x$ (respectively, $I$)—adding "with probability 1" for emphasis. (Note that reachability by mixed strategies has sometimes been defined differently in the literature.)

### *Two known facts*

We mention without proof the following two facts about histories in extensive games with perfect recall.

FACT 1. For all players $i$, nodes $x$, and pure strategy profiles $s$ and $t$, if both $H(s)$ and $H(t)$ reach $x$, then $H(t_i, s_{-i})$ also reaches $x$.

FACT 2. For all players $i$, information sets $I \in \mathcal{I}_i$, and pure strategy profiles $s$, $H(s)$ reaches $I$ if and only if both $s_i$ and $s_{-i}$ reach $I$. (Thus, $S(I) = S_i(I) \times S_{-i}(I)$.) Moreover, if two strategies $t_i$ and $t_i'$ both reach $I$, then they coincide at every information set of $i$ followed by $I$, and for all strategy subprofiles $t_{-i}$ reaching $I$, $H(t_i, t_{-i})$ and $H(t_i', t_{-i})$ reach the same decision node in $I$.

### *Sets of strategy subprofiles*

Following tradition, when talking about a set of strategy subprofiles $R_J$, we always implicitly mean that $R_J$ is a Cartesian product, $R_J = \times_{j \in J} R_j$. Following tradition again, the only exceptions in this paper are the already defined $S_{-i}(I)$ and $S(I)$, where $I \in \mathcal{I}_i$. (Indeed, although $S(I) = S_i(I) \times S_{-i}(I)$, $S_{-i}(I)$ and thus $S(I)$ may not be Cartesian products.)

## 4. DISTINGUISHABLE DOMINANCE

We break the notion of distinguishable dominance into simpler components.

DEFINITION 1 (Distinguishability and indistinguishability). Let $\sigma_i$ and $\sigma_i'$ be two different strategies of player $i$ and let $R_{-i}$ be a set of pure strategy subprofiles. A strategy subprofile $t_{-i} \in R_{-i}$ *distinguishes* $\sigma_i$ *and* $\sigma_i'$ (over $R_{-i}$) if

$$H(\sigma_i, t_{-i}) \neq H(\sigma_i', t_{-i}).$$

The strategies $\sigma_i$ and $\sigma_i'$ are *distinguishable* over $R_{-i}$ if there exists a strategy subprofile $t_{-i} \in R_{-i}$ that distinguishes them; otherwise, they are *indistinguishable* (over $R_{-i}$).

If $\sigma_i$ and $\sigma_i'$ are distinguishable over $R_{-i}$, we write $\sigma_i \not\simeq \sigma_i'$ over $R_{-i}$ or $\sigma_i \not\simeq_{R_{-i}} \sigma_i'$; otherwise, we write $\sigma_i \simeq \sigma_i'$ over $R_{-i}$ or $\sigma_i \simeq_{R_{-i}} \sigma_i'$.

Notice that indistinguishability is a notion expressing history equivalence and is much stronger than just payoff equivalence.[6] Also notice that in a normal-form game, as long as $R_{-i} \neq \varnothing$, every pair of different strategies of player $i$ is distinguishable over $R_{-i}$.[7]

Definition 2 (Distinguishable dominance). Let $i$ be a player and let $R$ be a set of pure strategy profiles. A strategy $s_i \in S_i$ is *distinguishably dominated* (DD) *by* $\sigma_i \in \Delta(S_i)$ *over* $R_{-i}$, if

(i) $s_i$ and $\sigma_i$ are distinguishable over $R_{-i}$

(ii) $u_i(s_i, t_{-i}) < u_i(\sigma_i, t_{-i})$ for every strategy subprofile $t_{-i} \in R_{-i}$ that distinguishes $s_i$ and $\sigma_i$.

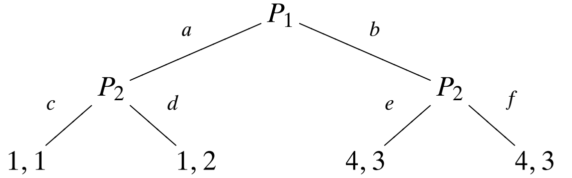
Further, $s_i$ is *distinguishably dominated by* $\sigma_i$ *within* $R$ if $s_i \in R_i$ and $\sigma_i \in \Delta(R_i)$.

We write

- $s_i \prec \sigma_i$ over $R_{-i}$ or $s_i \prec_{R_{-i}} \sigma_i$ if $s_i$ is DD by $\sigma_i$ over $R_{-i}$
- $s_i \preceq \sigma_i$ over $R_{-i}$ or $s_i \preceq_{R_{-i}} \sigma_i$ if either $s_i \simeq_{R_{-i}} \sigma_i$ or $s_i \prec_{R_{-i}} \sigma_i$
- $s_i \prec \sigma_i$ within $R$ or $s_i \prec_R \sigma_i$ if $s_i$ is DD by $\sigma_i$ within $R$.

Notice that $s_i \prec_R \sigma_i$ implies both $s_i \in R_i$ and $\sigma_i \in \Delta(R_i)$, while $s_i \prec_{R_{-i}} \sigma_i$ does not imply any of them. Notice also that $s_i \preceq_{R_{-i}} \sigma_i$ if for all $t_{-i} \in R_{-i}$, either $H(s_i, t_{-i}) = H(\sigma_i, t_{-i})$ or $u_i(s_i, t_{-i}) < u_i(\sigma_i, t_{-i})$.

Example 1. Consider the following game $G_1$.




In $G_1$, any two strategies of $P_2$ are distinguishable over $S_1$. In particular, $ce$ and $de$ are distinguished by $a$: indeed, $H(a, ce) = (a, c) \neq (a, d) = H(a, de)$. However, letting $R_1 = \{b\}$,

[6]Beyond determining (together with the opponents' strategies) a player's payoff, a strategy also determines the terminal node causing that payoff and thus the history of the game. But beyond that, a strategy has no further consequences. The fact that $\sigma_i \simeq_{R_{-i}} \sigma_i'$ thus guarantees that, as long as player $i$ is sure that all other players will choose their strategies from $R_{-i}$, $\sigma_i$ and $\sigma_i'$ are de facto *identical* to him. In concrete terms, if $i$ were far away from the "strategy buttons," but were able to observe the history of the game, and had instructed one of her subordinates to push button $\sigma_i$, while he pushed $\sigma_i'$ instead, then she could not tell the difference at all. Another notion, "outcome equivalence," also appears in the literature. However, sometimes (e.g., Battigalli and Friedenberg 2012) it is defined to mean payoff equivalence and sometimes (e.g., Osborne and Rubinstein 1994) to mean history equivalence. Accordingly, to avoid confusion, we do not use the term "outcome equivalence."

[7]By definition, in a normal-form game, the history of a strategy profile $\sigma$ coincides with $\sigma$ itself, so that any two different strategy profiles have different histories, and thus the notion of distinguishable dominance coincides with strict dominance, and so do their corresponding notions of iterated elimination.

the same strategies $ce$ and $de$ are indistinguishable over $R_1$. Indeed $H(b, ce) = (b, e) = H(b, de)$. Note that strategies $cf$ and $df$ are indistinguishable over $R_1$ too. Game $G_1$ thus illustrates that the notion of distinguishability is indeed dependent on the subprofile of sets of strategies under consideration.

Now turning our attention to distinguishable dominance, note the following details.

- $a$ is distinguishably dominated by $b$ over $S_2$. (Moreover, $a$ is strictly dominated by $b$ in game $G_1$.)
- $ce$ is distinguishably dominated by $de$ over $S_1$: the only strategy in $S_1$ distinguishing them is $a$, and $P_2$'s payoff is 2 under $(a, de)$ and only 1 under $(a, ce)$. (However, $ce$ is not strictly dominated in game $G_1$.)
- $ce$ is not distinguishably dominated by $df$ over $S_1$: although $b$ distinguishes $ce$ and $df$ over $S_1$, $P_2$'s payoff is the same under both $(b, ce)$ and $(b, df)$. (However, $ce$ is weakly dominated by $df$ in game $G_1$.)

Game $G_1$ thus illustrates that the notion of distinguishable dominance is different from both strict dominance and weak dominance. ♢

Definition 3 (Iterated elimination of DD strategies, and resilient solutions). A set of pure strategy profiles $R$ *survives iterated elimination of DD strategies* if there exists a sequence $\mathcal{R} = (R^0, \dots, R^K)$ of sets of strategy profiles such that

(i) $R^0 = S$ and $R^K = R$

(ii) for all $k < K$,

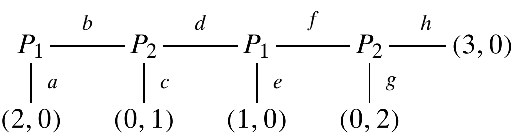
  (a) there is a player $i$ such that $R_i^k \setminus R_i^{k+1} \neq \varnothing$

  (b) for all players $j$, $R_j^{k+1} \subseteq R_j^k$ and every strategy in $R_j^k \setminus R_j^{k+1}$ is DD within $R^k$

(iii) each $R_i^K$ contains no strategy that is DD within $R^K$.

We refer to $\mathcal{R}$ as an *elimination order* of DD strategies and refer to $R$ as a *resilient solution*. Profile $\mathcal{R}$ is *maximal* if for all $k$ and $i$, $R_i^k \setminus R_i^{k+1}$ includes all strategies that are DD within $R^k$.

Example 2. The following game $G_2$, due to Reny (1992), is a classical example for illustrating different elimination orders.



In this game, one resilient solution corresponds to the maximal elimination of distinguishably dominated strategies: namely, $R = \{ae, af\} \times \{dg\}$.[8] Another resilient solution essentially corresponds to backward induction: namely, $T = \{ae, af\} \times \{cg, ch\}$.[9] Yet, notice that both $R$ and $T$ induce the same set of histories (namely, $\{(a)\}$). Our order-independence theorem implies that this is actually true in general. $\diamondsuit$

## 5. Our bridge lemma

Let us recall conditional dominance in our terminology, so as to facilitate a comparison with our notion.

Definition 4 (Conditional dominance). Let $R$ be a set of strategy profiles and let $i$ be a player. A strategy $s_i \in R_i$ is *conditionally dominated within* $R$ if there exists an information set $I \in \mathcal{I}_i$ and a strategy $\sigma_i \in \Delta(R_i)$ satisfying the requirements

(i) $s_i \in S_i(I)$, $\sigma_i \in \Delta(S_i(I))$, and $S_{-i}(I) \cap R_{-i} \neq \varnothing$

(ii) for each $t_{-i} \in S_{-i}(I) \cap R_{-i}$, $u_i(\sigma_i, t_{-i}) > u_i(s_i, t_{-i})$.

Note that iterated elimination, elimination order, and maximal elimination order are defined for conditional dominance exactly as for distinguishable dominance: just replace "distinguishably" with "conditionally" in Definition 3. The set of strategy profiles surviving the maximal elimination order of conditionally dominated strategies coincides with $\mathbb{ER}$, as proven by Shimoji and Watson (1998).

[8] Indeed, $S_1 = \{ae, af, be, bf\}$ and $S_2 = \{cg, ch, dg, dh\}$, and the maximal elimination of DD strategies works as follows:

1. Strategy $be \prec_S ae$ (distinguished by all strategies in $S_2$), $dh \prec_S dg$ (distinguished by $bf$), and nothing else is distinguishably dominated. Therefore, $R_1^1 = \{ae, af, bf\}$ and $R_2^1 = \{cg, ch, dg\}$.
2. Strategy $bf \prec_{R^1} ae$ (distinguished by all strategies in $R_2^1$), $cg \prec_{R^1} dg$ and $ch \prec_{R^1} dg$ (distinguished by $bf$), and nothing else. Therefore, $R_1^2 = \{ae, af\}$ and $R_2^2 = \{dg\}$.
3. No other strategy can be eliminated and thus $R^2$ survives the maximal elimination of DD strategies.

[9] Indeed, a different elimination order of DD strategies is as follows:

1. Strategy $dh$ is eliminated because $dh \prec_S dg$ (distinguished by $bf$). Therefore, $T_1^1 = S_1$ and $T_2^1 = \{cg, ch, dg\}$.
2. Strategy $bf$ is eliminated because $bf \prec_{T^1} be$ (distinguished by $dg$). Therefore, $T_1^2 = \{ae, af, be\}$ and $T_2^2 = T_2^1$.
3. Strategy $dg$ is eliminated because $dg \prec_{T^2} cg$ (distinguished by $be$). Therefore, $T_1^3 = T_1^2$ and $T_2^3 = \{cg, ch\}$.
4. Strategy $be$ is eliminated because $be \prec_{T^3} ae$ (distinguished by $cg$ and $ch$). Therefore, $T_1^4 = \{ae, af\}$ and $T_2^4 = T_2^3$.
5. No other strategy can be eliminated, and thus $T^4$ is a resilient solution.

*Differences between distinguishable and conditional dominance*

The definitions of distinguishable and conditional dominance are of course different. In particular, the notion of conditional dominance requires an additional component: namely, the information set $I$. Further, it allows for the possibility of some "circularity": namely, a pure strategy $s_i$ may be dominated by another pure strategy $s'_i$ within $R$ (relative to an information set $I$), while $s'_i$ is itself dominated by $s_i$ within the *same* $R$ (relative to a different information set $I'$). In this case, both strategies are eliminated simultaneously in the maximal elimination order. However, this circularity is innocuous: it is proved that it does not cause any problem to the notion of EFR. Such a circularity does not arise for distinguishable dominance.

Let us now explain that distinguishable and conditional dominance are indeed different concepts: distinguishable dominance implies conditional dominance, but not vice versa. To begin with, according to Definition 2, when $s_i \prec_{R_{-i}} \sigma_i$, we do not require $s_i \in R_i$ or $\sigma_i \in \Delta(R_i)$. When $s_i \notin R_i$ or $\sigma_i \notin \Delta(R_i)$, distinguishable dominance is quite unrelated to conditional dominance. However, we have the following proposition.

PROPOSITION 1. *For all sets of strategy profiles $R$, all players $i$, and all strategies $s_i$ and $\sigma_i$, $s_i \prec \sigma_i$ within $R$ implies that $s_i$ is conditionally dominated by $\sigma_i$ within $R$.*

PROOF. Because $s_i \not\approx_{R_{-i}} \sigma_i$, there exists $t_{-i} \in R_{-i}$ such that $H(s_i, t_{-i}) \neq H(\sigma_i, t_{-i})$. Considering one by one, starting with the root, the information sets of $i$ reached by $H(s_i, t_{-i})$, let $I$ be the first information set such that there exists $a_i$ in the support of $\sigma_i$ with $a_i(I) \neq s_i(I)$. (Such an $I$ exists, since otherwise $H(s_i, t_{-i}) = H(\sigma_i, t_{-i})$.) By definition, we have

$$s_i \in S_i(I) \quad \text{and} \quad S_{-i}(I) \cap R_{-i} \supseteq \{t_{-i}\} \neq \varnothing.$$

For each information set $I' \in \mathcal{I}_i$ followed by $I$, $H(s_i, t_{-i})$ reaches $I'$, because the game is with perfect recall. By the choice of $I$, for each $a_i$ in the support of $\sigma_i$ we have $a_i(I') = s_i(I')$. Accordingly, $\sigma_i$ coincides with $s_i$ at all information sets of $i$ followed by $I$, which implies that $H(\sigma_i, t_{-i})$ reaches $I$. Thus

$$\sigma_i \in \Delta(S_i(I)),$$

and requirement (i) of Definition 4 holds.

Because $\sigma_i$ and $s_i$ do not coincide at information set $I$, for each $t'_{-i} \in S_{-i}(I) \cap R_{-i}$, we have that $t'_{-i}$ distinguishes $s_i$ and $\sigma_i$, and thus $u_i(\sigma_i, t'_{-i}) > u_i(s_i, t'_{-i})$. Therefore, requirement (ii) of Definition 4 also holds, and $s_i$ is conditionally dominated by $\sigma_i$ within $R$.[10] □

[10]Actually, one can verify that $s_i \prec_R \sigma_i$ if and only if the following two requirements are satisfied:

1. Strategy $s_i$ is conditionally dominated by $\sigma_i$ within $R$.
2. For all $I \in \mathcal{I}_i$ such that
   2.1. $s_i \in S_i(I)$, $\sigma_i \in \Delta(S_i(I))$, $S_{-i}(I) \cap R_{-i} \neq \varnothing$
   2.2. $a_i(I) \neq s_i(I)$ for some $a_i$ in the support of $\sigma_i$,

   $s_i$ is conditionally dominated by $\sigma_i$ within $R$ with respect to $I$.

Let us now provide a simple counterexample proving that

$s_i$ being conditionally dominated by $\sigma_i$ within $R$ does not imply $s_i \prec \sigma_i$ within $R$.

Example 3. In game $G_1$ of Example 1, letting $R = \{a, b\} \times \{cf, de\}$, the strategy $cf$ is conditionally dominated by $de$ within $R$, with the desired information set being the decision node following $a$. However, $cf$ is not distinguishably dominated by $de$, because there exists $s_1 \in R_1$ (namely, strategy $b$) such that $H(s_1, cf) \neq H(s_1, de)$ and $u_2(s_1, cf) = u_2(s_1, de)$. Accordingly, $cf$ is not DD by any strategy in $\Delta(R_2)$ over $R_1$ and thus is not DD within $R$. ♢

Shimoji and Watson (1998) also put forward two variants of conditional dominance. These notions also are different from distinguishable dominance.[11]

*Bridging distinguishable and conditional dominance*

As we have just seen, relative to a particular set of strategy profiles $R$, a strategy may be conditionally dominated but not distinguishably dominated. However, for this to happen, we show that $R$ must be chosen somewhat "arbitrarily." That is, the two different notions of dominance considered here coincide with respect to all "naturally" obtained sets of strategy profiles $R$: namely, the set of all strategy profiles $S$ and all sets derived from $S$ solely by iteratively eliminating some conditionally or distinguishably dominated strategies. Indeed, in Example 3, the set $R = \{a, b\} \times \{cf, de\}$ cannot be obtained from $S$ by such iterated elimination. Let us now be more formal.

Lemma 1 (Bridge lemma). *Each elimination order of conditionally dominated strategies is also an elimination order of DD strategies and vice versa. Moreover, the maximal elimination order of conditionally dominated strategies is also the maximal elimination order of DD strategies.*

[11] The first variant is *conditional dominance by replacements*. For a strategy $s_i$ to be dominated in this sense within some set of strategy profiles $R$ by another strategy $\sigma_i$, not only should it be conditionally dominated by $\sigma_i$ within $R$, as per Definition 4, but $s_i$ and $\sigma_i$ must also be payoff equivalent with respect to each strategy subprofile $s_{-i} \in R_{-i} \setminus S_{-i}(I)$; that is, $(u_j(s_i, s_{-i}))_{j\in N} = (u_j(\sigma_i, s_{-i}))_{j\in N}$. The second variant is *conditional dominance by strong replacements*. For $s_i$ to be dominated in this sense within $R$ by $\sigma_i$, in addition to being conditionally dominated by $\sigma_i$ within $R$, $s_i$ and $\sigma_i$ must be history equivalent with respect to each $s_{-i} \in R_{-i} \setminus S_{-i}(I)$; that is, $H(s_i, s_{-i}) = H(\sigma_i, s_{-i})$. Among all three versions of conditional dominance, the last one is the closest to distinguishable dominance. However, although both consider some form of history equivalence, conditional dominance by strong replacements and distinguishable dominance are different. The former allows $s_i$ and $\sigma_i$ to differ only at one information set $I$ and every information set following $I$, but forces $s_i$ and $\sigma_i$ to coincide at every other information set that is reachable. The latter has no such restriction. In particular, if $s_i$ is distinguishably dominated by $\sigma_i$, then it is very possible that there exist two information sets $I$ and $I'$, neither following the other, such that $s_i$ and $\sigma_i$ differ at both of them and coincide everywhere else. The key idea of all three versions of conditional dominance is that, conditioned on a particular information set being reached, $s_i$ is strictly dominated by $\sigma_i$. By contrast, distinguishable dominance essentially compares $s_i$ and $\sigma_i$ wherever they differ (as reflected by item 2 of footnote 10). In a sense, it is "unconditional dominance."

The proof of Lemma 1 is given in Appendix A. Notice that the vice versa part of Lemma 1 is not necessary for proving that iterated elimination of conditionally dominated strategies is order-independent. But it establishes a closer connection between conditional dominance and distinguishable dominance. With this part, the lemma immediately implies that *the notion of a resilient solution does not depend on which of the two notions of dominance is chosen*. In light of the result of Shimoji and Watson, the second half of the lemma immediately implies the following alternative characterization of EFR.

Corollary 1 ($\mathbb{ER}$ is a resilient solution). *If $R$ is the set of strategy profiles surviving the maximal elimination order of DD strategies, then $R = \mathbb{ER}$.*

The corollary can be illustrated by the same game $G_1$ of Example 1. In this game, the maximal elimination order of DD strategies terminates after a single step, in which the strategies $a$, $ce$, and $cf$ are eliminated. Accordingly, the set of surviving strategy profiles is $\{b\} \times \{de, df\}$, and it is clear that (i) exactly the same set is obtained after one step of maximal elimination of conditionally dominated strategies and (ii) the strategies $b$, $de$, and $df$ are not conditionally dominated.

## 6. Main result

To extend the equivalence relation between strategies induced by the notion of indistinguishability (i.e., $\simeq_{R_{-i}}$ for given $R$ and player $i$) to sets of strategy profiles, we establish a suitable notation that lets us deal with equivalent strategies simultaneously.

*Notation.* If $R$ is a set of pure strategy profiles, then we can make the following definitions.

- The set $R_i^{\simeq_{R_{-i}}}$ denotes the partition of $R_i$ into equivalence classes under the relation $\simeq_{R_{-i}}$, and $R^{\simeq}$ denotes the profile of partitions $(R_1^{\simeq_{R_{-1}}}, \dots, R_n^{\simeq_{R_{-n}}})$.
- For all $s_i \in R_i$, $s_i^{\simeq_{R_{-i}}}$ denotes the equivalence class in $R_i^{\simeq_{R_{-i}}}$ to which $s_i$ belongs.
- For all $s \in R$, $s^{\simeq_R}$ denotes the profile of equivalence classes $(s_1^{\simeq_{R_{-1}}}, \dots, s_n^{\simeq_{R_{-n}}})$.

When the profile $R$ under consideration is clear, we may omit the symbols $R$ and $R_{-i}$ in superscripts, and simply write $R_i^{\simeq}$, $s_i^{\simeq}$, and $s^{\simeq}$.

Let us formally note that the *history of a profile of equivalence classes* is well defined.

Proposition 2. *For all sets of strategy profiles $R$, $s \in R$, and $s' \in s_1^{\simeq} \times \cdots \times s_n^{\simeq}$, we have $H(s') = H(s)$.*

The proof of Proposition 2 is a simple and standard argument: for completeness sake, see Appendix B. According to this proposition, if $R$ is a set of strategy profiles and $s \in R$, then we define $H(s^{\simeq_R})$ to be $H(s)$, without causing any ambiguity.

Definition 5 (Equivalence between sets of strategy profiles). Two sets of strategy profiles $R$ and $T$ are *equivalent* if there exists a profile $\phi$ of functions such that

- each $\phi_i$ is a bijection from $R_i^{\simeq_{R_{-i}}}$ to $T_i^{\simeq_{T_{-i}}}$
- for all strategy profiles $s \in R$, $H(s) = H(\phi_1(s_1^{\simeq_{R_{-1}}}), \dots, \phi_n(s_n^{\simeq_{R_{-n}}}))$.

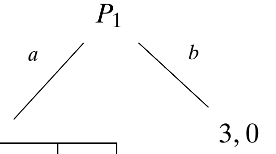
In this case, we further say that $R$ and $T$ are *equivalent under* $\phi$.

Notice that if $R$ and $T$ are equivalent, then $H(R) = H(T)$.

Theorem 1. *Any two sets of strategy profiles surviving iterated elimination of distinguishably dominated strategies are equivalent and thus equivalent to* $\mathbb{ER}$.

The proof of Theorem 1 is given in Appendix C. This theorem establishes a strong connection between EFR and resilient solutions (i.e., sets of strategy profiles surviving iterated elimination of distinguishably dominated strategies). This connection exists even when, as shown by Example 2 and the following example (which is a game with simultaneous moves), a player's strategies in some resilient solution are totally disjoint from his EFR strategies.

Example 4. Consider the following game $G_3$ introduced by Perea (2011).




| $P_1$ \ $P_2$ | $e$ | $f$ | $g$ |
|---|---|---|---|
| $c$ | 2, 2 | 2, 1 | 0, 0 |
| $d$ | 1, 1 | 1, 2 | 4, 0 |

In this game, the decision node following $P_1$'s action $a$ has $P_1$ and $P_2$ acting simultaneously, and is of height 1 (although its children are not explicitly drawn). One resilient solution corresponds to the maximal elimination of distinguishably (and by virtue of Lemma 1, conditionally) dominated strategies: namely, $\mathbb{ER} = \{bc, bd\} \times \{f\}$.[12] Accordingly, the *only* EFR strategy of $P_2$ is $f$. Another resilient solution is $T = \{bc, bd\} \times \{e\}$.[13]

[12] Indeed, $S_1 = \{ac, ad, bc, bd\}$ and $S_2 = \{e, f, g\}$, and the maximal elimination of DD strategies works as follows:

1. Strategy $ac \prec_S bc$ (distinguished by $e$, $f$, and $g$), $g \prec_S e$ (distinguished by $ac$ and $ad$), and nothing else is distinguishably dominated. Therefore, $R_1^1 = \{ad, bc, bd\}$ and $R_2^1 = \{e, f\}$.
2. Strategy $ad \prec_{R^1} bc$ (distinguished by $e$ and $f$), $e \prec_{R^1} f$ (distinguished by $ad$), and nothing else. Therefore, $R_1^2 = \{bc, bd\}$ and $R_2^2 = \{f\}$.
3. No other strategy can be eliminated and thus $R^2$ survives the maximal elimination of DD strategies.

[13] Indeed, a different elimination order of DD strategies is as follows:

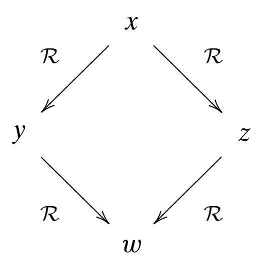
Notice that $\mathbb{ER}$ and $T$ generate the same histories: namely, $H(\mathbb{ER}) = H(T) = \{(b)\}$. In addition, $bc \simeq_{\mathbb{ER}_2} bd$ and $bc \simeq_{T_2} bd$. Thus, at least in this simple game, the profile $\phi$ guaranteed by Theorem 1 can be easily found: $\phi_1(\{bc, bd\}) = \{bc, bd\}$ and $\phi_2(\{f\}) = \{e\}$. Therefore, $\mathbb{ER}$ is equivalent to $T$. ♢

In the above example, the strategies of the unique subgame-perfect equilibrium $(bc, e)$ survive some elimination order. However, Example 7 of Appendix B shows that if a game has multiple subgame-perfect equilibria, then some of their strategies may not survive any elimination order.

### 6.1 *Some intuition behind our proof of Theorem 1*

Our precise line of reasoning is, of course, reflected in the proof itself. However, since the proof is of some complexity, in this subsection, we try to give the reader some (necessarily incomplete) intuition on how we proceed.

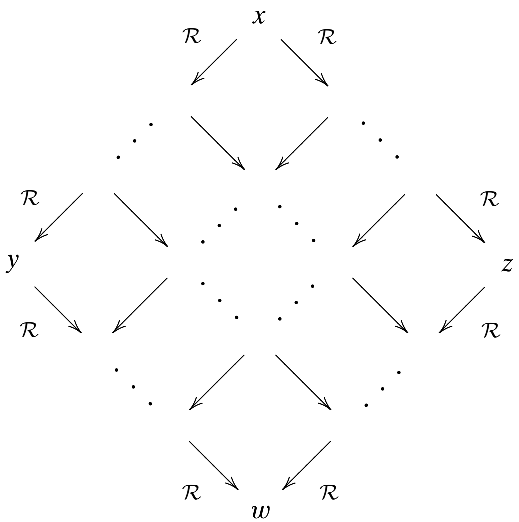
We prove Theorem 1 via the *strong Church–Rosser property* (Church and Rosser 1936), often referred to as the *diamond property*. This property is perhaps the most basic tool in the literature of abstract reduction systems (see, for instance, Klop 1992, Böhm and Micali 1980, and Huet 1980), and is implicitly used in Gilboa et al. (1990). Letting $\mathcal{S}$ be a finite set and $\mathcal{R}$ a binary relation over $\mathcal{S}$, the pair $(\mathcal{S}, \mathcal{R})$ satisfies the diamond property if, for all $x, y, z \in \mathcal{S}$, $x \mathcal{R} y$ and $x \mathcal{R} z$ imply that there exists $w \in \mathcal{S}$ such that $y \mathcal{R} w$ and $z \mathcal{R} w$. Pictorially,




A well known consequence of the diamond property is "unique termination" (in the formal parlance of reduction systems, "unique normal form"). Let $\mathcal{R}^*$ be the reflexive and transitive closure of $\mathcal{R}$. Then, for all $x, y, z \in \mathcal{S}$ such that $x \mathcal{R}^* y$ and $x \mathcal{R}^* z$, if both $y$ and $z$ are "terminal," that is, there does not exist any $w \in \mathcal{S}$ such that either $y \mathcal{R} w$ or

1. Strategy $g$ is eliminated because $g \prec_S e$. Therefore, $T_1^1 = S_1$ and $T_2^1 = \{e, f\}$.
2. Strategy $ad$ is eliminated because $ad \prec_{T^1} ac$ (distinguished by $e$ and $f$). Therefore, $T_1^2 = \{ac, bc, bd\}$ and $T_2^2 = \{e, f\}$.
3. Strategy $f$ is eliminated because $f \prec_{T^2} e$ (distinguished by $ac$). Therefore, $T_1^3 = \{ac, bc, bd\}$ and $T_2^3 = \{e\}$.
4. Strategy $ac$ is eliminated because $ac \prec_{T^3} bc$ (distinguished by $e$). Therefore, $T_1^4 = \{bc, bd\}$ and $T_2^4 = \{e\}$.
5. No other strategy can be eliminated and thus $T^4$ is a resilient solution.

$z \mathcal{R} w$, we have $y = z$. A formal proof can be found, for instance, in Klop (1992), but all the necessary intuition is contained in the following picture.




Now let $\mathcal{S}$ be the set of all sets of pure strategy profiles, and let $\mathcal{R}_{\mathtt{sim}}$ be the binary relation over $\mathcal{S}$ such that, for all $X$ and $Y$ in $\mathcal{S}$, $X \mathcal{R}_{\mathtt{sim}} Y$ if and only if $Y$ can be obtained from $X$ by (simultaneously) eliminating *one or more* DD strategies. If this particular pair $(\mathcal{S}, \mathcal{R}_{\mathtt{sim}})$ satisfied the diamond property, then by starting from $S$ (i.e., the set of all strategy profiles) and traveling through $\mathcal{S}$ following the relation $\mathcal{R}_{\mathtt{sim}}$, one would always terminate (because there are finitely many strategies to eliminate) and end up at the same set of strategy profiles. This would actually prove that all resilient solutions are not just equivalent to each other, but actually equal to each other. This, however, is too good to be true.

The so-defined pair $(\mathcal{S}, \mathcal{R}_{\mathtt{sim}})$ *does not* satisfy the diamond property. This can be derived from the fact that, as already shown, the game in Example 2 has two distinct resilient solutions. But a more detailed explanation is the following. Let $X$, $Y$, and $Z$ be sets in $\mathcal{S}$ such that $X \mathcal{R}_{\mathtt{sim}} Y$ and $X \mathcal{R}_{\mathtt{sim}} Z$. In particular, $Y$ could be obtained from $X$ by eliminating a strategy $s_i$ of player $i$ because it is distinguishably dominated by (and only by) a strategy $t_i$, and $Z$ could be obtained from $X$ by eliminating $s_j$ of player $j$. Further, assume that the only strategy subprofile that distinguishes $s_i$ and $t_i$ over $X_{-i}$ has $s_j$ as its $j$th component. Accordingly, $s_i$ and $t_i$ become equivalent over $Z_{-i}$, and $s_i$ cannot be eliminated from $Z$, implying that there does not exist any $W \in \mathcal{S}$ such that $Y \mathcal{R}_{\mathtt{sim}} W$ and $Z \mathcal{R}_{\mathtt{sim}} W$.

The latter problem is actually exacerbated when $Y$ and $Z$ are obtained from $X$ by simultaneously eliminating multiple DD strategies. Accordingly, we restrict the relation $\mathcal{R}_{\mathtt{sim}}$ by disallowing simultaneous elimination. In other words, we consider the binary relation $\mathcal{R}$ over $\mathcal{S}$, such that $X \mathcal{R} Y$ if and only if $Y$ can be obtained from $X$ by eliminating *a single* DD strategy. At this point, Theorem 1 follows from the following two properties.

- For all $X$ and $Y$ in $\mathcal{S}$, $X \mathcal{R}^*_{\mathtt{sim}} Y$ if and only if $X \mathcal{R}^* Y$.

- Relation $\mathcal{R}$ satisfies the diamond property.

Unfortunately, neither property holds. We do, however, enlarge the relation $\mathcal{R}$ to make both of them hold. Essentially, we let $X \mathcal{R} Y$ mean that the set $Y$ is obtained from $X$ by either

(1) eliminating a DD strategy as before

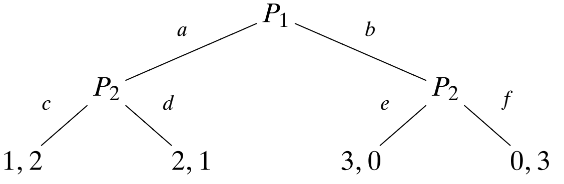
(2) eliminating a strategy indistinguishable to another one currently present

(3) replacing a strategy with an indistinguishable one (with respect to all other currently present strategies) that is not currently present.

With these changes, we "force" the desired properties to hold. However, with respect to the enlarged relation $\mathcal{R}$, unique termination is not well defined. This is so because, by solely replacing equivalent strategies, it is possible to go from a set $W$ to a different set $W'$ and back without ever terminating. Accordingly, the diamond property in our case does not imply that all resilient solutions are equal, because some of them may not be terminal with respect to $\mathcal{R}$. But, together with some other properties of the enlarged relation, it does imply that all resilient solutions are equivalent. In a sense, the slackness forced in the relation $\mathcal{R}$ translates equality into equivalence. In other words, if two resilient solutions are not equal outright, then we prove that it is possible to transform one into the other by adding/removing/replacing indistinguishable strategies, that is, via operations that produce only equivalent sets of strategy profiles.

### 6.2 *The convenience of using distinguishable dominance for proving Theorem 1*

Consider the following game $G_4$.




In this game, starting with the set of strategy profiles $X = \{a, b\} \times \{ce, df\}$ and eliminating conditionally dominated strategies, one can get

$$Y = \{a\} \times \{df\} \quad \text{and} \quad Z = \{b\} \times \{ce\}.$$[14]

[14]On one hand, starting with $X$ and eliminating $ce$ (which is conditionally dominated within $X$ by $df$, relative to the decision node following $b$), one obtains $Y' = \{a, b\} \times \{df\}$. Then, by eliminating $b$ (which is conditionally dominated within $Y'$ by $a$, relative to the root), one obtains $Y$. On the other hand, starting with $X$ and eliminating $df$ (which is conditionally dominated within $X$ by $ce$, relative to the decision node following $a$), one obtains $Z' = \{a, b\} \times \{ce\}$. Then, by eliminating $a$ (which is conditionally dominated within $Z'$ by $b$, relative to the root), one obtains $Z$.

Notice that $H(Y) = \{(a, d)\}$, $H(Z) = \{(b, e)\}$ and these two histories are not even payoff-equivalent. Accordingly, $Y$ and $Z$ are not at all equivalent: in other words, if $\widetilde{\mathcal{R}}$ is the (properly enlarged[15]) relation corresponding to the elimination of *conditionally dominated* strategies, then

$$X \,\widetilde{\mathcal{R}}^*\, Y \wedge X \,\widetilde{\mathcal{R}}^*\, Z \text{ does not imply that there exists } W \text{ such that } Y \,\widetilde{\mathcal{R}}^*\, W \wedge Z \,\widetilde{\mathcal{R}}^*\, W.$$

Note too, however, that it is *not* possible to obtain $X$ in this game by eliminating conditionally dominated strategies starting with $S$. This is, in general, the case. Indeed, we say that $X$ *is reachable from* $S$ if $S \,\widetilde{\mathcal{R}}^*\, X$. Following the bridge lemma and the fact that our enlarged relation for distinguishable dominance satisfies the diamond property, we have that

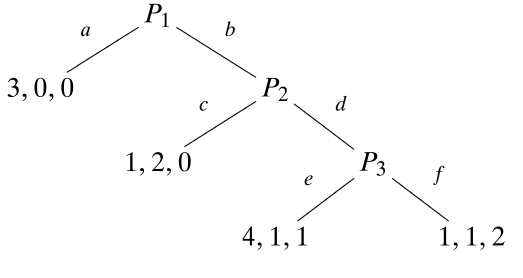
$$\widetilde{\mathcal{R}} \text{ satisfies the diamond property for all sets of strategy profiles } X \text{ reachable from } S.$$

In the absence of our results, however, the above statement was not known to be true. Further, any direct proof would have to leverage the hypothesis that "$X$ is reachable from $S$." By contrast, distinguishable dominance satisfies the diamond property for all $X$, thus allowing for a more abstract and uniform proof: the one intuitively sketched in the previous subsection.

## 7. A more general order-independence result

As shown by the following example, when a game is played, if the players iteratively eliminate DD strategies according to *different* orders and each player chooses strategies from his own surviving set, then the resulting set of possible strategy profiles need not be a resilient solution at all.

Example 5. Consider the following game $G_5$.




In this game, there are (at least) the following three elimination orders of DD strategies.

1. Eliminating strategy $e$ (dominated by $f$) then $d$ (dominated by $c$) and finally $b$ (dominated by $a$) yields a resilient solution $R^1 = \{a\} \times \{c\} \times \{f\}$.

2. Eliminating strategy $e$ and then $b$ yields a resilient solution $R^2 = \{a\} \times \{c, d\} \times \{f\}$.

[15]We do not know how to enlarge the elimination of conditionally dominated strategies without introducing our notion of distinguishable dominance, because the enlargement we have in mind is to allow elimination and replacement of indistinguishable strategies as we see in the last paragraph of Section 6.1.

3. Eliminating strategy $d$ and then $b$ yields a resilient solution $R^3 = \{a\} \times \{c\} \times \{e, f\}$.

Accordingly, $R^1_1 \times R^2_2 \times R^3_3 = \{a\} \times \{c, d\} \times \{e, f\}$. But this set of strategy profiles is not a resilient solution: indeed, one can verify that the strategies $d$ and $e$ never appear together in any resilient solution.

Notice, however, that the product set $R^1_1 \times R^2_2 \times R^3_3$ is equivalent to $R^1$ (and thus to every resilient solution of $G_5$). A consequence of Theorem 1, stated below without proof, is that this is always the case for games with perfect recall. ♢

Theorem 2. *For all resilient solutions $R^1, \dots, R^n$, the set of strategy profiles $\times_i R^i_i$ is equivalent to every resilient solution (and thus to $\mathbb{ER}$).*

## 8. Connection between EFR and nice weak dominance

Letting $U$ be the function mapping a strategy profile $s$ to the payoff profile $(u_1(s), \dots, u_n(s))$, below we recall the notion of *nice weak dominance* that Marx and Swinkels (1997) propose.

Definition 6. Let $R$ be a set of strategy profiles and let $i$ be a player. A strategy $s_i \in R_i$ is *nicely weakly dominated within* $R$ if there exists a strategy $\sigma_i \in \Delta(R_i)$ such that (i) for all $s_{-i} \in R_{-i}$, either $u_i(s_i, s_{-i}) < u_i(\sigma_i, s_{-i})$ or $U(s_i, s_{-i}) = U(\sigma_i, s_{-i})$, and (ii) there exists $s_{-i} \in R_{-i}$ such that $u_i(s_i, s_{-i}) < u_i(\sigma_i, s_{-i})$.

The notions of iterated elimination, elimination order, and maximal elimination order are defined for nice weak dominance exactly in the same way as for distinguishable dominance. As Marx and Swinkels (1997) prove, for each pair of elimination orders of nicely weakly dominated strategies, letting $R$ and $T$ be the corresponding sets of surviving strategy profiles, we have

$$U(R) = U(T).$$

Using this result and our bridge lemma, we can prove the following theorem.

Theorem 3. *For every set of strategy profiles $\mathbb{NW}$ that survives some elimination order of nicely weakly dominated strategies, we have*

$$U(\mathbb{ER}) \supseteq U(\mathbb{NW}).$$

Proof. By the definitions of distinguishable dominance and nice weak dominance, we have that for all sets of strategy profiles $T$, players $i$, and strategies $s_i \in T_i$ and $\sigma_i \in \Delta(T_i)$,

$$s_i \prec_T \sigma_i \text{ implies that } s_i \text{ is nicely weakly dominated by } \sigma_i \text{ within } T.$$

To see why this is true, assume $s_i \prec_T \sigma_i$. By definition, the following two conditions hold.

(i) For all $s_{-i} \in T_{-i}$, either $u_i(s_i, s_{-i}) < u_i(\sigma_i, s_{-i})$ or $H(s_i, s_{-i}) = H(\sigma_i, s_{-i})$.

(ii) There exists $s_{-i} \in T_{-i}$ such that $u_i(s_i, s_{-i}) < u_i(\sigma_i, s_{-i})$.

Because $H(s_i, s_{-i}) = H(\sigma_i, s_{-i})$ implies $U(s_i, s_{-i}) = U(\sigma_i, s_{-i})$, by definition $s_i$ is nicely weakly dominated by $\sigma_i$ within $T$.

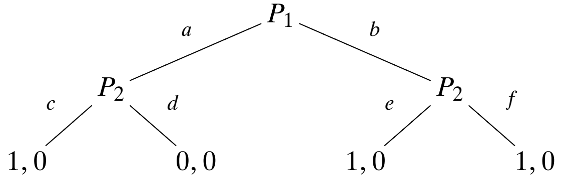
Accordingly, letting $R^0 = S, R^1, \ldots, R^K$ be the maximal elimination order of distinguishably dominated strategies, we have that for each $k < K$ and each player $i$, the strategies in $R_i^k \setminus R_i^{k+1}$ are all nicely weakly dominated within $R^k$. Therefore, $R^0, R^1, \ldots, R^K$ is the start of some particular elimination order $R^0, R^1, \ldots, R^K, \ldots, R^L$ of nicely weakly dominated strategies, where $L \geq K$. (Although $R^K$ does not contain any strategy that is DD within $R^K$, it may still contain some strategies that are nicely weakly dominated within $R^K$, and thus $L$ may be greater than $K$.) Notice that $R^0, \ldots, R^L$ may not be the elimination order that leads to $\mathbb{NW}$. But according to Marx and Swinkels (1997), we have

$$U(R^L) = U(\mathbb{NW}).$$

By Lemma 1, $R^K = \mathbb{ER}$. Because $R^K \supseteq R^L$, we finally have $U(\mathbb{ER}) = U(R^K) \supseteq U(R^L) = U(\mathbb{NW})$ and Theorem 3 holds. □

Because distinguishable dominance coincides with strict dominance in normal-form games, and it is well known that iterated elimination of nicely weakly dominated strategies can lead to a smaller set of surviving payoff profiles than iterated elimination of strictly dominated strategies, we immediately have that the inclusion in Theorem 3 can be strict. The following example shows that this continues to be the case even for extensive games (of height greater than 1).

Example 6. Consider the following game $G_6$.




In this game, on one hand, no strategy is distinguishably dominated, which implies that $\mathbb{ER} = S$ and $U(\mathbb{ER}) = \{(1, 0), (0, 0)\}$. On the other hand, the strategy $a$ of $P_1$ is nicely weakly dominated by $b$ within $S$: indeed, for $s_2 \in \{ce, cf\}$, $U(a, s_2) = U(b, s_2) = (1, 0)$, and for $s_2 \in \{de, df\}$, $u_1(a, s_2) = 0 < 1 = u_1(b, s_2)$. After $a$ is eliminated, no strategy is nicely weakly dominated, and the set of surviving strategy profiles is $\mathbb{NW} = \{b\} \times \{ce, cf, de, df\}$. Therefore, $U(\mathbb{NW}) = \{(1, 0)\} \subsetneq U(\mathbb{ER})$. ◊

## Appendix A: Proof of Lemma 1

We start by proving the following lemma, which also is used later in proving other theorems.

Lemma 2. *Let $X$ and $Y$ be two sets of strategy profiles such that $Y \subsetneq X$, and for each player $i$ and each $s_i \in X_i \setminus Y_i$, $s_i$ is distinguishably dominated within $X$. Then, for each player $i$ and each $s_i \in X_i \setminus Y_i$, there exists $\sigma_i \in \Delta(Y_i)$ such that $s_i \prec_{X_{-i}} \sigma_i$.*

Proof. Consider an arbitrary player $i$. Without loss of generality, assume that $X_i \setminus Y_i \neq \varnothing$. Let $k = |X_i \setminus Y_i|$ and $X_i \setminus Y_i = \{s_{i,1}, \dots, s_{i,k}\}$. To prove Lemma 2, it suffices to show that

$$\text{for each } \ell \leq k \text{, there exists a strategy } \sigma_{i,\ell} \in \Delta(Y_i) \text{ such that } s_{i,\ell} \prec_{X_{-i}} \sigma_{i,\ell}. \qquad (\star)$$

To prove statement ($\star$), notice that by hypothesis, for each $\ell \leq k$, there exists $\tau_{i,\ell} \in \Delta(X_i)$ such that $s_{i,\ell} \prec_{X_{-i}} \tau_{i,\ell}$. If all those $\tau_{i,\ell}$ are in $\Delta(Y_i)$, then letting $\sigma_{i,\ell} = \tau_{i,\ell}$ for each $\ell$, we are done immediately. Otherwise, we construct $\sigma_{i,1}, \dots, \sigma_{i,k}$ explicitly, and in $k$ steps.

For $j = 1, \dots, k$, the goal of the $j$th step is to construct $\sigma^j_{i,1}, \dots, \sigma^j_{i,k}$, such that for each $\ell \leq k$, $s_{i,\ell} \prec_{X_{-i}} \sigma^j_{i,\ell}$ and $\sigma^j_{i,\ell} \in \Delta(X_i \setminus \{s_{i,1}, \dots, s_{i,j}\})$. (Intuitively, we want to gradually remove $s_{i,1}, \dots, s_{i,k}$ from the support of each $\tau_{i,\ell}$, while preserving the corresponding distinguishable dominance relation.) Notice that once all $k$ steps are done successfully, we obtain $\sigma^k_{i,1}, \dots, \sigma^k_{i,k}$ such that (by the goal of the $k$th step) for each $\ell \leq k$, $s_{i,\ell} \prec_{X_{-i}} \sigma^k_{i,\ell}$ and $\sigma^k_{i,\ell} \in \Delta(X_i \setminus \{s_{i,1}, \dots, s_{i,k}\}) = \Delta(Y_i)$. Thus by taking $\sigma_{i,\ell} = \sigma^k_{i,\ell}$ for each $\ell \leq k$, statement ($\star$) holds, and so does Lemma 2.

Now we implement the above proposed $k$-steps. In the first step, we construct $\sigma^1_{i,1}, \dots, \sigma^1_{i,k}$ based on $\tau_{i,1}, \dots, \tau_{i,k}$. We start from $\sigma^1_{i,1}$. Notice that $\tau_{i,1} \neq s_{i,1}$ (in other words, $\tau_{i,1}(s_{i,1}) \neq 1$), because $s_{i,1} \not\approx_{X_{-i}} \tau_{i,1}$. Therefore, we take $\sigma^1_{i,1}$ to be $\tau_{i,1}$ conditioned on $s_{i,1}$ not occurring, that is,

$$\sigma^1_{i,1}(s_i) = \frac{\tau_{i,1}(s_i)}{1 - \tau_{i,1}(s_{i,1})} \quad \text{for all } s_i \neq s_{i,1}.$$

In particular, if $\tau_{i,1}(s_{i,1}) = 0$, then $\sigma^1_{i,1} = \tau_{i,1}$. By construction, $\sigma^1_{i,1} \in \Delta(X_i \setminus \{s_{i,1}\})$: indeed,

$$\sum_{s_i \in X_i \setminus \{s_{i,1}\}} \sigma^1_{i,1}(s_i) = \frac{1}{1 - \tau_{i,1}(s_{i,1})} \sum_{s_i \in X_i \setminus \{s_{i,1}\}} \tau_{i,1}(s_i) = \frac{1 - \tau_{i,1}(s_{i,1})}{1 - \tau_{i,1}(s_{i,1})} = 1.$$

Also by construction, for each strategy subprofile $t_{-i}$, $t_{-i}$ distinguishes $s_{i,1}$ and $\sigma^1_{i,1}$ if and only if it distinguishes $s_{i,1}$ and $\tau_{i,1}$. Because $s_{i,1} \not\approx_{X_{-i}} \tau_{i,1}$, we have $s_{i,1} \not\approx_{X_{-i}} \sigma^1_{i,1}$. Further, because for all distinguishing strategy subprofiles $t_{-i} \in X_{-i}$,

$$u_i(s_{i,1}, t_{-i}) < u_i(\tau_{i,1}, t_{-i}) = (1 - \tau_{i,1}(s_{i,1}))u_i(\sigma^1_{i,1}, t_{-i}) + \tau_{i,1}(s_{i,1})u_i(s_{i,1}, t_{-i}),$$

we have $u_i(s_{i,1}, t_{-i}) < u_i(\sigma^1_{i,1}, t_{-i})$. Accordingly, $s_{i,1} \prec_{X_{-i}} \sigma^1_{i,1}$.

Now for each $\ell \neq 1$, we construct $\sigma^1_{i,\ell}$ based on $\tau_{i,\ell}$ and $\sigma^1_{i,1}$. To do so, for each $s_i \in X_i \setminus \{s_{i,1}\}$, let

$$\sigma^1_{i,\ell}(s_i) = \tau_{i,\ell}(s_i) + \tau_{i,\ell}(s_{i,1}) \cdot \sigma^1_{i,1}(s_i).$$

That is, $\sigma^1_{i,\ell}$ is obtained from $\tau_{i,\ell}$ by replacing $s_{i,1}$ with $\sigma^1_{i,1}$. By construction, we have $\sigma^1_{i,\ell} \in \Delta(X_i \setminus \{s_{i,1}\})$: indeed,

$$\begin{aligned}\sum_{s_i \in X_i \setminus \{s_{i,1}\}} \sigma^1_{i,\ell}(s_i) &= \sum_{s_i \in X_i \setminus \{s_{i,1}\}} \tau_{i,\ell}(s_i) + \tau_{i,\ell}(s_{i,1}) \cdot \sum_{s_i \in X_i \setminus \{s_{i,1}\}} \sigma^1_{i,1}(s_i) \\ &= (1 - \tau_{i,\ell}(s_{i,1})) + \tau_{i,\ell}(s_{i,1}) = 1.\end{aligned}$$

Next we prove that $s_{i,\ell} \prec_{X_{-i}} \sigma^1_{i,\ell}$, given the hypothesis $s_{i,\ell} \prec_{X_{-i}} \tau_{i,\ell}$. To do so, notice that when $\tau_{i,\ell}(s_{i,1}) = 0$, we have $\tau_{i,\ell} = \sigma^1_{i,\ell}$, which together with the hypothesis clearly implies $s_{i,\ell} \prec_{X_{-i}} \sigma^1_{i,\ell}$. When $\tau_{i,\ell}(s_{i,1}) > 0$, we have $\tau_{i,\ell} \neq \sigma^1_{i,\ell}$, and that for each $t_{-i}$, $t_{-i}$ distinguishes $\tau_{i,\ell}$ and $\sigma^1_{i,\ell}$ if and only if it distinguishes $s_{i,1}$ and $\sigma^1_{i,1}$. Because $s_{i,1} \prec_{X_{-i}} \sigma^1_{i,1}$, when $\tau_{i,\ell}(s_{i,1}) > 0$, we have (i) there exists $t_{-i} \in X_{-i}$ distinguishing $\tau_{i,\ell}$ and $\sigma^1_{i,\ell}$, and (ii) for all such $t_{-i}$,

$$\begin{aligned}u_i(\tau_{i,\ell}, t_{-i}) &= \sum_{s_i \in X_i \setminus \{s_{i,1}\}} \tau_{i,\ell}(s_i) u_i(s_i, t_{-i}) + \tau_{i,\ell}(s_{i,1}) u_i(s_{i,1}, t_{-i}) \\ &< \sum_{s_i \in X_i \setminus \{s_{i,1}\}} \tau_{i,\ell}(s_i) u_i(s_i, t_{-i}) + \tau_{i,\ell}(s_{i,1}) u_i(\sigma^1_{i,1}, t_{-i}) = u_i(\sigma^1_{i,\ell}, t_{-i}).\end{aligned}$$

Accordingly, when $\tau_{i,\ell}(s_{i,1}) > 0$, we have $\tau_{i,\ell} \prec_{X_{-i}} \sigma^1_{i,\ell}$. Because the $\prec_{X_{-i}}$ relation is transitive, together with the hypothesis we have $s_{i,\ell} \prec_{X_{-i}} \sigma^1_{i,\ell}$ and we are done with the first step.

The remaining steps are very similar. In particular, in the $j$th step for each $j > 1$, we construct $\sigma^j_{i,1}, \dots, \sigma^j_{i,k}$ based on $\sigma^{j-1}_{i,1}, \dots, \sigma^{j-1}_{i,k}$. We start from $\sigma^j_{i,j}$, and take it to be $\sigma^{j-1}_{i,j}$ conditioned on $s_{i,j}$ not occurring. For each $\ell \neq j$, $\sigma^j_{i,\ell}$ is obtained from $\sigma^{j-1}_{i,\ell}$ by replacing $s_{i,j}$ with $\sigma^j_{i,j}$. By similar analysis, we have that for each $\ell \leq k$, $\sigma^j_{i,\ell} \in \Delta(X_i \setminus \{s_{i,1}, \dots, s_{i,j}\})$ and $s_{i,\ell} \prec_{X_{-i}} \sigma^j_{i,\ell}$, as desired.

As already mentioned, after the $k$th step, we have $\sigma^k_{i,1}, \dots, \sigma^k_{i,k}$ such that for each $\ell \leq k$, $s_{i,\ell} \prec_{X_{-i}} \sigma^k_{i,\ell}$ and $\sigma^k_{i,\ell} \in \Delta(X_i \setminus \{s_{i,1}, \dots, s_{i,k}\}) = \Delta(Y_i)$. Taking $\sigma_{i,\ell} = \sigma^k_{i,\ell}$ for each $\ell \leq k$, statement ($\star$) holds, and so does Lemma 2. □

We now proceed to prove Lemma 1. The proof consists of three parts. In the first part, which is the most complicated one, we prove that each elimination order of conditionally dominated strategies is also an elimination order of DD strategies. To do so, letting $S^0 = S, S^1, \dots, S^K$ be an arbitrary elimination order of conditionally dominated strategies, we prove the following statement:

For all $k \leq K$, all $i$, and all $s_i \in S^k_i$, $s_i$ is conditionally dominated within $S^k$ if and only if it is distinguishably dominated within $S^k$. (∗)

Indeed, statement (∗) implies that for all $k < K$ and all players $i$, every strategy in $S^k_i \setminus S^{k+1}_i$ is distinguishably dominated with $S^k$. Further, because each $S^K_i$ contains no strategy that is conditionally dominated within $S^K$, statement (∗) further implies that each $S_i^K$ contains no strategy that is distinguishably dominated within $S^K$. Since $S^0 = S$, by definition $S^0, S^1, \dots, S^K$ is an elimination order of distinguishably dominated strategies, as desired.

Let us now prove statement (∗) by induction on $k$.

*Base Case*: $k = 0$. Assume that $s_i$ is conditionally dominated within $S$ by strategy $\sigma_i$. Then there exists an information set $I \in \mathcal{I}_i$ together with which $s_i$ and $\sigma_i$ satisfy Definition 4. In particular, we have $s_i \in S_i(I)$, $\sigma_i \in \Delta(S_i(I))$, $S_{-i}(I) \neq \emptyset$, and $\forall t_{-i} \in S_{-i}(I) u_i(s_i, t_{-i}) < u_i(\sigma_i, t_{-i})$.

We construct a mixed strategy $\sigma_i'$ as follows. For each pure strategy $a_i$ in the support of $\sigma_i$, let $a_i'$ be the pure strategy such that

(i) $a_i'(I) = a_i(I)$

(ii) $a_i'(I') = a_i(I')$ for all information sets $I' \in \mathcal{I}_i$ following $I$

(iii) $a_i'(I') = s_i(I')$ for all other information sets $I' \in \mathcal{I}_i$.

Notice that $a_i'$ is a well defined pure strategy because the game is with *perfect recall*. Let $\sigma_i'(a_i') = \sigma_i(a_i)$.

We prove $s_i \prec_S \sigma_i'$. First consider an arbitrary $t_{-i}$ in $S_{-i} \setminus S_{-i}(I)$. By Fact 2 of Section 3, neither $H(s_i, t_{-i})$ nor any $H(a_i', t_{-i})$ with $a_i'$ in the support of $\sigma_i'$ reaches $I$. Accordingly, $H(s_i, t_{-i}) = H(a_i', t_{-i})$ for each $a_i'$, which implies that $H(s_i, t_{-i}) = H(\sigma_i', t_{-i})$. Therefore, such a $t_{-i}$ does not distinguish $s_i$ and $\sigma_i'$.

Now consider an arbitrary $t_{-i}$ in $S_{-i}(I)$. (Because $S_{-i}(I) \neq \emptyset$, such a $t_{-i}$ always exists.) Because $\sigma_i$ is in $\Delta(S_i(I))$, by construction so is $\sigma_i'$. Accordingly, Fact 2 of Section 3 implies that the three (distributions of) histories $H(s_i, t_{-i})$, $H(\sigma_i, t_{-i})$, and $H(\sigma_i', t_{-i})$ not only all reach $I$, but actually all reach the same decision node in $I$. For each $a_i$ in the support of $\sigma_i$, because $a_i$ and the corresponding $a_i'$ coincide at $I$ and at every information set following $I$, we have $H(a_i, t_{-i}) = H(a_i', t_{-i})$. Thus $H(\sigma_i, t_{-i}) = H(\sigma_i', t_{-i})$, which further implies $u_i(\sigma_i, t_{-i}) = u_i(\sigma_i', t_{-i})$. Since $u_i(\sigma_i, t_{-i}) > u_i(s_i, t_{-i})$ (by the definition of conditional dominance), we have $u_i(\sigma_i', t_{-i}) > u_i(s_i, t_{-i})$, which of course implies that $t_{-i}$ distinguishes $s_i$ and $\sigma_i'$.

Since apparently $\sigma_i' \in \Delta(S_i)$, we have $s_i \prec_S \sigma_i'$ as we wanted to show.

The other direction is quite easy. Indeed, if $s_i \prec_S \sigma_i$, then by Proposition 1, $s_i$ is conditionally dominated by $\sigma_i$ within $S$.

*Induction Step*: $k > 0$. Assume that $s_i$ is conditionally dominated within $S^k$ by $\sigma_i \in \Delta(S_i^k)$. We prove $s_i \prec_{S_{-i}^k} \bar{\sigma}_i$ for some $\bar{\sigma}_i \in \Delta(S_i^k)$. To do so, let $I \in \mathcal{I}_i$ be the information set as per Definition 4. Constructing the mixed strategy $\sigma_i'$ from $\sigma_i$ as in the base case, we have

$$s_i \prec_{S_{-i}^k} \sigma_i'.$$

The remaining question is where the support of $\sigma_i'$ lies. If $\sigma_i' \in \Delta(S_i^k)$, then we are done. If $\sigma_i' \notin \Delta(S_i^k)$, then we construct the desired strategy $\bar{\sigma}_i \in \Delta(S_i^k)$ from $\sigma_i'$, as follows.

Because $\sigma_i' \in \Delta(S_i^0)$ and $\sigma_i' \notin \Delta(S_i^k)$, there exists an integer $\ell < k$ such that $\sigma_i' \in \Delta(S_i^\ell)$ and $\sigma_i' \notin \Delta(S_i^{\ell+1})$. Accordingly, there exists $a_i'$ in the support of $\sigma_i'$ such that $a_i' \in S_i^\ell \setminus S_i^{\ell+1}$.

By definition, $a_i'$ is conditionally dominated within $S^\ell$. Without loss of generality, assume that there is only one such $a_i'$; that is, $S_i^\ell \setminus S_i^{\ell+1} = \{a_i'\}$. By the induction hypothesis, $a_i'$ is distinguishably dominated within $S^\ell$ and thus there exists $\tau_i \in \Delta(S_i^\ell)$ such that $a_i' \prec_{S_{-i}^\ell} \tau_i$. According to Lemma 2, again without loss of generality, we can assume $\tau_i \in \Delta(S_i^{\ell+1})$; that is, $\tau_i(a_i') = 0$. Because $\ell < k$, we have $S_{-i}^k \subseteq S_{-i}^\ell$ and thus

$$a_i' \preceq_{S_{-i}^k} \tau_i.$$

We construct a new mixed strategy $\widehat{\sigma}_i$ from $\sigma_i'$ as follows: for all $t_i \in S_i^{\ell+1}$,

$$\widehat{\sigma}_i(t_i) = \sigma_i'(t_i) + \sigma_i'(a_i') \cdot \tau_i(t_i);$$

that is, $\widehat{\sigma}_i$ is obtained from $\sigma_i'$ by replacing $a_i'$ with $\tau_i$, as we have done in the proof of Lemma 2. Notice that $\widehat{\sigma}_i$ is a well defined mixed strategy in $\Delta(S_i^{\ell+1})$: indeed,

$$\sum_{t_i \in S_i^{\ell+1}} \widehat{\sigma}_i(t_i) = \sum_{t_i \in S_i^{\ell+1}} \sigma_i'(t_i) + \sigma_i'(a_i') \cdot \sum_{t_i \in S_i^{\ell+1}} \tau_i(t_i) = \sum_{t_i \in S_i^{\ell+1}} \sigma_i'(t_i) + \sigma_i'(a_i') = \sum_{t_i \in S_i^\ell} \sigma_i'(t_i) = 1.$$

Because $a_i' \preceq_{S_{-i}^k} \tau_i$, and by the construction of $\widehat{\sigma}_i$, we have $\sigma_i' \preceq_{S_{-i}^k} \widehat{\sigma}_i$, as we have seen in the proof of Lemma 2. Because $s_i \prec_{S_{-i}^k} \sigma_i'$, we finally have

$$s_i \prec_{S_{-i}^k} \widehat{\sigma}_i.$$

Comparing with $\sigma_i'$, we have brought the support of $\widehat{\sigma}_i$ from $S_i^\ell$ to $S_i^{\ell+1}$.

Repeat the above procedure, with the role of $\sigma_i'$ replaced by $\widehat{\sigma}_i$, and we finally get a mixed strategy $\bar{\sigma}_i \in \Delta(S_i^k)$ such that $s_i \prec_{S_{-i}^k} \bar{\sigma}_i$, as we wanted to do.

Again by Proposition 1 it is easy to see that the other direction is true; that is, if $s_i$ is distinguishably dominated within $S^k$ (by $\sigma_i$), then $s_i$ is also conditionally dominated within $S^k$ (by the same $\sigma_i$). Therefore, statement ($*$) holds, concluding the first part of our proof of Lemma 1.

In the second part, we prove that any elimination order of DD strategies is also an elimination order of conditionally dominated strategies. To do so, letting $R^0 = S, R^1, \ldots, R^K$ be an arbitrary elimination order of DD strategies, following Proposition 1, we already have that for all $k < K$ and all players $i$, every strategy in $R_i^k \setminus R_i^{k+1}$ is conditionally dominated within $R^k$. Accordingly, the first part of the proof of Lemma 1 implies that any strategy that is conditionally dominated within $R^K$ must be distinguishably dominated within $R^K$. Because each $R_i^K$ contains no strategy that is distinguishably dominated within $R^K$, each $R_i^K$ contains no strategy that is conditionally dominated within $R^K$ either. Therefore, $R^0, \ldots, R^K$ is an elimination order of conditionally dominated strategies, concluding the second part of our proof of Lemma 1.

In the last part, we prove that the maximal elimination of conditionally dominated strategies, denoted by the sequence $M^0 = S, M^1, \ldots, M^K$, is also the maximal elimination of DD strategies. This follows almost directly from the first part. Indeed, the conclusion of the first part guarantees that $M^0, \ldots, M^K$ is an elimination order of DD strategies.

Moreover, because for each $k < K$ and each player $i$, $M_i^k \setminus M_i^{k+1}$ consists of all strategies that are conditionally dominated within $M^k$, statement $(*)$ implies that $M_i^k \setminus M_i^{k+1}$ also consists of all strategies that are distinguishably dominated within $M^k$, which means that $M^0, \ldots, M^K$ is the maximal elimination of DD strategies, as desired.

In sum, Lemma 1 holds. □

## Appendix B: Proof of Proposition 2 and Example 7

Proof of Proposition 2. For each player $i$, because $s_i' \in s_i^{\simeq}$, we have $s_i \simeq_{R_{-i}} s_i'$ by definition. Therefore,

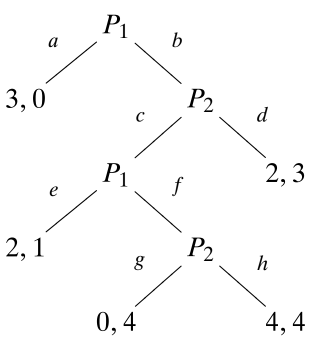
$$\begin{aligned} H(s'_{\{1,\ldots,i-1\}}, s_{\{i,\ldots,n\}}) &= H(s'_{\{1,\ldots,i-1\}}, s_i, s_{\{i+1,\ldots,n\}}) \\ &= H(s'_{\{1,\ldots,i-1\}}, s_i', s_{\{i+1,\ldots,n\}}) = H(s'_{\{1,\ldots,i\}}, s_{\{i+1,\ldots,n\}}). \end{aligned}$$

Applying this equation repeatedly, from $i = 1$ to $i = n$, we have

$$H(s) = H(s_1', s_{-1}) = H(s'_{\{1,2\}}, s_{-\{1,2\}}) = \cdots = H(s'_{\{1,\ldots,n-1\}}, s_n) = H(s'),$$

and Proposition 2 holds. □

Example 7. Consider the following game $G_7$.




In this game, $P_2$'s strategy $dg$ is part of a subgame-perfect equilibrium: namely, $(ae, dg)$. However, $dg$ is not part of *any* resilient solution.[16] (Note that the game above is of perfect information. The same phenomenon can also be illustrated by a classical game with simultaneous moves: namely, the Battle-of-the-Sexes game with an outside option.) ◊

[16]There are precisely three elimination orders of DD strategies, namely,

1. $be$, followed by $dg$, followed by $dh$
2. $be$, followed by $dh$, followed by $dg$
3. $be$, followed by a simultaneous elimination of $dh$ and $dg$.

Accordingly, there is only one resilient solution, namely, $R = \{ae, af, bf\} \times \{cg, ch\}$, and $dg \notin R_2$.

# Appendix C: Proof of Theorem 1

## C.1 *Important relations between sets of strategy profiles*

Recall that if $\mathcal{R}$ is a binary relation between sets of strategy profiles, then $\mathcal{R}^*$ denotes the reflexive and transitive closure of $\mathcal{R}$. We first define a particular binary relation between sets of strategy profiles, which expresses the operation of eliminating precisely one DD strategy.

Definition 7. Among sets of strategy profiles, the *strict elimination relation*, denoted by $\xrightarrow[\prec]{e}$, is defined as $R \xrightarrow[\prec]{e} T$ if there exists a player $i$ such that

(i) $T_{-i} = R_{-i}$

(ii) $T_i = R_i \setminus \{s_i\}$, where $s_i \in R_i$ and $s_i \prec_{R_{-i}} \tau_i$ for some $\tau_i \in \Delta(T_i)$.

To emphasize the role of $s_i$ and $\tau_i$, we may write $R \xrightarrow[s_i \prec \tau_i]{e} T$.

If $R$ is a set of strategy profiles, then $R$ is *strict-elimination-free* if there exists no $T$ such that $R \xrightarrow[\prec]{e} T$.

Notice that if $R$ is a resilient solution, then it is strict-elimination-free. Before defining the enlarged relation, below we briefly discuss what properties we want it to satisfy.

*Properties wanted for the enlarged relation.* As mentioned in Section 6.1, to prove that any two resilient solutions $R$ and $T$ are equivalent, we enlarge the relation $\xrightarrow[\prec]{e}$ to a relation $\longrightarrow$ such that the set $\mathcal{S}$ of all sets of strategy profiles together with the relation $\longrightarrow$ satisfies the *diamond property*. But also recall from Section 6.1 that the relation $\longrightarrow$ has to satisfy some other properties. In particular, if $R$ and $T$ are resilient solutions, then we want

1. $S \longrightarrow^* R$ and $S \longrightarrow^* T$
2. from the relationships in item 1 and the diamond property, we can deduce that
   2.1. $R \longrightarrow^* W$ and $T \longrightarrow^* W$ for some $W$, and more importantly
   2.2. the paths from $R$ to $W$ and $T$ to $W$ are both "equivalence-preserving."

Toward the above desired properties, we define two equivalence-preserving relations between sets of strategy profiles, and the desired relation $\longrightarrow$ is obtained by combining them together with the relation $\xrightarrow[\prec]{e}$. The first relation expresses the operation of eliminating precisely one strategy because it is indistinguishable from another one that is currently present.

Definition 8. Among sets of strategy profiles, we define two relations.

- The *indistinguishable elimination relation*, denoted by $\xrightarrow[\simeq]{e}$, is defined as $R \xrightarrow[\simeq]{e} T$ if there exists a player $i$ such that

  (i) $T_{-i} = R_{-i}$

  (ii) $T_i = R_i \setminus \{s_i\}$, where $s_i \in R_i$ and $s_i \simeq_{R_{-i}} t_i$ for some $t_i \in T_i$.

To emphasize the role of $s_i$ and $t_i$, we may write $R \xrightarrow[s_i \simeq t_i]{e} T$.[17]

- The *elimination relation*, denoted by $\xrightarrow{e}$, encompasses the $\xrightarrow[\prec]{e}$ and $\xrightarrow[\simeq]{e}$ relations as follows:

$$R \xrightarrow{e} T \text{ if and only if either } R \xrightarrow[s_i \prec \tau_i]{e} T \text{ or } R \xrightarrow[s_i \simeq \tau_i]{e} T.$$

To emphasize the role of $s_i$ and $\tau_i$ in $\xrightarrow{e}$, we may write $R \xrightarrow[s_i \preceq \tau_i]{e} T$ or simply $R \xrightarrow[s_i, \tau_i]{e} T$.

The second equivalence-preserving relation expresses the operation of replacing one strategy with an indistinguishable one that is not currently present.

Definition 9. Among sets of strategy profiles, the *replacement relation*, denoted by $\xrightarrow{r}$, is defined as $R \xrightarrow{r} T$ if either (i) $R = T$ or (ii) there exists a player $i$ such that

(ii.a) $T_{-i} = R_{-i}$

(ii.b) $R_i \setminus T_i = \{s_i\}$ and $T_i \setminus R_i = \{t_i\}$, where $s_i \simeq_{R_{-i}} t_i$.

We may write $R \xrightarrow[\epsilon]{r} T$ to emphasize that we are in case (i) and $R \xrightarrow[s_i, t_i]{r} T$ that we are in case (ii).

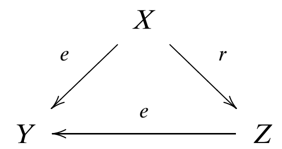
The relation $\xrightarrow{\simeq}$ is defined as $R \xrightarrow{\simeq} T$ if either $R \xrightarrow[\simeq]{e} T$ or $R \xrightarrow{r} T$.

The relation $\longrightarrow$ is defined as $R \longrightarrow T$ if either $R \xrightarrow{e} T$ or $R \xrightarrow{r} T$.

Notice that the replacement relation requires that both $s_i$ and $t_i$ be pure strategies. As we prove later, for all sets of strategy profiles $R$ and $T$, if $R \xrightarrow{\simeq}{}^{*} T$, then $R$ and $T$ are equivalent.

[17]Note that one could define $s_i \simeq_{R_{-i}} \tau_i$, where $\tau_i \in \Delta(T_i)$. Indeed, $s_i \simeq_{R_{-i}} \tau_i$ if and only if $s_i \simeq_{R_{-i}} t_i$ for every $t_i$ in the support of $\tau_i$.

Remark. Our results can certainly be proved without relying on the (sub)relation $\xrightarrow[\epsilon]{r}$. Our reason for introducing the "empty-replacement" relation is ensuring uniformity in our proofs. Without it, the diamond property may sometimes become a "triangle property": pictorially,

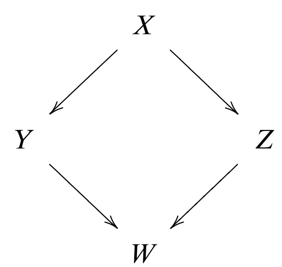



where $Y$ is obtained from $X$ by eliminating some strategy $s_i$ dominated by $t_i$, and $Z$ is obtained from $X$ by replacing $t_i$ with an equivalent strategy $t_i'$. (Recall that we are defining the diamond property for a relation $\mathcal{R}$, not for its reflexive and transitive closure $\mathcal{R}^*$.)

## C.2 *Useful lemmas*

Having defined proper relations, we now have the following lemma.

Lemma 3. *For all sets of strategy profiles $X$, $Y$, and $Z$, if $X \longrightarrow Y$ and $X \longrightarrow Z$, then there exists a set of strategy profiles $W$ such that $Y \longrightarrow W$ and $Z \longrightarrow W$. In a picture,*




Proof. By symmetry, we need to analyze only three cases.

*Case 1*: $Y \xleftarrow[s_i \preceq \sigma_i]{e} X \xrightarrow[t_j \preceq \tau_j]{e} Z$. If $i \neq j$, then we have $Y_i = X_i \setminus \{s_i\} = Z_i \setminus \{s_i\}$, $Z_j = X_j \setminus \{t_j\} = Y_j \setminus \{t_j\}$, and $Y_{-\{i,j\}} = X_{-\{i,j\}} = Z_{-\{i,j\}}$. Let $W$ be the set of strategy profiles where $W_i = Y_i$, $W_j = Z_j$, and $W_{-\{i,j\}} = X_{-\{i,j\}}$. We prove

$$Y \xrightarrow[t_j \preceq \tau_j]{e} W \xleftarrow[s_i \preceq \sigma_i]{e} Z.$$

To do so, we focus on the $Y \xrightarrow[t_j \preceq \tau_j]{e} W$ part (the other part is by symmetry). Since $t_j \in X_j$ and $\tau_j \in \Delta(Z_j)$, we have $t_j \in Y_j$ and $\tau_j \in \Delta(W_j)$. Since $t_j \preceq_{X_{-j}} \tau_j$ and $Y_{-j} \subseteq X_{-j}$, we have that for all $s_{-j} \in Y_{-j}$,

$$\text{either} \quad H(t_j, s_{-j}) = H(\tau_j, s_{-j}) \quad \text{or} \quad u_j(t_j, s_{-j}) < u_j(\tau_j, s_{-j}).$$

Therefore, $t_j \preceq_{Y_{-j}} \tau_j$ and $Y \xrightarrow[t_j \preceq \tau_j]{e} W$ as desired. (From the analysis, one can see that the choice of $W$ is actually inevitable.)

If $i = j$ and $s_i = t_i$, then $Y = Z$, and letting $W = Y$, we have

$$Y \xrightarrow[\epsilon]{r} W \xleftarrow[\epsilon]{r} Z.$$

If $i = j$, $s_i \neq t_i$, and $s_i \simeq_{X_{-i}} t_i$, then we have $Y_{-i} = Z_{-i} = X_{-i}$, $Y_i \setminus Z_i = \{t_i\}$, $Z_i \setminus Y_i = \{s_i\}$, and $s_i \simeq_{Y_{-i}} t_i$. Therefore letting $W = Z$, we have

$$Y \xrightarrow[t_i, s_i]{r} W \xleftarrow[\epsilon]{r} Z.$$

(The case before and this case differ only at the relations between $Y$ and $W$: one is doing nothing and the other is replacement.)

If $i = j$, $s_i \neq t_i$, and $s_i \not\simeq_{X_{-i}} t_i$, then we have $Y_{-i} = Z_{-i} = X_{-i}$, $Y_i = X_i \setminus \{s_i\}$, $Z_i = X_i \setminus \{t_i\}$, $t_i \preceq_{Y_{-i}} \tau_i$, and $s_i \preceq_{Z_{-i}} \sigma_i$. Letting $W$ be such that $W_i = X_i \setminus \{s_i, t_i\}$ and $W_{-i} = X_{-i}$, we prove that there exists $\tau'_i \in \Delta(W_i)$ and $\sigma'_i \in \Delta(W_i)$ such that

$$Y \xrightarrow[t_i \preceq \tau'_i]{e} W \xleftarrow[s_i \preceq \sigma'_i]{e} Z.$$

To do so, we focus on the $Y \xrightarrow[t_i \preceq \tau'_i]{e} W$ part (the other part is by symmetry). Indeed, if $\tau_i(s_i) = 0$, then $\tau_i \in \Delta(Z_i \setminus \{s_i\}) = \Delta(X_i \setminus \{s_i, t_i\}) = \Delta(Y_i \setminus \{t_i\}) = \Delta(W_i)$. Take $\tau'_i = \tau_i$ and we are done. If $\tau_i(s_i) > 0$, then $\tau_i \notin \Delta(Y_i)$, and we construct a strategy $\tau''_i$ based on $\tau_i$, by replacing $s_i$ with $\sigma_i$, as we have done in the proof of Lemma 2. Indeed,

$$\forall s'_i \neq s_i \quad \tau''_i(s'_i) \triangleq \tau_i(s'_i) + \tau_i(s_i) \cdot \sigma_i(s'_i).$$

Because $\tau_i \in \Delta(Z_i) = \Delta(X_i \setminus \{t_i\}) = \Delta((Y_i \cup \{s_i\}) \setminus \{t_i\})$, we have $\tau_i \in \Delta(Y_i \cup \{s_i\})$. Further because $\sigma_i \in \Delta(Y_i)$, the so constructed $\tau''_i$ is in $\Delta(Y_i)$. Because $s_i \preceq_{Y_{-i}} \sigma_i$, we have $\tau_i \preceq_{Y_{-i}} \tau''_i$. Because $t_i \preceq_{X_{-i}} \tau_i$ and $X_{-i} = Y_{-i}$, we have

$$t_i \preceq_{Y_{-i}} \tau''_i.$$

If $\tau''_i(t_i) = 0$, then $\tau''_i \in \Delta(Y_i \setminus \{t_i\}) = \Delta(W_i)$ and we are done by taking $\tau'_i = \tau''_i$. Otherwise, notice that $\tau''_i(t_i) < 1$; indeed, assuming $\tau''_i = t_i$, we have $\tau''_i(t_i) = 1 = \tau_i(t_i) + \tau_i(s_i) \cdot \sigma_i(t_i)$, which together with the fact $\tau_i(t_i) = 0$ implies that $\tau_i = s_i$ and $\sigma_i = t_i$, which together with the facts $s_i \preceq_{X_{-i}} \sigma_i$ and $t_i \preceq_{X_{-i}} \tau_i$ further imply $s_i \simeq_{X_{-i}} t_i$, contradicting the hypothesis. Accordingly, $\tau''_i(t_i) < 1$, and by taking $\tau'_i$ to be $\tau''_i$ conditioned on $t_i$ not occurring, we have $\tau'_i \in \Delta(Y_i \setminus \{t_i\}) = \Delta(W_i)$ and $t_i \preceq_{Y_{-i}} \tau'_i$, and we are done as well.

*Case 2*: $Y \xleftarrow[s_i \preceq \sigma_i]{e} X \xrightarrow{r} Z$. In this case, if $X \xrightarrow[\epsilon]{r} Z$ then letting $W = Y$, we have $Y \xrightarrow[\epsilon]{r} W \xleftarrow[s_i \preceq \sigma_i]{e} Z$.

Now assume $X \xrightarrow[t_j, t_j']{r} Z$, that is,

$$Y \xleftarrow[s_i \preceq \sigma_i]{e} X \xrightarrow[t_j, t_j']{r} Z.$$

We consider three subcases.

*Subcase 2.1*: $i \neq j$. In this case, we have $t_j \in Y_j$, $t_j' \notin Y_j$, and $t_j \simeq_{Y_{-j}} t_j'$ (because $Y_{-j} \subsetneq X_{-j}$). Letting $W$ be $Y$ with $t_j$ replaced by $t_j'$, that is, $W_j = (Y_j \setminus \{t_j\}) \cup \{t_j'\}$ and $W_{-j} = Y_{-j}$, we have $Y \xrightarrow[t_j, t_j']{r} W$. We now show that $Z \xrightarrow[s_i \preceq \sigma_i]{e} W$. To see why this is true, notice that (i) $W_i = Y_i = X_i \setminus \{s_i\} = Z_i \setminus \{s_i\}$, (ii) $W_j = (Y_j \setminus \{t_j\}) \cup \{t_j'\} = (X_j \setminus \{t_j\}) \cup \{t_j'\} = Z_j$, (iii) $W_{-\{i,j\}} = Y_{-\{i,j\}} = X_{-\{i,j\}} = Z_{-\{i,j\}}$, (iv) $s_i \in Z_i$ $(= X_i)$, and (v) $\sigma_i \in \Delta(W_i)$ $(= \Delta(Y_i))$.

Therefore, it suffices to show that $s_i \preceq_{Z_{-i}} \sigma_i$. To do so, notice that for all $y_{-i} \in Z_{-i}$, if $y_j \neq t_j'$, then $y_{-i} \in X_{-i}$ as well, and thus

$$\text{either} \quad H(s_i, y_{-i}) = H(\sigma_i, y_{-i}) \quad \text{or} \quad u_i(s_i, y_{-i}) < u_i(\sigma_i, y_{-i}),$$

because $s_i \preceq_{X_{-i}} \sigma_i$. If $y_j = t_j'$, then

$$H(s_i, t_j', y_{-\{i,j\}}) = H(s_i, t_j, y_{-\{i,j\}}) \quad \text{and} \quad H(\sigma_i, t_j', y_{-\{i,j\}}) = H(\sigma_i, t_j, y_{-\{i,j\}}),$$

because $t_j \simeq_{X_{-j}} t_j'$. Since $(t_j, y_{-\{i,j\}}) \in X_{-i}$, we have

$$\text{either} \quad H(s_i, t_j, y_{-\{i,j\}}) = H(\sigma_i, t_j, y_{-\{i,j\}}) \quad \text{or} \quad u_i(s_i, t_j, y_{-\{i,j\}}) < u_i(\sigma_i, t_j, y_{-\{i,j\}}),$$

which together with the two equations above imply that

$$\text{either} \quad H(s_i, t_j', y_{-\{i,j\}}) = H(\sigma_i, t_j', y_{-\{i,j\}}) \quad \text{or} \quad u_i(s_i, t_j', y_{-\{i,j\}}) < u_i(\sigma_i, t_j', y_{-\{i,j\}}),$$

that is,

$$\text{either} \quad H(s_i, y_{-i}) = H(\sigma_i, y_{-i}) \quad \text{or} \quad u_i(s_i, y_{-i}) < u_i(\sigma_i, y_{-i}).$$

Therefore, $s_i \preceq_{Z_{-i}} \sigma_i$ and $Z \xrightarrow[s_i \preceq \sigma_i]{e} W$ as desired. Accordingly, we have

$$Y \xrightarrow[t_j, t_j']{r} W \xleftarrow[s_i \preceq \sigma_i]{e} Z.$$

*Subcase 2.2*: $i = j$ *but* $s_i \neq t_i$. In this case, letting $W$ be $Y$ with $t_i$ replaced by $t_i'$, with similar analysis we have that there exists $\sigma_i' \in \Delta(W_i)$ such that $Y \xrightarrow[t_i, t_i']{r} W \xleftarrow[s_i \preceq \sigma_i']{e} Z$. Indeed, $\sigma_i' = \sigma_i$ if $\sigma_i(t_i) = 0$, and $\sigma_i'$ is obtained from $\sigma_i$ by replacing $t_i$ with $t_i'$ otherwise.

*Subcase 2.3*: $i = j$ *and* $s_i = t_i$. In this case, $Y_i = Z_i \setminus \{t_i'\}$, $Y_{-i} = Z_{-i}$, $\sigma_i \in \Delta(Y_i)$, and $t_i' \preceq_{Z_{-i}} \sigma_i$. Accordingly, letting $W = Y$, we have $Y \xrightarrow[\epsilon]{r} W \xleftarrow[t_i' \preceq \sigma_i]{e} Z$.

*Case 3*: $Y \xleftarrow{r} X \xrightarrow{r} Z$. In this case, letting $W = X$, we have $Y \xrightarrow{r} W \xleftarrow{r} Z$, because the replacement relation is clearly symmetric.

Thus Lemma 3 holds in all cases. □

Lemma 3 guarantees that the set of all sets of strategy profiles and the relation $\longrightarrow$ together satisfy the diamond property. To use this lemma, we need to show that $S \longrightarrow^{*} R$ for all resilient solutions $R$. Notice that this is not directly implied by the definition of resilient solutions, because iterated elimination of DD strategies allows simultaneous elimination of multiple strategies in each step, while the relation $\longrightarrow$ does not allow such operation.[18] Fortunately we have the following lemma.

Lemma 4. *For all resilient solutions* $R$, $S \xrightarrow{e}^{*} R$.

Proof. Let $R^0 = S, R^1, \ldots, R^K = R$ be the elimination order of DD strategies corresponding to $R$. To prove Lemma 4, it suffices to prove that $R^k \xrightarrow{e}^{*} R^{k+1}$ for each $k < K$. We actually prove a more general result, namely:

> for all sets of strategy profiles $X$ and $Y$, if $Y$ is obtained from $X$ by simultaneously eliminating several strategies that are distinguishably dominated within $X$, then $X \xrightarrow{e}^{*} Y$.

To see why this is true, assume that $\ell$ pure strategies are eliminated from $X$ so as to get $Y$, and denote them by $s_{i_1}, \ldots, s_{i_\ell}$. (Notice that these strategies, respectively, belong to players $i_1, \ldots, i_\ell$, some of which may be the same one.) Let $\tau_{i_1}, \ldots, \tau_{i_\ell}$ be the mixed strategies "responsible for these eliminations," that is, $s_{i_j} \prec_{X_{-i_j}} \tau_{i_j}$ and $\tau_{i_j} \in \Delta(X_{i_j})$ for $j = 1, \ldots, \ell$. According to Lemma 2, we can assume that $\tau_{i_j} \in \Delta(Y_{i_j})$ for each $j$. We prove that $Y$ can be obtained from $X$ by eliminating $s_{i_1}, \ldots, s_{i_\ell}$ one by one, that is, in $\ell$ steps, and in that order. More specifically, letting $X^1 = X$ and $X^{\ell+1} = Y$, and for each $j \in \{2, \ldots, \ell\}$, letting $X^j$ be the set of strategy profiles obtained from $X$ by eliminating $s_{i_1}, \ldots, s_{i_{j-1}}$, we prove that for each $j \leq \ell$,

$$X^j \xrightarrow[s_{i_j} \preceq \tau_{i_j}]{e} X^{j+1}.$$

[18] In principle, problems may arise when eliminating strategies simultaneously. For instance, when a player $i$ eliminates $s_i$ from $R_i$ because $s_i \prec_R \sigma_i$ and there exists a unique $t_{-i} \in R_{-i}$ distinguishing the two, another player $j$ may simultaneously eliminate $t_j$, causing the elimination of $s_i$ to be problematic.

To see why this is true, notice that for each $j \leq \ell$, $Y_{i_j} \subseteq X^{j+1}_{i_j}$ and thus $\tau_{i_j} \in \Delta(X^{j+1}_{i_j})$. Because $s_{i_j} \prec_{X_{-i_j}} \tau_{i_j}$ and $X^j_{-i_j} \subseteq X_{-i_j}$, we have $s_{i_j} \preceq_{X^j_{-i_j}} \tau_{i_j}$. Therefore, $X^j \xrightarrow[s_{i_j} \preceq \tau_{i_j}]{e} X^{j+1}$ for each $j \leq \ell$, which implies that $X \xrightarrow{e}{}^{*} Y$.

Applying this rule to $R^k$ and $R^{k+1}$ for each $k < K$, we have $S \xrightarrow{e}{}^{*} R$ as desired. □

Lemmas 3 and 4 together are enough for us to deduce that for all resilient solutions $R$ and $T$, there exists a set of strategy profiles $W$ such that $R \longrightarrow^{*} W$ and $T \longrightarrow^{*} W$. But to further deduce that $R$ and $T$ are equivalent, we need three additional properties for relations $\xrightarrow[\prec]{e}$, $\xrightarrow[\simeq]{e}$, and $\xrightarrow{r}$, as stated and proved in the following three lemmas.

Lemma 5. *For all sets of strategy profiles $R$ and $X$, if $R \xrightarrow{r} X$ and $R$ is strict-elimination-free, then $X$ is strict-elimination-free.*

Proof. We proceed by contradiction. Assume that $R \xrightarrow{r} X$ and $R$ is strict-elimination-free, yet $X$ is not strict-elimination-free, that is, there exists $T$ such that $X \xrightarrow[t_j \prec \tau_j]{e} T$.

We derive a contradiction by proving that there exists $W$ such that $R \xrightarrow[\prec]{e} W \xrightarrow{r} T$, which implies that $R$ is not strict-elimination-free.

If $R \xrightarrow[\epsilon]{r} X$, then letting $W = T$, we are done immediately, with $R \xrightarrow[t_j \prec \tau_j]{e} W \xrightarrow[\epsilon]{r} T$. Therefore, we assume $R \xrightarrow[s_i, s'_i]{r} X$, that is,

$$R \xrightarrow[s_i, s'_i]{r} X \xrightarrow[t_j \prec \tau_j]{e} T.$$

Because the replacement relation is symmetric, we have

$$R \xleftarrow[s'_i, s_i]{r} X \xrightarrow[t_j \prec \tau_j]{e} T,$$

which is what we see in Case 2 of Lemma 3, with notations changed (in particular, $Z$ becomes $R$, $Y$ becomes $T$, $\preceq$ becomes $\prec$, and $i$ and $j$ are exchanged). We consider three cases here.

*Case 1*: $i \neq j$. In this case, following Subcase 2.1 of the proof of Lemma 3 and letting $W$ be $R$ with $t_j$ removed, we have $R \xrightarrow[t_j \preceq \tau_j]{e} W \xleftarrow[s'_i, s_i]{r} T$. We prove that $t_j \not\simeq_{R_{-j}} \tau_j$, that is, there exists $t_{-j} \in R_{-j}$ such that

$$H(t_j, t_{-j}) \neq H(\tau_j, t_{-j}).$$

To do so, recall that $t_j \prec_{X_{-j}} \tau_j$, which implies that there exists $\hat{t}_{-j} \in X_{-j}$ such that $H(t_j, \hat{t}_{-j}) \neq H(\tau_j, \hat{t}_{-j})$. If $\hat{t}_i \neq s'_i$, then $\hat{t}_{-j} \in R_{-j}$, because $X_{-\{i,j\}} = R_{-\{i,j\}}$ and $X_i \setminus \{s'_i\} = R_i \setminus \{s_i\} \subseteq R_i$. Letting $t_{-j} = \hat{t}_{-j}$, we are done. Otherwise, we have $\hat{t}_i = s'_i$ and

$$H(t_j, s'_i, \hat{t}_{-\{i,j\}}) \neq H(\tau_j, s'_i, \hat{t}_{-\{i,j\}}).$$

Let $t_i = s_i$ and $t_{-\{i,j\}} = \hat{t}_{-\{i,j\}}$. On one hand, we have $t_{-j} \in R_{-j}$. On the other hand, we have $s_i \simeq_{R_{-i}} s'_i$, which implies that

$$H(t_j, s_i, t_{-\{i,j\}}) = H(t_j, s'_i, t_{-\{i,j\}}) \quad \text{and} \quad H(\tau_j, s_i, t_{-\{i,j\}}) = H(\tau_j, s'_i, t_{-\{i,j\}}).$$

Because the right-hand sides of the two equations are not equal, the left-hand sides are not equal either. That is, $H(t_j, s_i, t_{-\{i,j\}}) \neq H(\tau_j, s_i, t_{-\{i,j\}})$ or, equivalently, $H(t_j, t_{-j}) \neq H(\tau_j, t_{-j})$ as desired.

Thus $R \xrightarrow[t_j \prec \tau_j]{e} W \xleftarrow[s'_i, s_i]{r} T$. Again because the replacement relation is symmetric, we have

$$R \xrightarrow[t_j \prec \tau_j]{e} W \xrightarrow[s_i, s'_i]{r} T.$$

*Case 2*: $i = j$, $s'_i \neq t_i$. In this case, following Subcase 2.2 of the proof of Lemma 3 and letting $W$ be $R$ with $t_i$ removed, we have $R \xrightarrow[t_i \preceq \tau'_i]{e} W \xleftarrow[s'_i, s_i]{r} T$. In particular, $\tau'_i = \tau_i$ if $\tau_i(s'_i) = 0$, and $\tau'_i$ is obtained from $\tau_i$ by replacing $s'_i$ with $s_i$ otherwise.

Again we prove that $t_i \not\simeq_{R_{-i}} \tau'_i$. To do so, notice that $\tau'_i$ is either $\tau_i$ itself or obtained from $\tau_i$ by replacing $s'_i$ with $s_i$ such that $s_i \simeq_{R_{-i}} s'_i$. Therefore, we have $\tau'_i \simeq_{R_{-i}} \tau_i$. Because $t_i \not\simeq_{X_{-i}} \tau_i$ and $R_{-i} = X_{-i}$, we have $t_i \not\simeq_{R_{-i}} \tau'_i$ and thus $R \xrightarrow[t_i \prec \tau'_i]{e} W \xleftarrow[s'_i, s_i]{r} T$. By symmetry we have

$$R \xrightarrow[t_i \prec \tau'_i]{e} W \xrightarrow[s_i, s'_i]{r} T.$$

*Case 3*: $i = j$ *and* $s'_i = t_i$. In this case, $T$ is obtained from $R$ by first replacing $s_i$ with $s'_i$, and then eliminating $s'_i$ because $s'_i \prec_X \tau_i$. Therefore, the elimination can be done directly without any replacement. That is, letting $W = T$, $W$ can be obtained from $R$ by eliminating $s_i$, because $s_i \prec_R \tau_i$. Accordingly, $R \xrightarrow[s_i \prec \tau_i]{e} W \xrightarrow[\epsilon]{r} T$.

In sum, Lemma 5 follows. □

Lemma 6. *For all sets of strategy profiles $R$ and $X$, if $R \xrightarrow[\simeq]{e} X$ and $R$ is strict-elimination-free, then $X$ is strict-elimination-free.*

PROOF. We actually prove a more general result, namely,

$$\text{if } T, W, \text{ and } Y \text{ are sets of strategy profiles such that } T \xrightarrow[s_i \simeq s'_i]{e} W \xrightarrow[t_j \prec \tau_j]{e} Y,$$

$$\text{then there exists a set of strategy profiles } Z \text{ such that } T \xrightarrow[t_j \prec \tau_j]{e} Z \xrightarrow[s_i \preceq \sigma'_i]{e} Y.$$

That is, if an indistinguishable elimination is followed by a strict elimination, then we can exchange these two eliminations.

To see why this is true, notice that, by definition, we have $s_i \simeq_{T_{-i}} s'_i$ and $t_j \prec_{W_{-j}} \tau_j$. Because the only change from $W$ to $T$ is that a strategy $s_i$ (which is equivalent to some present ones) is added, we have $t_j \prec_{T_{-j}} \tau_j$. Because $\tau_j \in \Delta(Y_j)$ and $Y_j = W_j \setminus \{t_j\} \subseteq T_j \setminus \{t_j\}$, we have $\tau_j \in \Delta(T_j \setminus \{t_j\})$. Accordingly, letting $Z$ be the set of strategy profiles obtained from $T$ by removing $t_j$, we have

$$T \xrightarrow[t_j \prec \tau_j]{e} Z$$

and that $Y$ is obtained from $Z$ by removing $s_i$.

Now we construct $\sigma'_i$ as follows. If $j \neq i$ or if $j = i$ but $t_i \neq s'_i$, then letting $\sigma'_i = s'_i$, we have $\sigma'_i \in \Delta(Z_i \setminus \{s_i\})$ and $s_i \simeq_{Z_{-i}} \sigma'_i$. Otherwise (that is, $j = i$ and $t_i = s'_i$), letting $\sigma'_i = \tau_i$, we have $\sigma'_i \in \Delta(Z_i \setminus \{s_i\})$ and $s_i \prec_{Z_{-i}} \sigma'_i$. Accordingly, we have $Z \xrightarrow[s_i \preceq \sigma'_i]{e} Y$.

Given this general result, Lemma 6 follows easily. Indeed, if $X$ is not strict-elimination-free, then there exists $W$ such that $R \xrightarrow[\simeq]{e} X \xrightarrow[\prec]{e} W$, which implies that there exists $Z$ such that $R \xrightarrow[\prec]{e} Z \xrightarrow{e} W$, contradicting the fact that $R$ is strict-elimination-free. □

LEMMA 7. *For all sets of strategy profiles $R$ and $X$, if $R \xrightarrow{\simeq} X$, more generally if $R \xrightarrow{\simeq}{}^{*} X$, then $R$ and $X$ are equivalent.*

PROOF. Since equivalence between sets of strategy profiles is clearly reflexive, symmetric, and transitive, it suffices to prove that $R \xrightarrow{\simeq} X$ implies that $R$ and $X$ are equivalent. To do so, we first prove that $R \xrightarrow{r} X$ implies that $R$ and $X$ are equivalent. To this end, notice that if $R \xrightarrow[\epsilon]{r} X$, then $R$ and $X$ are trivially equivalent (since they are equal). Now let $R \xrightarrow[s_i, t_i]{r} X$. Then the profile of functions required by the equivalence relation is simply the profile $\phi$ such that $\phi_i(s_i^{\simeq_{R_{-i}}}) = t_i^{\simeq_{X_{-i}}}$, $\phi_i(a_i^{\simeq_{R_{-i}}}) = a_i^{\simeq_{X_{-i}}}$ for each $a_i \not\simeq_{R_{-i}} s_i$, and for each $j \neq i$ and each strategy $s_j$, $\phi_j(s_j^{\simeq_{R_{-j}}}) = s_j^{\simeq_{X_{-j}}}$. To prove that $R \xrightarrow[\simeq]{e} X$ implies that $R$ and $X$ are equivalent, we can construct a similar profile of functions. □

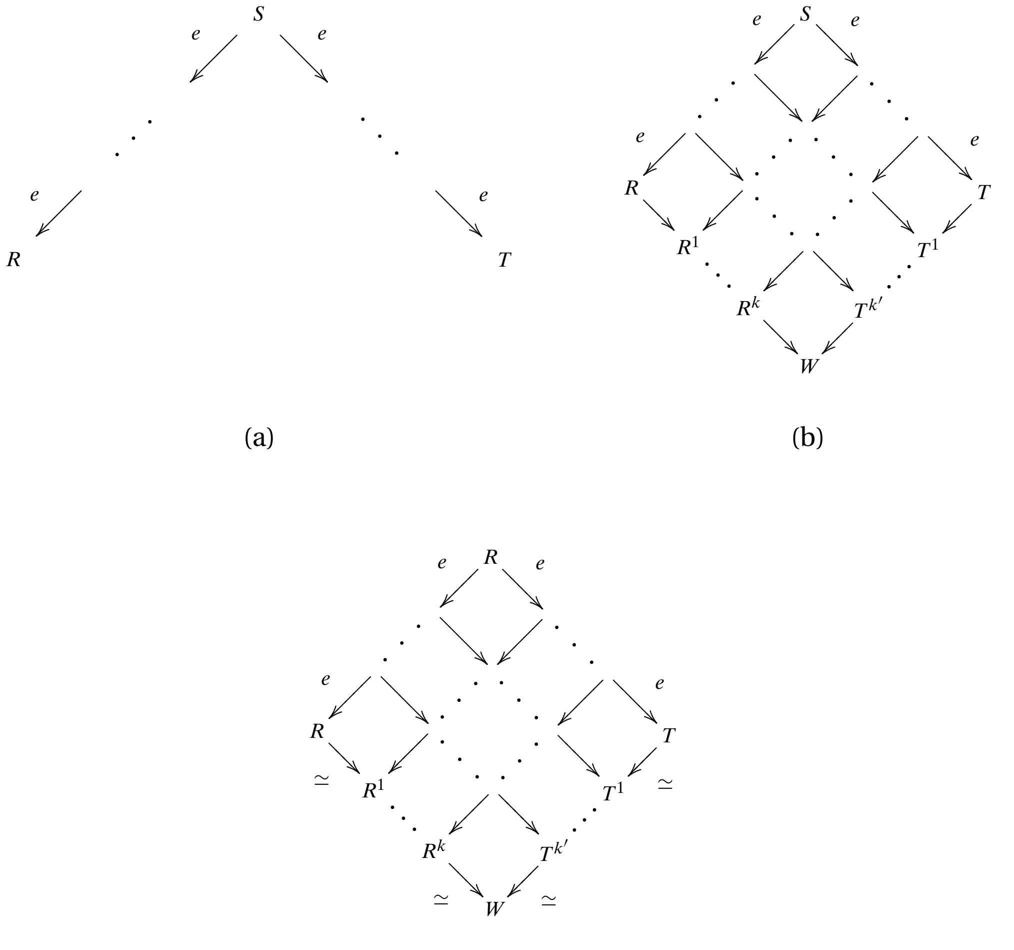



FIGURE 1. Proof of Theorem 1.

### C.3 *Proof of Theorem 1*

At this point we can easily prove our main theorem. Let $R$ and $T$ be two resilient solutions. According to Lemma 4, we have $S \xrightarrow{e}{}^{*} R$ and $S \xrightarrow{e}{}^{*} T$. Pictorially, we have Figure 1(a).[19] By applying Lemma 3 repeatedly, starting from $S$, there exists a set of strategy profiles $W$ such that $R \longrightarrow{}^{*} W$ and $T \longrightarrow{}^{*} W$. Pictorially, we have Figure 1(b). Since $R$ is strict-elimination-free, we have $R \xrightarrow{\simeq} R^1$. Then Lemmas 5 and 6 imply that $R^1$ is also strict-elimination-free. By continued usage of Lemmas 5 and 6, we have $R \xrightarrow{\simeq}{}^{*} W$ and $T \xrightarrow{\simeq}{}^{*} W$, as illustrated by Figure 1(c).

[19]Without loss of generality, in Figure 1, we assume that there are at least two steps from $S$ to $R$ and to $T$.

Finally, accordingly to Lemma 7, $R$ and $W$ are equivalent, and so are $T$ and $W$. Because the equivalence relation between sets of strategy profiles is reflexive, symmetric, and transitive, $R$ and $T$ are equivalent, as desired. □

## Appendix D: Additional properties

Below we prove two additional properties of the equivalence relation between sets of strategy profiles. Proposition 3 guarantees that the required profile $\phi$ in the definition of equivalence between sets of strategy profiles $R$ and $T$, if it exists, is unique, and that $\phi$ maps each strategy that appears in both $R$ and $T$ to itself.

Proposition 3. *For all sets of strategy profiles $R$ and $T$, and all profiles of functions $\phi$ and $\psi$, if $R$ and $T$ are equivalent under both $\phi$ and $\psi$, then $\phi = \psi$. Moreover, for all players $i$ and strategies $s_i \in R_i \cap T_i$, $\phi_i(s_i^{\simeq R_{-i}}) = s_i^{\simeq T_{-i}}$.*

Proof. Proving $\phi = \psi$ is equivalent to proving that for all $i$ and $s_i \in R_i$, $\phi_i(s_i^{\simeq R_{-i}}) = \psi_i(s_i^{\simeq R_{-i}})$. Arbitrarily fixing $i$ and $s_i$, and arbitrarily fixing $a_i \in \phi_i(s_i^{\simeq R_{-i}})$ and $b_i \in \psi_i(s_i^{\simeq R_{-i}})$, it suffices to prove that $a_i \simeq_{T_{-i}} b_i$. To do so, $\forall t_{-i} \in T_{-i}$, let the strategy subprofile $s_{-i} \in R_{-i}$ be such that $s_j \in \phi_j^{-1}(t_j^{\simeq T_{-j}})$ $\forall j \neq i$ and let the strategy subprofile $t'_{-i} \in T_{-i}$ be such that $t'_j \in \psi_j(s_j^{\simeq R_{-j}})$ $\forall j \neq i$. Because $R$ and $T$ are equivalent under $\phi$, we have

$$H(s_i, s_{-i}) = H(a_i, t_{-i}).$$

Because $R$ and $T$ are equivalent under $\psi$, we have

$$H(s_i, s_{-i}) = H(b_i, t'_{-i}).$$

Accordingly, we have $H(a_i, t_{-i}) = H(b_i, t'_{-i})$, which implies that $H(a_i, t_{-i}) = H(b_i, t_{-i})$ by Fact 1 of Section 3 (for games with perfect recall). Therefore, $a_i \simeq_{T_{-i}} b_i$ and we have $\phi = \psi$.

To prove the remaining part, arbitrarily fixing $i$, $s_i \in R_i \cap T_i$, and $t_i \in \phi_i(s_i^{\simeq R_{-i}})$, it suffices to prove that $s_i \simeq_{T_{-i}} t_i$. Again $\forall t_{-i} \in T_{-i}$, let $s_{-i}$ be such that $s_j \in \phi_j^{-1}(t_j^{\simeq T_{-j}})$ $\forall j \neq i$. Because $R$ and $T$ are equivalent under $\phi$, we have $H(s_i, s_{-i}) = H(t_i, t_{-i})$. By Fact 1 of Section 3, this implies that $H(s_i, t_{-i}) = H(t_i, t_{-i})$. Therefore, $s_i \simeq_{T_{-i}} t_i$ and $\phi_i(s_i^{\simeq R_{-i}}) = s_i^{\simeq T_{-i}}$. □

Proposition 4 guarantees that the union of two equivalent sets of strategy profiles is still equivalent to each one of them, with the desired profile of functions being naturally defined. This property helps to establish another connection between resilient solutions and EFR.

Proposition 4. *For all sets of strategy profiles $R$ and $T$, and all profiles of functions $\phi$ such that $R$ and $T$ are equivalent under $\phi$, letting $R \cup T = (R_1 \cup T_1, \ldots, R_n \cup T_n)$, we have* *that $R \cup T$ and $T$ are equivalent under a profile of functions $\psi$. Moreover, for each player $i$, $(R_i \cup T_i)^{\simeq (R\cup T)_{-i}} = \{s_i^{\simeq R_{-i}} \cup \phi_i(s_i^{\simeq R_{-i}}) : s_i \in R_i\}$ and $\psi_i(s_i^{\simeq R_{-i}} \cup \phi_i(s_i^{\simeq R_{-i}})) = \phi_i(s_i^{\simeq R_{-i}})$ for each $s_i \in R_i$.*

The proof is done by repeatedly applying the definition of equivalence between $R$ and $T$ and Fact 1 of Section 3, so it is omitted here.

Definition 10. We denote by $\mathbb{SR}$ the set of strategy profiles such that, for all strategies $s_i$ of a player $i$, $s_i \in \mathbb{SR}_i$ if and only if there exists a resilient solution $R$ such that $s_i \in R_i$.

In a sense, $\mathbb{SR}$ is the union of all resilient solutions. Theorem 1 and Proposition 4 together immediately imply the following connection between resilient solutions and $\mathbb{ER}$, whose proof is omitted.

Corollary 2. *The set $\mathbb{SR}$ is equivalent to every resilient solution (and thus to $\mathbb{ER}$).*

Let us emphasize that $\mathbb{SR}$ may happen to be a resilient solution, but *need not be* one. Recall the game $G_3$ of Example 4, where $\mathbb{ER} = \{bc, bd\} \times \{f\}$ and $T = \{bc, bd\} \times \{e\}$ are two distinct resilient solutions. For $G_3$, it is easy to verify that the only resilient solution different from the above two is $R = \{bc, bd\} \times \{e, f\}$; that is, the strategies $ac$, $ad$, and $g$ never survives any elimination order. (For instance, another elimination order is $g$ followed by a simultaneous elimination of $ac$ and $ad$, yielding $R$.) Therefore, $\mathbb{SR} = \{bc, bd\} \times \{e, f\} = R$. Recall now game $G_5$ of Example 5. For $G_5$, it is easy to verify that $\mathbb{SR} = \{a\} \times \{c, d\} \times \{e, f\}$, which is not a resilient solution itself. Yet it is also easy to verify that $\mathbb{ER} = \{a\} \times \{c\} \times \{f\}$ and that $\mathbb{SR}$ is equivalent to $\mathbb{ER}$.

## References


Apt, Krzysztof R. (2004), "Uniform proofs of order independence for various strategy elimination procedures." *Contributions to Theoretical Economics*, 4, Article 5. [128]

Battigalli, Pierpaolo (1997), "On rationalizability in extensive games." *Journal of Economic Theory*, 74, 40–61. [126, 128, 129]

Battigalli, Pierpaolo and Amanda Friedenberg (2012), "Forward induction reasoning revisited." *Theoretical Economics*, 7, 57–98. [132]

Battigalli, Pierpaolo and Marciano Siniscalchi (2002), "Strong belief and forward induction reasoning." *Journal of Economic Theory*, 106, 356–391. [129]

Bernheim, B. Douglas (1984), "Rationalizable strategic behavior." *Econometrica*, 52, 1007–1028. [126]

Böhm, Corrado and Silvio Micali (1980), "Minimal forms in $\lambda$-calculus computations." *Journal of Symbolic Logic*, 45, 165–171. [139]

Brandenburger, Adam and Amanda Friedenberg (2011), "The relationship between rationality on the matrix and the tree." Unpublished paper. [128]

Chen, Jing and Silvio Micali (2012), "The order independence of iterated dominance in extensive games, with connections to mechanism design and backward induction." Technical Report 2012-023, Computer Science and Artificial Intelligence Laboratory, MIT. [129]

Church, Alonzo and J. Barkley Rosser (1936), "Some properties of conversion." *Transactions of the American Mathematical Society*, 39, 472–482. [139]

Dufwenberg, Martin and Mark Stegeman (2002), "Existence and uniqueness of maximal reductions under iterated strict dominance." *Econometrica*, 70, 2007–2023. [126]

Gilboa, Itzhak, Ehud Kalai, and Eitan Zemel (1990), "On the order of eliminating dominated strategies." *Operation Research Letters*, 9, 85–89. [126, 139]

Gretlein, Rodney J. (1982), "Dominance solvable voting schemes: A comment." *Econometrica*, 50, 527–528. [128]

Gretlein, Rodney J. (1983), "Dominance elimination procedures on finite alternative games." *International Journal of Game Theory*, 12, 107–113. [128]

Huet, Gérard P. (1980), "Confluent reductions: Abstract properties and applications to term rewriting systems." *Journal of ACM*, 27, 797–821. [139]

Klop, Jan W. (1992), "Term rewriting systems." In *Handbook of Logic in Computer Science*, volume 2 (Samson Abramsky, Dov M. Gabbay, and Thomas S. E. Maibaum, eds.), 1–116, Clarendon Press, Oxford. [139, 140]

Marx, Leslie M. and Jeroen M. Swinkels (1997), "Order independence for iterated weak dominance." *Games and Economic Behavior*, 18, 219–245. [127, 128, 143, 144]

Moulin, Hervé (1979), "Dominance solvable voting schemes." *Econometrica*, 47, 1337–1351. [128]

Osborne, Martin J. and Ariel Rubinstein (1994), *A Course in Game Theory*. MIT Press, Cambridge, Massachusetts. [126, 130, 132]

Pearce, David G. (1984), "Rationalizable strategic behavior and the problem of perfection." *Econometrica*, 52, 1029–1050. [126]

Perea, Andrés (2011), "Belief in the opponents' future rationality." Unpublished paper, Maastricht University. [138]

Reny, Philip J. (1992), "Backward induction, normal form perfection and explicable equilibria." *Econometrica*, 60, 627–649. [133]

Robles, Jack (2006), "Order independence of conditional dominance." Unpublished paper. [128]

Rochet, Jean-Charles (1980), *Selection of an Unique Equilibrium Pay Off for Extensive Games With Perfect Information*. Ph.D. thesis, Université de Paris IX Dauphine. [128]

Shimoji, Makoto (2004), "On the equivalence of weak dominance and sequential best response." *Games and Economic Behavior*, 48, 385–402. [128]

Shimoji, Makoto and Joel Watson (1998), "Conditional dominance, rationalizability, and game forms." *Journal of Economic Theory*, 83, 161–195. [125, 126, 134, 136]

---

Submitted 2011-2-14. Final version accepted 2012-1-18. Available online 2012-1-18.