

# More Reasoning, Fewer Outcomes: A Monotonicity Result for Reasoning in Dynamic Games\*

Andrés Perea<sup>†</sup>  
Maastricht University

Revised version: August 2026

## Abstract

A focus function in a dynamic game describes, for every player and each of the information sets, the collection of opponents' information sets he reasons about. Every focus function induces a rationalizability procedure in which a player believes, whenever possible, that his opponents choose rationally at those information sets he reasons about. Under certain conditions, we show that if the players start reasoning about more information sets, then the set of outcomes induced by the associated rationalizability procedure becomes smaller or stays the same. This result does not hold on the level of strategies, unless the players only reason about present and future information sets. The monotonicity result enables us to derive existing theorems, such as the relation in terms of outcomes between forward and backward induction reasoning, but also paves the way for new results.

*JEL Classification:* C72

*Keywords:* Dynamic games, reasoning, rationalizability, forward induction, backward induction, epistemic priority, focus function

## 1 Introduction

To reach a meaningful decision in a dynamic game it is important to reason, at each of your information sets, about the opponents' rationality at some, or all, of their information sets. This is reflected in many equilibrium and rationalizability concepts for dynamic games. Consider, for instance, the forward induction concept of *strong rationalizability*

---

\*I wish to thank the co-editor, three anonymous referees, Emiliano Catonini, Bram Driesen, Arkadi Predtetchinski, Elias Tsakas, and three anonymous reviewers at the TARK 2025 conference for their valuable feedback.

<sup>†</sup>*E-mail:* a.perea@maastrichtuniversity.nl *Affiliation:* Maastricht University, Department of Quantitative Economics and EpiCenter, The Netherlands. *Web:* [www.epicenter.name/Perea/](http://www.epicenter.name/Perea/)

(Pearce (1984), Battigalli (1997)), also known as *extensive-form rationalizability*, in which a player must believe, whenever possible, that his opponents choose rationally at all of their information sets. It thus requires the player to always reason about *all* the opponents' information sets. Compare this to the backward induction concepts of *backward dominance* (Perea (2014)) and *backwards rationalizability*<sup>1</sup> (Penta (2015), Catonini and Penta (2026), Perea (2014)) in which a player must always believe that his opponents will choose rationally at all present and future information sets. It only requires a player to actively reason about the *present* and *future* information sets of his opponents.

As strong rationalizability requires reasoning about more information sets than backward dominance, does this mean that the former concept will also be more restrictive than the latter? From Perea (2017), Catonini (2020) and the arguments in Chen and Micali (2013) we know that this is only true in terms of *outcomes*, but not in terms of *strategies*: Every outcome that is possible under strong rationalizability is also possible under backward dominance, but there may be strongly rationalizable strategies that do not survive the backward dominance procedure.<sup>2</sup>

This naturally raises the question whether, in general, reasoning about more information sets will always lead to a smaller (or equal) set of outcomes. The goal of this paper is to establish conditions under which such an outcome monotonicity result indeed holds. To this purpose we introduce a family of rationalizability procedures parametrized by the collections of opponents' information sets that the players reason about at the various information sets. Formally, the parameter is a *focus function*  $f$  (Perea and Tsakas (2019)) which specifies for every player  $i$  and every information set  $h$ , the collection of opponents' information sets that  $i$  reasons about while being at  $h$ . The associated belief restriction is that every player believes, whenever possible, that his opponents choose rationally at those information sets he reasons about. By iterating this condition we obtain a recursive reduction procedure that we call *f-rationalizability*.

It turns out that many of the existing rationalizability concepts in the literature, including strong rationalizability, backward dominance, and the Ben-Porath procedure (Ben-Porath (1997)), can be phrased as *f-rationalizability* concepts with an appropriately

---

<sup>1</sup>The difference between the two concepts is that the latter requires the beliefs to satisfy *forward consistency* (also referred to as Bayesian updating), whereas the former does not. The term forward consistency is adopted from Battigalli, Catonini and Manili (2023).

<sup>2</sup>Catonini (2020) shows this result for the stronger concept of backwards rationalizability.

chosen focus function  $f$ . But we can also use the notion of  $f$ -rationalizability to introduce new interesting reasoning concepts for dynamic games.

In Theorem 4.1 we prove the existence of  $f$ -rationalizable strategies under appropriate conditions on the focus function  $f$ . The conditions are the following: (a)  $f$  must be *individually forward decreasing*, which means that as the game moves on, a player should not start to reason about information sets he did not reason about before; and (b)  $f$  must *individually preserve focus on past information sets*, which means that if a player reasons, at a particular information set, about a past information set  $h$ , then he will still reason about  $h$  at all future information sets. The conditions above rule out the possibility of dynamically inconsistent beliefs, which would endanger the existence of  $f$ -rationalizable strategies. Many rationalizability concepts for dynamic games in the literature satisfy these conditions, and their existence is thus guaranteed by Theorem 4.1.

Our main result, Theorem 5.1, compares two focus functions  $f$  and  $g$  such that  $f$  requires reasoning about more information sets than  $g$ , and shows that under certain conditions every outcome induced by a combination of  $f$ -rationalizable strategies is also induced by a combination of  $g$ -rationalizable strategies. The proof is constructive, as it explicitly shows how to transform a combination of  $f$ -rationalizable strategies into a combination of  $g$ -rationalizable strategies that yields the same outcome. The theorem also shows that if in  $f$  the players only reason about present and future information sets, then every  $f$ -rationalizable strategy is  $g$ -rationalizable. That is, in this case the set inclusion even holds on the level of strategies.

In fact, Theorem 5.1 states something stronger than the above monotonicity result, by showing that the outcomes induced by combinations of  $f$ -rationalizable strategies are the *same* as the outcomes that result if we first apply the  $g$ -rationalizability procedure and then the  $f$ -rationalizability procedure. In the latter combined procedure we thus give epistemic priority to reasoning according to  $g$  above reasoning according to  $f$ . Again, the proof is constructive as it shows how to transform a combination of strategies surviving the first procedure into a combination of strategies surviving the second combined procedure yielding the same outcome, and *vice versa*.

The conditions under which the theorem holds may be viewed as “interactive versions” of the “individual” conditions (a) and (b) above. First, we require that  $f$  must be *forward decreasing*, which means that as the game proceeds a player  $i$  will never start reasoning

about an information set by player  $j$  that a player  $k$  moving before him (possibly  $k = i$ ) did not reason about. We also require  $f$  to *preserve focus on past information sets*, which means that if a player reasons, at a particular information set, about a past information set  $h$ , then the players at later information sets will still reason about  $h$ . Together, these two conditions thus state that the reasoning patterns of the different players should be “sufficiently aligned”. For the other focus function  $g$  it is sufficient to require that it is *individually* forward decreasing and *individually* preserving focus on past information sets.

Theorem 5.1 can be used to establish existing results, such as the inclusion in terms of outcomes between strong rationalizability and backward dominance, as discussed above. This result, in turn, implies Battigalli’s theorem in Battigalli (1997), which shows that for all games with perfect information where there are no relevant ties<sup>3</sup>, the unique outcome under strong rationalizability is the backward induction outcome. Alternative proofs for Battigalli’s theorem can be found in Heifetz and Perea (2015) and Perea (2018). Reny (1992), in Proposition 3, has shown a similar result for the alternative forward induction concept of *explicable equilibrium*.

However, Theorem 5.1 also implies new interesting results that cannot be derived based on existing theorems or arguments. Consider, for instance, a focus function  $f$  where a player always reasons, at every information set, about the opponents’ past information sets in  $F$  and the opponents’ present and future information sets in  $F'$ , where  $F$  and  $F'$  are some fixed collections of information sets, and  $F$  is included in  $F'$ . Here,  $F$  and  $F'$  may be viewed as “focal” information sets that seem particularly relevant to the players in the dynamic game at hand, for instance because mistakes at these information sets would have drastic consequences compared to other information sets. The induced  $f$ -rationalizability procedure represents a version of *focal forward induction reasoning* in which every player always attempts to rationalize the opponents’ past choices at  $F$ , and uses this to draw inferences about the opponents’ future behavior at  $F'$ . Now, compare  $f$  to a focus function  $g$  that is similarly generated by collections of information sets  $G$  and  $G'$ . Theorem 5.1 implies that, whenever  $G$  and  $G'$  are included in  $F$  and  $F'$ , respectively, then every outcome that is possible under  $f$ -rationalizability will also be possible under  $g$ -rationalizability. See Corollary 6.4.

---

<sup>3</sup>No relevant ties means that for every player, different actions always lead to different utilities for that player.

In intuitive terms, this result thus states that when players attempt to rationalize more instances of past behavior, or draw more inferences about future behavior from observed past behavior, then the set of possible outcomes will become smaller or stay the same. Note that backward dominance and strong rationalizability constitute extreme cases of focal forward induction reasoning, where the former is induced by taking  $F$  equal to the empty set and  $F'$  equal to the collection of all information sets, whereas the latter is obtained by taking both  $F$  and  $F'$  equal to the collection of all information sets. However, if we consider focal forward induction concepts where  $F$  and  $F'$  contain some, but not all, information sets in the game, then the result in Corollary 6.4 does not follow from existing theorems or arguments.

The outline of the paper is as follows. In Section 2 we present our model of a dynamic game, together with some derived objects such as strategies, strict dominance, beliefs, and expected utility. In Section 3 we lay out the notion of a focus function, explain how it gives rise to a family of rationalizability procedures, and show how many of the existing rationalizability procedures can be modelled as members of this family. In Section 4 we present the conditions under which we can guarantee the existence of rationalizable strategies. In Section 5 we present our main result, Theorem 5.1, and provide a sketch of the proof. In Section 6 we show how this theorem can be used to prove existing and new results. In Section 7 we provide some concluding remarks. All the proofs can be found in Section 8. The working paper version Perea (2024) contains some additional material. In particular, it discusses why the concept of  $f$ -rationalizability does not change if we impose *forward consistency* (also known as Bayesian updating) on the beliefs of a player, and require this player to believe that his opponents' beliefs are forward consistent at those information sets he reasons about.

## 2 Games, Strategies and Strict Dominance

### 2.1 Dynamic Games

A *finite dynamic game with observable actions* (or simply *dynamic game*) is a tuple  $\Gamma = (I, (A_i^*, A_i, u_i)_{i \in I})$ , where

- (a)  $I$  is the finite set of *players*, and where for every player  $i$
- (b)  $A_i^*$  is the finite set of *actions* for player  $i$ . By  $\bar{A}$  we denote the set of *action sequences*,

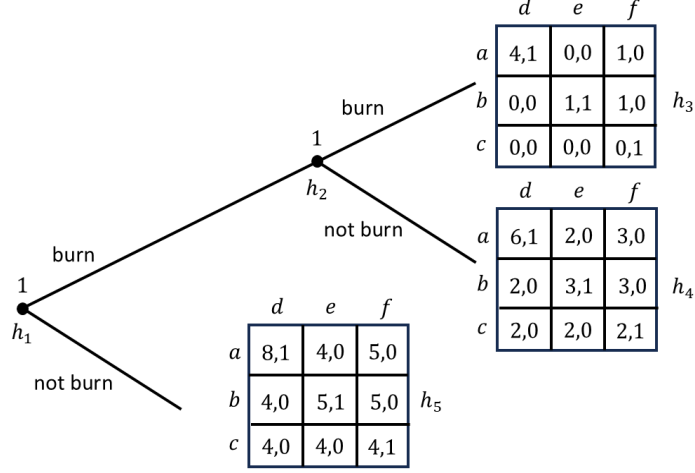


Figure 1: Two-step burning money game

containing the empty sequence  $\emptyset$ , and the finite sequences  $\bar{a} = (a^1, \dots, a^n)$ , where  $n \in \mathbb{N}$  and  $a^m = (a_i^m)_{i \in I} \in \times_{i \in I} A_i^*$  for every  $m \in \{1, \dots, n\}$ ;

(c) the mapping  $A_i$  assigns to every action sequence  $\bar{a} \in \bar{A}$  the nonempty set  $A_i(\bar{a}) \subseteq A_i^*$  of *available actions* for player  $i$  after  $\bar{a}$ . Player  $i$  is said to be *active* at  $\bar{a}$  if  $|A_i(\bar{a})| \geq 2$ , and is said to be *inactive* otherwise. The non-empty action sequence  $\bar{a} = (a^1, \dots, a^n) \in \bar{A}$  is called a *history* if (i)  $a_i^1 \in A_i(\emptyset)$  for every player  $i$ , (ii)  $a_i^m \in A_i(a^1, \dots, a^{m-1})$  for every  $m \in \{2, \dots, n\}$  and every player  $i$ , and (iii) there is at least one active player at  $\emptyset$ , and at every subsequence  $(a^1, \dots, a^m)$  with  $m \in \{1, \dots, n-1\}$ . The empty action sequence  $\emptyset$  is also considered to be a history. The history  $\bar{a}$  is *non-terminal* if there is at least one active player at  $\bar{a}$ , and is called a *terminal history* (or *outcome*) otherwise. By  $H$  and  $Z$  we denote the set of non-terminal and terminal histories, respectively. We assume that the mappings  $(A_i)_{i \in I}$  are such that the sets  $H$  and  $Z$  are finite; and

(d) the mapping  $u_i : Z \rightarrow \mathbb{R}$  is player  $i$ 's *utility function*.

An illustration of a dynamic game can be found in Figure 1, where the set of players is  $I = \{1, 2\}$ , the non-terminal histories are  $h_1 = \emptyset$ ,  $h_2 = \text{burn}$ ,  $h_3 = (\text{burn}, \text{burn})$ ,  $h_4 = (\text{burn}, \text{not burn})$  and  $h_5 = (\text{not burn})$ , only player 1 is active at  $h_1$  and  $h_2$ , both players are active at  $h_3, h_4$  and  $h_5$ , and there are 27 terminal histories. The story behind the game is the following: At the beginning, player 1 must decide whether to burn two units of money

or not. If he does not burn any money, he faces a simultaneous move game with player 2 in which player 1's actions are in the rows and player 2's actions are in the columns. If he burns money at the beginning, he enters the second step of the game in which he must decide whether to burn two additional units of money or not. In both cases, he will face the same simultaneous move game with player 2 as above, but with two or four units of money less for player 1, depending on whether he burned money once or twice. If the two players are firms, then the event of burning money may be viewed as a metaphor for incurring sunk costs.

In a general dynamic game, we denote for a given player  $i$  by  $H_i$  the set of non-terminal histories at which  $i$  is active. We assume that at every such history  $\bar{a} \in H_i$ , player  $i$  knows which actions have been played in  $\bar{a}$ . As such, the collection of *information sets* for player  $i$  coincides with  $H_i$ . We say that a non-terminal history  $h \in H$  *precedes* a terminal or non-terminal history  $h' = (a^1, \dots, a^n)$  with  $n \in \mathbb{N}$  or, equivalently,  $h'$  *follows*  $h$ , if  $h = \emptyset$  or  $h = (a^1, \dots, a^m)$  for some  $m \in \{1, \dots, n-1\}$ . We say that  $h$  *weakly precedes*  $h'$  (or  $h'$  *weakly follows*  $h$ ) if either  $h$  precedes  $h'$ , or  $h = h'$ .

For an information set  $h \in H$ , we denote by  $H^+(h)$  and  $H^-(h)$  the collections of information sets that weakly follow  $h$ , and precede  $h$ , respectively. By  $H(h) := H^+(h) \cup H^-(h)$  we denote the collection of information sets that weakly follow or precede  $h$ . For a given player  $i$  we define  $H_i^+(h) := H^+(h) \cap H_i$  and  $H_i^-(h) := H^-(h) \cap H_i$ . For a collection  $\hat{H}_i \subseteq H_i$  we denote by  $\hat{H}_i^{first}$  the collection of information sets in  $\hat{H}_i$  that are not preceded by any other information set in  $\hat{H}_i$ . We say that a collection  $\hat{H} \subseteq H$  of information sets is *closed under weak followers* (*closed under predecessors*) if for every  $h \in \hat{H}$  we have that  $H^+(h) \subseteq \hat{H}$  ( $H^-(h) \subseteq \hat{H}$ ).

## 2.2 Strategies

A strategy<sup>4</sup> for player  $i$  is a mapping  $s_i$  that assigns to every information set  $h_i \in H_i$  some action  $s_i(h_i) \in A_i(h_i)$ . By  $S_i$  we denote the set of strategies for player  $i$ , and by  $S_{-i} := \times_{j \neq i} S_j$  the set of strategy combinations for  $i$ 's opponents. Every strategy profile  $s$  in  $\times_{i \in I} S_i$  induces a unique outcome  $z(s)$ . We say that the strategy profile  $s$  *reaches* an information set  $h$  if  $h$  precedes  $z(s)$ . For a given information set  $h \in H$  we define the set

---

<sup>4</sup>We could alternatively have used the notion of *plan of action* (Rubinstein (1991)) in this paper, as it makes no difference for our analysis. In the examples we always use plans of action instead of strategies.

$S(h) := \{s \in \times_{i \in I} S_i \mid s \text{ reaches } h\}$ , and for a given player  $i$  we denote by  $S_i(h)$  and  $S_{-i}(h)$  the projection of  $S(h)$  on  $S_i$  and  $S_{-i}$ , respectively. For a strategy  $s_i \in S_i$  we denote by  $H_i(s_i) := \{h_i \in H_i \mid s_i \in S_i(h_i)\}$  the collection of information sets for player  $i$  that the strategy  $s_i$  allows to be reached.

A *decision problem* for player  $i$  at an information set  $h_i \in H_i$  is a pair  $(D_i(h_i), D_{-i}(h_i))$ , where  $D_i(h_i) \subseteq S_i(h_i)$  and  $D_{-i}(h_i) \subseteq S_{-i}(h_i)$  are non-empty sets.

### 2.3 Strict Dominance and Optimal Strategies

For a finite set  $X$ , let  $\Delta(X)$  denote the set of probability distributions on  $X$ . For a given player  $i$  and information set  $h_i \in H_i$ , let  $(D_i(h_i), D_{-i}(h_i))$  be a decision problem for player  $i$  at  $h_i$ . A strategy  $s_i \in D_i(h_i)$  is said to be *strictly dominated* in  $(D_i(h_i), D_{-i}(h_i))$  if there is a randomized strategy  $\mu_i \in \Delta(D_i(h_i))$  for player  $i$  such that  $\sum_{s'_i \in D_i(h_i)} \mu_i(s'_i) \cdot u_i(z(s'_i, s_{-i})) > u_i(z(s_i, s_{-i}))$  for all  $s_{-i} \in D_{-i}(h_i)$ .

From Pearce (1984) it is well-known that a strategy is not strictly dominated in a decision problem precisely when it is optimal for some probabilistic belief there. Formally, a (probabilistic) *belief* for player  $i$  is a probability distribution  $b_i \in \Delta(S_{-i})$  on the opponents' strategy combinations. For a given strategy  $s_i \in S_i$  and belief  $b_i \in \Delta(S_{-i})$ , we denote by  $u_i(s_i, b_i) := \sum_{s_{-i} \in S_{-i}} b_i(s_{-i}) \cdot u_i(z(s_i, s_{-i}))$  the *expected utility* induced by strategy  $s_i$  under the belief  $b_i$ . Now, consider a player  $i$ , an information set  $h_i \in H_i$ , a decision problem  $(D_i(h_i), D_{-i}(h_i))$  at  $h_i$ , and a belief  $b_i \in \Delta(D_{-i}(h_i))$ . A strategy  $s_i \in D_i(h_i)$  is said to be *optimal* in  $D_i(h_i)$  for the belief  $b_i$  if  $u_i(s_i, b_i) \geq u_i(s'_i, b_i)$  for all  $s'_i \in D_i(h_i)$ . The following result corresponds to Lemma 3 in Pearce (1984).

**Lemma 2.1 (Pearce's lemma)** *Consider a player  $i$ , an information set  $h_i \in H_i$ , and a decision problem  $(D_i(h_i), D_{-i}(h_i))$  for player  $i$  at  $h_i$ . Then, a strategy  $s_i \in D_i(h_i)$  is not strictly dominated in  $(D_i(h_i), D_{-i}(h_i))$ , if and only if,  $s_i$  is optimal in  $D_i(h_i)$  for some belief  $b_i \in \Delta(D_{-i}(h_i))$ .*

The proof follows from a simple adaptation of the proof in Pearce (1984), and is therefore omitted.

### 3 Rationalizability Procedures

In this section we present a family of rationalizability procedures parametrized by a *focus function* – a mapping that specifies for every player, and each of the information sets, the collection of opponents’ information sets he reasons about.

#### 3.1 Decision Problem Maps

A *decision problem map*  $D$  assigns to every player  $i$  and information set  $h_i \in H_i$  a decision problem  $(D_i(h_i), D_{-i}(h_i))$  for player  $i$  at  $h_i$ . By  $D^{full}$  we denote the map that specifies for every player  $i$  and information set  $h_i \in H_i$  the full decision problem  $(S_i(h_i), S_{-i}(h_i))$ . For two decision problem maps  $D, E$  we write  $D \subseteq E$  if for every player  $i$  and every information set  $h_i \in H_i$  we have that  $D_i(h_i) \subseteq E_i(h_i)$  and  $D_{-i}(h_i) \subseteq E_{-i}(h_i)$ .

#### 3.2 Reduction Operators

A *reduction operator*  $r$  assigns to every decision problem map  $D$  a reduced decision problem map  $r(D) \subseteq D$ . For a given  $D$  we set  $r^0(D) := D$ , and for every  $k \geq 1$  we inductively define by  $r^k(D) := r(r^{k-1}(D))$  the  $k$ -fold application of the reduction operator  $r$  to  $D$ . Suppose that  $K$  is such that  $r^{K+1}(D) = r^K(D)$ . Then, we denote by  $r^\infty(D) := r^K(D)$  the *iterated* application of the reduction operator  $r$  to  $D$ .

#### 3.3 Focus-Based Rationalizability Procedures

We will now consider some special reduction operators for dynamic games in which a player, at each of the information sets, reasons about the rationality of his opponents at *certain* – but not necessarily all – information sets. Based on Perea and Tsakas (2019) we formalize this type of reasoning by a *focus function*.

**Definition 3.1 (Focus function)** A **focus function**  $f$  assigns to every player  $i$ , information set  $h \in H$ , and opponent  $j \neq i$ , a collection of information sets  $f_{ij}(h) \subseteq H_j$ .

Intuitively,  $f_{ij}(h)$  contains the player  $j$  information sets that player  $i$  reasons about<sup>5</sup> when being at information set  $h$ .<sup>6</sup> For two focus functions  $f, g$  we write  $f \subseteq g$  if for every

<sup>5</sup>More formally, player  $i$  believes at  $h$ , whenever possible, that  $j$  chooses rationally at all  $h_j \in f_{ij}(h)$ .

<sup>6</sup>Note that a player is also required to reason at information sets at which he is not active. Our results would not change if we require a player to only reason at information sets at which he is active. However,

player  $i$ , information set  $h \in H$ , and opponent  $j \neq i$ , we have that  $f_{ij}(h) \subseteq g_{ij}(h)$ .

Every focus function  $f$  induces a *focus based reduction operator*  $rf$  as follows. For every decision problem map  $D$ , the output will be the reduced decision problem map  $E = rf(D)$  where

$$D_{-i}^+(h_i, f) := \{(s_j)_{j \neq i} \in D_{-i}(h_i) \mid \text{for all } j \neq i, s_j \in D_j(h_j) \text{ for all } h_j \in f_{ij}(h_i) \cap H_j(s_j)\},$$

$$E_{-i}(h_i) := \begin{cases} D_{-i}^+(h_i, f), & \text{if } D_{-i}^+(h_i, f) \neq \emptyset \\ D_{-i}(h_i), & \text{if } D_{-i}^+(h_i, f) = \emptyset \end{cases}, \text{ and}$$

$$E_i(h_i) := \{s_i \in D_i(h_i) \mid s_i \text{ not strictly dominated in } (D_i(h_i), E_{-i}(h_i))\}$$

for all players  $i$  and information sets  $h_i \in H_i$ . By definition of  $E_{-i}(h_i)$ , player  $i$  believes, whenever possible, at  $h_i$  that every opponent  $j$  is choosing from  $D_j(h_j)$  at all  $h_j \in f_{ij}(h_i)$  that  $i$  reasons about at  $h_i$ . If believing so is not possible at  $h_i$ , then no additional restrictions are imposed on the beliefs at  $h_i$ , and thus  $E_{-i}(h_i) = D_{-i}(h_i)$ . By construction, both sets  $E_i(h_i)$  and  $E_{-i}(h_i)$  are non-empty.

For a given decision problem map  $D$  we call  $(rf^k(D))_{k=0}^\infty$  the *f-rationalizability procedure* starting at  $D$ . Let  $(D^k)_{k=0}^\infty$  be the *f-rationalizability procedure* starting at  $D^{full}$ , and let  $K$  be such that  $D^K = D^{K+1}$ . By construction, the sets  $D_i^K(h_i)$  and  $D_{-i}^K(h_i)$  are always non-empty for every player  $i$  and  $h_i \in H_i$ . For every player  $i$ , a *strategy*  $s_i \in S_i$  is called *f-rationalizable* if  $s_i \in D_i^K(h_i)$  for all  $h_i \in H_i(s_i)$ . An *outcome*  $z \in Z$  is called *f-rationalizable* if for every player  $i$  there is an *f-rationalizable* strategy  $s_i$  such that  $z = z((s_i)_{i \in I})$ .

For two focus functions  $g, f$ , the *(g, f)-rationalizability procedure* is obtained by first performing *g-rationalizability* until the decision problem map cannot be refined further, after which we perform *f-rationalizability*. If we start from  $D^{full}$ , it thus yields  $D^K = (rf^\infty \circ rg^\infty)(D^{full})$  as the final decision problem map. We can then define *(g, f)-rationalizable* strategies and outcomes in the same way as above.

In view of Lemma 2.1, the sets  $E_i(h_i)$  above in the definition of  $rf(D)$  can equivalently be defined as

$$E_i(h_i) := \{s_i \in D_i(h_i) \mid s_i \text{ optimal in } D_i(h_i) \text{ for some belief } b_i \in \Delta(E_{-i}(h_i))\}.$$

---

the present definition facilitates our analysis.

| concept                      | focus function $f$                                                                                                 |
|------------------------------|--------------------------------------------------------------------------------------------------------------------|
| strong rationalizability     | $f_{ij}^{all}(h) = H_j$                                                                                            |
| backward dominance procedure | $f_{ij}^{future}(h) = H_j^+(h)$                                                                                    |
| Ben-Porath procedure         | $f_{ij}^{initial}(\emptyset) = H_j$<br>$f_{ij}^{initial}(h) = \emptyset$ for all $h \in H \setminus \{\emptyset\}$ |

Table 1: Rationalizability concepts and their associated focus functions

With our definition above, the  $f$ -rationalizability procedure is phrased as an iterated strict dominance procedure, whereas the latter definition of  $E_i(h_i)$  would describe it as an iterated never-a-best-reply procedure.<sup>7</sup>

### 3.4 Special Cases

It turns out that many of the existing rationalizability concepts for dynamic games can be phrased as  $f$ -rationalizability, for some appropriately chosen focus function  $f$ . Table 1 illustrates this for strong rationalizability (Pearce (1984), Battigalli (1997)), the backward dominance procedure (Perea (2014)) and the Ben-Porath procedure (1997). In the table, recall that  $\emptyset$  denotes the beginning of the game. The working paper Perea (2024) contains a more detailed discussion of these rationalizability concepts and their associated focus functions.

To illustrate these rationalizability procedures, we will now apply them to the two-step burning money game of Figure 1 and see which strategies and outcomes result. We start with strong rationalizability. Player 2 will believe at information set  $h_4$  that player 1 will not choose  $b$  or  $c$  at  $h_4$ , as not burning any money and subsequently choosing  $a$  would give player 1 more at  $h_1$ . Therefore, player 2 will choose  $d$  at  $h_4$ , and player 1 believes he will get utility 6 for sure if he burns two units of money. But then, player 2 will believe at  $h_5$  that player 1 will not choose  $b$  or  $c$  at  $h_5$ , as it would give player 1 at most 5. As

---

<sup>7</sup>The elimination procedure in Perea and Tsakas (2019) is formulated as an iterated never-a-best-reply procedure. Another difference is that Perea and Tsakas (2019) allow players to entertain belief hierarchies about focus functions in which they, or their opponents, are (believed to be) wrong about the players' actual focus functions, whereas we implicitly assume that the players' actual focus functions are transparent to everyone.

|                          | strategies<br>for player 1                    | strategies<br>for player 2                          | outcomes                                                                            |
|--------------------------|-----------------------------------------------|-----------------------------------------------------|-------------------------------------------------------------------------------------|
| strong rationalizability | $\{(nb, a)\}$                                 | $\{d, e, f\} \times \{d\} \times \{d\}$             | $\{(nb, a, d)\}$                                                                    |
| focal forward induction  | $\{(bu, nb, a), (nb, a)\}$                    | $\{d\} \times \{d\} \times \{d, e\}$                | $\{(bu, nb, a, d)\} \cup \{(nb, a)\} \times \{d, e\}$                               |
| backward dominance       | $\{(bu, nb, a)\} \cup \{nb\} \times \{a, b\}$ | $\{d, e\} \times \{d, e\} \times \{d, e\}$          | $\{(bu, nb, a)\} \times \{d, e\} \cup \{nb\} \times \{a, b\} \times \{d, e\}$       |
| Ben-Porath procedure     | $\{(bu, nb, a)\} \cup \{nb\} \times \{a, b\}$ | $\{d, e, f\} \times \{d, e, f\} \times \{d, e, f\}$ | $\{(bu, nb, a)\} \times \{d, e, f\} \cup \{nb\} \times \{a, b\} \times \{d, e, f\}$ |

Table 2: Strategies and outcomes in the two-step burning money game

such, player 2 will choose  $d$  at  $h_5$  as well. Player 1, anticipating this, would thus not burn any money and choose  $a$ , believing that he will obtain the utility of 8. Hence, the unique outcome under strong rationalizability is  $(not\ burn, a, d)$ , whereas the induced strategies can be found in Table 2. In the table,  $bu$  and  $nb$  stand for *burn* and *not burn*, respectively.

We next turn to backward dominance. Then, player 2 believes at  $h_3, h_4$  and  $h_5$  that player 1 does not choose  $c$  there, as  $c$  is never optimal for any belief at these information sets. Moreover, this is the *only* conclusion that player 2 will draw there, as he need not believe that player 1 chose optimally in the past. Player 2 can therefore either choose  $d$  or  $e$  at these three information sets. Player 1, anticipating this, will then either burn two units of money and choose  $a$ , or not burn any money and choose  $a$  or  $b$  afterwards. The induced outcomes and strategies can be found in Table 2.

Consider now the Ben-Porath procedure. As player 2 is only active at  $h_3, h_4$  and  $h_5$  where he is not required to reason about any information set of player 1, player 2 can hold any belief at these information sets, and can therefore choose any action there. Player 1, anticipating this, will then again either burn two units of money and choose  $a$ , or not burn any money and choose  $a$  or  $b$  afterwards. See Table 2 for the induced outcomes and strategies.

We finally consider a new rationalizability procedure with focus function  $f^{focal}$ , where  $f_1^{focal}(h) = H_2$  and  $f_2^{focal}(h) = H_1 \setminus \{h_1\}$  for all  $h \in H$ . That is, player 1 reasons about all information sets where player 2 is active, whereas player 2 believes, whenever possible,

that player 1 chose rationally at  $h_2$ , but not necessarily at  $h_1$ , in the past, and that player 1 will choose rationally at  $h_3, h_4$  and  $h_5$ . Such reasoning is natural if player 2 expects player 1 to be more careful at  $h_2$  compared to  $h_1$ , since making a mistake at  $h_2$  may yield player 1 some very low payoffs compared to those he could have achieved by not burning money at the beginning. The associated  $f^{focal}$ -rationalizability concept is called *focal forward induction*.

Under focal forward induction, player 2 must believe at  $h_3$  that player 1's strategy was optimal at  $h_2$ . But then, player 2 must believe at  $h_3$  that player 1 will not choose  $b$  or  $c$  at  $h_3$ , as not burning money at  $h_2$  and subsequently choosing  $a$  would always be better. As such, player 2 will choose  $d$  at  $h_3$ . Player 1, anticipating this, will expect a utility of 4 at  $h_2$  if he would burn money there. But then, player 2 must believe at  $h_4$  that player 1 will not choose  $b$  or  $c$  at  $h_4$ , as burning money at  $h_2$  and subsequently choosing  $a$  would always be better. As such, player 2 will choose  $d$  at  $h_4$  as well. Hence, if player 1 anticipates all this, he will either choose strategy  $(burn, not\ burn, a)$  or strategy  $(not\ burn, a)$ . As player 2, at  $h_5$ , need only believe that player 1 chooses optimally at  $h_5$  (but not at  $h_1$ ), player 2 will believe that player 1 does not choose  $c$  at  $h_5$ . Hence, player 2 must choose either the strategy  $(d, d, d)$  or the strategy  $(d, d, e)$ . The associated strategies and outcomes can be found in Table 2.

## 4 Existence

We will show that, under some conditions on the focus function  $f$ , there will always be at least one  $f$ -rationalizable strategy for every player.

**Definition 4.1 (Individual conditions on focus functions)** (a) A focus function  $f$  is **individually forward decreasing** if for every two players  $i \neq j$ , and every two information sets  $h, h' \in H$  where  $h'$  follows  $h$ , it holds that  $f_{ij}(h') \subseteq f_{ij}(h)$ .

(b) A focus function  $f$  **individually preserves focus on past information sets** if for every two players  $i \neq j$ , and every two information sets  $h, h' \in H$  where  $h'$  follows  $h$ , it holds that  $f_{ij}(h) \cap H^-(h) \subseteq f_{ij}(h')$ .

The individual forward decreasing property states that a player must not start reasoning about new information sets as the game proceeds, whereas individual preservation of

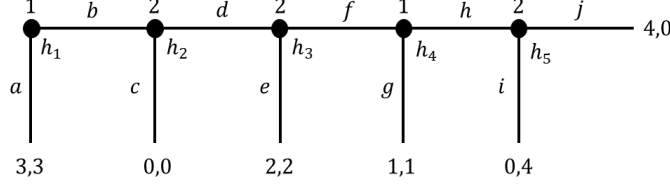


Figure 2: Why conditions in Theorem 4.1 are needed

focus on past information sets requires a player to always reason about all past information sets he focused on before.

**Theorem 4.1 (Existence)** *Let  $f$  be a focus function that is individually forward decreasing and individually preserving focus on past information sets. Then, for every player there is at least one  $f$ -rationalizable strategy.*

It may be verified that each of the three focus functions in Table 1 satisfies the conditions above. Therefore, Theorem 4.1 guarantees the existence of the three associated rationalizability concepts.

To see why the conditions in Theorem 4.1 are needed to guarantee existence, consider the game in Figure 2. Consider first the focus function  $f$  where player 1 always reasons about all information sets of player 2, player 2 reasons at  $h_1$  and  $h_2$  about all information sets of player 1, and player 2 reasons at  $h_3, h_4$  and  $h_5$  only about player 1's information set  $h_4$ . Then,  $f$  does not individually preserve focus on past information sets, as player 2 no longer reasons about the past information set  $h_1$  if the game moves from  $h_2$  to  $h_3$ . However,  $f$  is individually forward decreasing.

At  $h_2$ , player 2 must believe, whenever possible, that player 1 chose optimally at  $h_1$ . As  $(b, h)$  is the only strategy for player 1 that could give him more than 3, player 2 must believe at  $h_2$  that player 1 chooses strategy  $(b, h)$ . At  $h_3$ , however, player 2 no longer needs to believe that player 1 chose optimally at  $h_1$ , but he must believe that player 1 will choose optimally at  $h_4$ , and that player 1 believes at  $h_4$  that player 2 will choose optimally at  $h_5$ . But then, player 2 must believe at  $h_3$  that player 1 chooses  $(b, g)$ . The beliefs for player 2 at  $h_2$  and  $h_3$  are “dynamically inconsistent”, as there is no strategy for player 2 that is

optimal for both of these beliefs. As such, there is no  $f$ -rationalizable strategy for player 2.

Consider next the focus function  $g$  where player 1 always reasons about all information sets of player 2, player 2 reasons at  $h_1$  and  $h_2$  only about player 1's information set  $h_4$ , and player 2 reasons at  $h_3, h_4$  and  $h_5$  about all of player 1's information sets. Then,  $g$  is not individually forward decreasing, as player 2 reasons at  $h_3$  about more information sets than at  $h_2$ . However,  $g$  individually preserves focus on past information sets. By a similar line of reasoning as above, player 2 must believe at  $h_2$  that player 1 chooses  $(b, g)$ , and must believe at  $h_3$  that player 1 chooses  $(b, h)$ . Again, there is no strategy for player 2 that is optimal for both of these beliefs, and hence there is no  $g$ -rationalizable strategy for player 2.

## 5 Monotonicity Result

Consider two focus functions  $f$  and  $g$  where  $g \subseteq f$ . We will show that, under certain conditions, every  $f$ -rationalizable outcome is also  $g$ -rationalizable.

### 5.1 Theorem

The conditions we need are “interactive” extensions of the “individual” conditions in Definition 4.1.

**Definition 5.1 (Interactive conditions on focus functions)** (a) A focus function  $f$  is **forward decreasing** if for all players  $i, j, k$  with  $i \neq k$  and  $j \neq k$ , and every two information sets  $h, h' \in H$  where  $h'$  weakly follows  $h$ , it holds that  $f_{jk}(h') \subseteq f_{ik}(h)$ .

(b) A focus function  $f$  **preserves focus on past information sets** if for all players  $i, j, k$  with  $i \neq k$  and  $j \neq k$ , and every two information sets  $h, h' \in H$  where  $h'$  weakly follows  $h$ , it holds that  $f_{ik}(h) \cap H^-(h) \subseteq f_{jk}(h')$ .

Conditions (a) and (b) require that the focus functions of the different players should not only be individually forward decreasing and individually preserving focus on past information sets (as defined in Section 4) – which is the case when  $i = j$  – but should satisfy these requirements collectively. As such, the various focus functions of the different players should be “appropriately aligned”.

To state part (c) of the theorem below we need the following definition: We say that a focus function  $f$  is *future oriented* if  $f_{ij}(h) \subseteq H_j^+(h)$  for every two players  $i \neq j$  and every information set  $h \in H$ . That is, players only reason about information sets that weakly follow (but not necessarily all of these).

**Theorem 5.1 (Monotonicity theorem)** *Let  $f, g$  be two focus functions with  $g \subseteq f$  such that  $f$  is forward decreasing and preserving focus on past information sets, and  $g$  is individually forward decreasing and individually preserving focus on past information sets. Then,*

- (a) *the set of  $f$ -rationalizable outcomes is the same as the set of  $(g, f)$ -rationalizable outcomes,*
- (b) *every  $f$ -rationalizable outcome is  $g$ -rationalizable, and*
- (c) *if  $f$  is future oriented, then every  $f$ -rationalizable strategy is  $g$ -rationalizable.*

Part (a) thus reveals that, under the conditions of the theorem, it does not matter for the induced outcomes whether we only use the focus function  $f$ , or first iteratedly use  $g$  and then  $f$ , thereby giving epistemic priority to  $g$  over  $f$ . Part (b) immediately follows from (a) as, by construction, every  $(g, f)$ -rationalizable outcome is in particular  $g$ -rationalizable. Part (c) states that the monotonicity property in (b) even holds in terms of strategies if  $f$  only involves backward induction reasoning, and no forward induction reasoning.

As an illustration of this theorem, consider the two-step burning money game from Figure 1. Let  $f^{initial}, f^{future}, f^{focal}, f^{all}$  be the focus functions associated with the Ben-Porath procedure, backward dominance procedure, the focal forward induction procedure from Section 3.4, and the strong rationalizability procedure, respectively. Then,  $f^{initial} \subseteq f^{future} \subseteq f^{focal} \subseteq f^{all}$ , and all four focus functions satisfy the conditions of our theorem.

To illustrate part (a), recall from Table 2 that the only  $f^{all}$ -rationalizable outcome is *(not burn, a, d)*. By (a) in our theorem, it must thus be the case that *(not burn, a, d)* is the only  $(f^{focal}, f^{all})$ -rationalizable outcome, which is indeed true. However, this equality does not hold in terms of strategies: It turns out that *(not burn, a)* and *(d, d, d)* are the only  $(f^{focal}, f^{all})$ -rationalizable strategies, whereas also the strategies *(e, d, d)* and *(f, d, d)* are  $f^{all}$ -rationalizable.

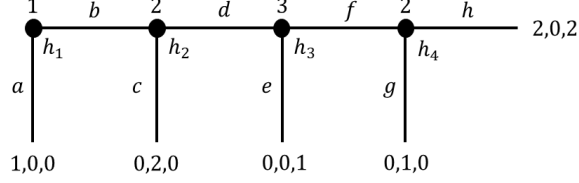


Figure 3: Why conditions in Theorem 5.1 are needed

For part (b), note from Table 2 that the set of  $f^{focal}$ -rationalizable outcomes is strictly nested between the set of  $f^{all}$ -rationalizable outcomes and the set of  $f^{future}$ -rationalizable outcomes, and that the set of  $f^{future}$ -rationalizable outcomes is strictly nested between the set of  $f^{focal}$ -rationalizable outcomes and the set of  $f^{initial}$ -rationalizable outcomes. Again, Table 2 shows that these inclusions do not hold in terms of strategies.

As to part (c), note that  $f^{initial} \subseteq f^{future}$ . As  $f^{future}$  is future oriented, part (c) thus guarantees that every  $f^{future}$ -rationalizable strategy is  $f^{initial}$ -rationalizable, which is indeed confirmed by Table 2.

To show that the conditions in Theorem 5.1 are needed for the result, consider the game in Figure 3. Let  $f$  be the focus function where (i) player 1, at all information sets, only reasons about player 3's information set  $h_3$ , (ii) player 2 never reasons about any information set, and (iii) player 3, at all information sets, only reason about player 2's information sets  $h_2$  and  $h_4$ . Note that  $f$  is not forward decreasing as player 1, at  $h_1$ , does not reason about player 2's information sets  $h_2$  and  $h_4$ , whereas player 3 does so at  $h_3$ .

Under  $f$ -rationalizability player 3 must believe at  $h_3$ , whenever possible, that player 2 chose optimally at  $h_2$  and will choose optimally at  $h_4$ . Since there is no strategy for player 2 that is optimal at  $h_2$  and  $h_4$  for some belief, player 3 may believe that player 2 will suboptimally choose  $h$  at  $h_4$ . As such, player 3 can optimally choose  $f$  at  $h_3$ . As player 1 need only believe at  $h_1$  that player 3 (but not player 2) will choose optimally, player 1 may believe at  $h_1$  that player 2 chooses  $(d, h)$  and player 3 chooses  $f$ . As such, player 1 can optimally choose  $b$  at  $h_1$ , and hence strategy  $b$  is  $f$ -rationalizable. Since strategy  $c$  is clearly  $f$ -rationalizable for player 2, the outcome  $(b, c)$  is  $f$ -rationalizable.

Next, consider the focus function  $g$  where player 3, at all information sets, only reasons about player 2's information set  $h_4$ , and that otherwise coincides with  $f$ . Then,  $g \subseteq f$ , and  $g$  is individually forward decreasing and individually preserving focus on past information sets. Under  $g$ -rationalizability, player 3 need not believe at  $h_3$  that player 2 chose optimally at  $h_2$ , but he must believe at  $h_3$  that player 2 will choose optimally at  $h_4$ . That is, player 3 must believe at  $h_3$  that player 2 will choose  $g$  at  $h_4$ , and thus player 3's only optimal strategy is  $e$ . Player 1, anticipating this, would only opt for strategy  $a$ , and hence  $a$  is the only  $g$ -rationalizable strategy for player 1. As such,  $a$  is the only  $g$ -rationalizable outcome.

This means, however, that the set of  $f$ -rationalizable outcomes is not contained in the set of  $g$ -rationalizable outcomes, although  $g \subseteq f$ . The reason is that  $f$  is not forward decreasing.

This example also highlights the main difficulty in proving Theorem 5.1: If the focus function  $f$  is compared to a “less fine-grained” focus function  $g$ , then there may be information sets (like  $h_3$  in the example) where the active player may be able to rationalize the opponents' moves when reasoning in accordance with  $g$ -rationalizability, but not when he reasons in accordance with  $f$ -rationalizability. In such situations,  $g$ -rationalizability may impose more restrictions on the beliefs at that information set compared to  $f$ -rationalizability, as was the case in the example above. The major challenge is to prove that, under appropriate conditions, this phenomenon does not cause the set of  $g$ -rationalizable outcomes to become smaller than the set of  $f$ -rationalizable outcomes.

## 5.2 Proof Sketch

The proof of part (a) in Theorem 5.1 is constructive: We explicitly show how to transform a combination  $(s_i)_{i \in I}$  of  $f$ -rationalizable strategies into a combination  $(\sigma_i(s_i))_{i \in I}$  of  $(g, f)$ -rationalizable strategies that induces the same outcome, and *vice versa*. Here,  $\sigma_i$  represents a transformation mapping for every player  $i$ , transforming an  $f$ -rationalizable strategy  $s_i$  into a  $(g, f)$ -rationalizable strategy  $\sigma_i(s_i)$  that preserves the behavior of  $s_i$  at those information sets that can be reached under  $(s_i)_{i \in I}$ .

To formally describe  $\sigma_i(s_i)$  we need some new terminology. We say that a strategy  $s_i$  is  $f$ -rationalizable ( $(g, f)$ -rationalizable) at information set  $h_i \in H_i(s_i)$  if  $s_i$  survives all the elimination steps of  $f$ -rationalizability ( $(g, f)$ -rationalizability) at  $h_i$ . An information set  $h$  is called *explicable* under  $f$ -rationalizability if it is possible for every player  $i$  to believe

at  $h$  that every opponent  $j$  chooses a strategy that leads to  $h$  and that is  $f$ -rationalizable at every information set for  $j$  that  $i$  reasons about at  $h$ .

Now, take an  $f$ -rationalizable strategy  $s_i$ . We then transform  $s_i$  into a new strategy  $\sigma_i(s_i)$  that (i) agrees with  $s_i$  at all explicable information sets under  $f$ -rationalizability, and (ii) agrees at the non-explicable information sets with an arbitrary strategy that is  $(g, f)$ -rationalizable there. Our proof shows that the new strategy  $\sigma_i(s_i)$  is  $(g, f)$ -rationalizable, and that every combination  $(s_i)_{i \in I}$  of  $f$ -rationalizable strategies is transformed into a combination  $(\sigma_i(s_i))_{i \in I}$  of  $(g, f)$ -rationalizable strategies that yields the same outcome as  $(s_i)_{i \in I}$ . Moreover, a similar construction can be used to transform every combination  $(s_i)_{i \in I}$  of  $(g, f)$ -rationalizable strategies into a combination  $(\tau_i(s_i))_{i \in I}$  of  $f$ -rationalizable strategies that yields the same outcome as  $(s_i)_{i \in I}$ .

As an illustration for how to construct the transformed strategies  $\sigma_i(s_i)$ , consider the two-step burning money game from Figure 1. Let  $f$  be the focus function that belongs to strong rationalizability, and  $g$  the focus function associated with the focal forward induction concept in Table 2. From this table we see that the only  $f$ -rationalizable strategy for player 1 is  $s_1 = (\text{not burn}, a)$ , which implies that the only explicable information sets under  $f$ -rationalizability are  $h_1$  and  $h_5$ . Moreover, we know from the same table that the strategy  $s_2 = (e, d, d)$  is  $f$ -rationalizable, but not  $g$ -rationalizable (and hence not  $(g, f)$ -rationalizable) for player 2.

We will now use the construction above to transform the pair  $(s_1, s_2)$  of  $f$ -rationalizable strategies into a pair  $(\sigma_1(s_1), \sigma_2(s_2))$  of  $(g, f)$ -rationalizable strategies yielding the same outcome. As  $s_1 = (\text{not burn}, a)$  only reaches explicable information sets under  $f$ -rationalizability, we set  $\sigma_1(s_1) := s_1$ .

Next, take  $s_2 = (e, d, d)$ . As  $h_5$  is an explicable information set under  $f$ -rationalizability, we let  $\sigma_2(s_2)$  agree with  $s_2$  at  $h_5$ . That is,  $\sigma_2(s_2)$  prescribes  $d$  at  $h_5$ . Since  $h_3$  and  $h_4$  are non-explicable under  $f$ -rationalizability, we let  $\sigma_2(s_2)$  agree at  $h_3$  and  $h_4$  with an arbitrary  $(g, f)$ -rationalizable strategy. Since we know by Table 2 that  $(d, d, d)$  and  $(d, d, e)$  are the only  $g$ -rationalizable strategies for player 2, it follows that every  $(g, f)$ -rationalizable strategy for player 2 must prescribe  $d$  at  $h_3$  and  $h_4$ . Hence, we let  $\sigma_2(s_2)$  prescribe  $d$  at  $h_3$  and  $h_4$ . Altogether, we thus set  $\sigma_2(s_2) := (d, d, d)$ .

It can be verified that the transformed strategies  $\sigma_1(s_1)$  and  $\sigma_2(s_2)$  are  $(g, f)$ -rationalizable, and that they induce the same outcome  $(\text{not burn}, a, d)$  as the original strategies  $s_1$  and

## 6 Implications

We now show how Theorem 5.1 can be used to prove some existing results in the literature, but also how it generates new interesting results.

### 6.1 Forward versus Backward Induction Reasoning

Perea (2017) and Catonini (2020) have shown, for dynamic games with observable actions, that every outcome induced by strong rationalizability is also induced by the backward dominance procedure. It also follows from the arguments in Chen and Micali (2013). This property is an immediate consequence of Theorem 5.1: It may be verified that  $f := f^{all}$  and  $g := f^{future}$  satisfy all the conditions in Theorem 5.1. As such, by Theorem 5.1 (b), every  $f^{all}$ -rationalizable outcome is  $f^{future}$ -rationalizable, leading to the following corollary.

**Corollary 6.1 (Forward versus backward induction reasoning)** *Every outcome induced by strong rationalizability is also induced by the backward dominance procedure.*

On the basis of Theorem 5.1 (a) we can say even more:

**Corollary 6.2 (Epistemic priority)** *The outcomes induced by strong rationalizability are the same as those induced by first applying the backward dominance procedure, followed by applying strong rationalizability.*

That is, for the induced outcomes it does not matter whether we give epistemic priority to the backward induction concept of backward dominance over the forward induction concept of strong rationalizability, or whether we restrict ourselves to strong rationalizability all the way. This result is similar, in spirit, to Catonini (2020) where it is shown that under certain conditions, first imposing exogenous restrictions on the first-order beliefs and then applying the strong rationalizability procedure yields the same outcomes as reversing this order.

Finally, let us apply Corollary 6.1 to a game with perfect information without relevant ties. As the backward dominance procedure leads to the unique backward induction

strategies (see Perea (2014)), and hence to the unique backward induction outcome, and strongly rationalizable strategies always exist, we obtain the following result which has first been shown by Battigalli (1997).

**Corollary 6.3 (Battigalli’s theorem)** *Consider a dynamic game with perfect information and without relevant ties. Then, strong rationalizability uniquely induces the backward induction outcome.*

## 6.2 Focal Forward Induction Reasoning

In principle, each of the corollaries above could also have been proven on the basis of Chen and Micali (2013), who show that the iterated conditional dominance procedure (Shimaji and Watson (1998)), which in turn is equivalent to the strong rationalizability procedure, is *order independent with respect to outcomes*. That is, changing the order of elimination may alter the sets of strategies, but not the set of induced outcomes. Indeed, the procedure in which we first apply the backward dominance procedure followed by the strong rationalizability procedure, can be viewed as an alternative order of elimination for the strong rationalizability procedure. By Chen and Micali (2013) both procedures should thus yield the same set of outcomes, establishing Corollary 6.2. The other two corollaries above follow from this result.

However, Theorem 5.1 also enables us to prove new interesting results that cannot be shown on the basis of Chen and Micali (2013). Consider a focus function  $f$  and the associated  $f$ -rationalizability procedure. Then, we say that this procedure is a *focal forward induction procedure* if there are collections of focal information sets  $H^{past}$  and  $H^{future}$  with  $H^{past} \subseteq H^{future}$ , such that

$$f_{ij}(h) = (H_j^-(h) \cap H^{past}) \cup (H_j^+(h) \cap H^{future})$$

for all information sets  $h \in H$  and players  $i \neq j$ . A special case is the focal forward induction concept from Section 3.4 for the two-step burning money game, where we take  $H^{past} = H \setminus \{h_1\}$  and  $H^{future} = H$ . In a focal forward induction procedure, a player always believes, whenever possible, that the opponents have chosen optimally at all focal past information sets in  $H^{past}$ , and will choose optimally at all focal future information sets in  $H^{future}$ . The interpretation is that the players collectively deem the information sets in  $H^{past}$  and  $H^{future}$  as particularly important, which may naturally happen if mistakes at

those information sets can have drastic consequences for the associated players compared to other information sets. We call  $f$ -rationalizability the *focal forward induction concept induced by  $H^{past}$  and  $H^{future}$* . Since  $H^{past} \subseteq H^{future}$ , it may be verified that  $f$  satisfies the conditions in Theorem 5.1. On the basis of Theorem 5.1 we can thus establish the following result.

**Corollary 6.4 (Focal forward induction)** *Let  $f$ -rationalizability and  $g$ -rationalizability be two focal forward induction concepts induced by  $H^{past,f}, H^{future,f}$  and  $H^{past,g}, H^{future,g}$ , respectively. If  $H^{past,g} \subseteq H^{past,f}$  and  $H^{future,g} \subseteq H^{future,f}$ , then every  $f$ -rationalizable outcome is also  $g$ -rationalizable.*

That is, if under  $f$ -rationalizability, compared to  $g$ -rationalizability, a player always attempts to rationalize more instances of past behavior, and draws more inferences about future behavior from observed past behavior, then the more refined focal forward induction concept induces a smaller (or possibly the same) set of outcomes compared to the less refined concept. If  $f$ -rationalizability does not coincide with strong rationalizability, then this result does not follow from Chen and Micali (2013).

## 7 Concluding Remarks

**Extension to context dependent focus functions.** In this paper we have concentrated on “context independent” focus functions  $f$  where at every information set, the collection of opponents’ information sets that a player reasons about is independent of the strategies that have been eliminated in the associated  $f$ -rationalizability procedure so far. In the working paper version Perea (2024) we consider a more general class of “context dependent” focus functions, where the collection of opponents’ information sets a player reasons about may depend on the strategies that have been eliminated so far. Theorem 4.1 has been extended to that setting, by additionally imposing *monotonicity* on the focus functions. Similarly, the working paper generalizes Theorem 5.1 to context-dependent focus functions, by adding the requirements of *monotonicity*, *monotonicity with respect to another focus function* and being *transitively closed*. It turns out that the last property can be dispensed with for the generalization of Theorem 5.1, which is the reason why it does not figure in the present paper.

These generalizations enable us to also capture concepts like *forward and backward rationalizability* in Meier and Perea (2023), where we first apply the strong rationalizability procedure to the last information sets in the game, followed by applying strong rationalizability to the last and second-to-last information sets, keeping the restrictions at the last information sets, and so on, until we reach the beginning of the game. It thus reveals an epistemic priority over different modes of reasoning, where we give the highest priority to forward induction reasoning about the last information sets, the second-highest priority to forward induction reasoning about the last and second-to-last information sets, and so on. Such epistemic priority based concepts can be modelled as an instance of  $f$ -rationalizability for an appropriately chosen context dependent focus function  $f$ .

**Extension to games with non-observable actions.** In the working paper version Perea (2024), it is shown that the Theorems 4.1 and 5.1, and therefore all the derived results from Section 6, also hold for games with non-observable actions.<sup>8</sup> That is, some players may not fully observe which actions have been chosen by the opponents in the past. In this paper we restrict to observable actions as to make the exposition and the analysis easier.

**Incomplete information.** A route that remains to be explored is how to extend our analysis to games with incomplete information. Many of the rationalizability concepts, such as strong  $\Delta$ -rationalizability (Battigalli (2003)), which includes strong rationalizability as a special case, and backwards rationalizability (Penta (2015), Catonini and Penta (2026), Perea (2014)), have already been defined for games with incomplete information. Extending the monotonicity result to the context of incomplete information may be very valuable for implementation theory and mechanism design. In that light, Battigalli and Catonini (2026) prove an outcome-monotonicity result that is similar, in spirit, to ours: It is shown that when the planner evaluates the agents' behavior by strong rationalizability, then imposing more restrictions on the players' initial beliefs about types leads to fewer, or the same, outcomes.

---

<sup>8</sup>For Theorem 5.1, the working paper version also requires the focus function to be *transitively closed*. However, as stated above, this property is not needed.

## 8 Proofs

### 8.1 Preparatory Results for Theorem 4.1

For a decision problem map  $D$ , player  $i$  and information set  $h_i \in H_i$  we define the set

$$D_i^*(h_i) := \{s_i \in S_i(h_i) \mid s_i \text{ not strictly dominated in } (S_i(h_i), D_{-i}(h_i))\}.$$

**Lemma 8.1 (Optimal strategies in rationalizability procedure)** *Let  $f, g$  be two focus functions, and let  $D = (rf^k \circ rg^m)(D^{full})$  for some  $k, m$  with  $k + m \geq 1$ . Then,  $D_i(h_i) = D_i^*(h_i)$ .*

**Proof.** We prove the statement by induction on  $k + m$ .

**Induction start.** Let  $k + m = 1$ . Suppose first that  $k = 0$  and  $m = 1$ , which means that  $D = rg(D^{full})$ . But then, by definition,  $D_i(h_i) = D_i^*(h_i)$ . If  $k = 1$  and  $m = 0$ , then  $D = rf(D^{full})$ , which similarly implies that  $D_i(h_i) = D_i^*(h_i)$ .

**Induction step.** Suppose next that  $k, m$  are such  $k + m \geq 2$ , and that the statement is true for  $k + m - 1$ .

**Case 1.** Assume first that  $k \geq 1$ . Then,  $D = rf(E)$ , where  $E := (rf^{k-1} \circ rg^m)(D^{full})$ . We thus know, by induction, that  $E_i(h_i) = E_i^*(h_i)$ . We start by showing that  $D_i(h_i) \subseteq D_i^*(h_i)$ . Take some  $s_i \in D_i(h_i)$ . Then, by definition,  $s_i \in E_i(h_i)$  and  $s_i$  is not strictly dominated in  $(E_i(h_i), D_{-i}(h_i))$ . By Lemma 2.1 we then know that  $s_i$  is optimal in  $E_i(h_i)$  for a belief  $b_i \in \Delta(D_{-i}(h_i))$ . We will show that  $s_i$  is optimal in  $S_i(h_i)$  for  $b_i$ . Suppose not. Then, there is some  $s_i^* \in S_i(h_i)$  such that

$$u_i(s_i, b_i) < u_i(s_i^*, b_i) \tag{8.1}$$

and  $s_i^*$  is optimal in  $S_i(h_i)$  for  $b_i$ . As  $D_{-i}(h_i) \subseteq E_{-i}(h_i)$  it follows that  $b_i \in \Delta(E_{-i}(h_i))$ . Hence, we conclude by Lemma 2.1 that  $s_i^*$  is not strictly dominated in  $(S_i(h_i), E_{-i}(h_i))$ , and thus  $s_i^* \in E_i^*(h_i)$ . Since, by the induction assumption,  $E_i^*(h_i) = E_i(h_i)$  it follows that  $s_i^* \in E_i(h_i)$ . However, together with (8.1) this would contradict the fact that  $s_i$  is optimal in  $E_i(h_i)$  for  $b_i$ . We thus conclude that  $s_i$  is optimal in  $S_i(h_i)$  for  $b_i$ . Since  $b_i \in \Delta(D_{-i}(h_i))$  we know by Lemma 2.1 that  $s_i$  is not strictly dominated in  $(S_i(h_i), D_{-i}(h_i))$ , and hence  $s_i \in D_i^*(h_i)$ . As this holds for every  $s_i \in D_i(h_i)$  we see that  $D_i(h_i) \subseteq D_i^*(h_i)$ .

We next show that  $D_i^*(h_i) \subseteq D_i(h_i)$ . Take some  $s_i \in D_i^*(h_i)$ . Then, by definition,  $s_i$  is not strictly dominated in  $(S_i(h_i), D_{-i}(h_i))$ . As  $D_{-i}(h_i) \subseteq E_{-i}(h_i)$  it follows that  $s_i$  is not

strictly dominated in  $(S_i(h_i), E_{-i}(h_i))$  and hence  $s_i \in E_i^*(h_i)$ . By the induction assumption we know that  $E_i^*(h_i) = E_i(h_i)$  and hence  $s_i \in E_i(h_i)$ . We thus see that  $s_i \in E_i(h_i)$  and  $s_i$  is not strictly dominated in  $(S_i(h_i), D_{-i}(h_i))$ . In particular,  $s_i$  is not strictly dominated in  $(E_i(h_i), D_{-i}(h_i))$ , which implies that  $s_i \in D_i(h_i)$ . As this holds for every  $s_i \in D_i^*(h_i)$  we conclude that  $D_i^*(h_i) \subseteq D_i(h_i)$ . Together with the insight above that  $D_i(h_i) \subseteq D_i^*(h_i)$  we conclude that  $D_i(h_i) = D_i^*(h_i)$ .

**Case 2.** Suppose that  $k = 0$ . Then,  $m \geq 2$ , and we have that  $D = rg(E)$  where  $E := rg^{m-1}(D^{full})$ . By the induction assumption we know that  $E_i(h_i) = E_i^*(h_i)$ . In the same way as in Case 1 it can then be shown that  $D_i(h_i) = D_i^*(h_i)$ . By induction on  $k + m$ , the proof is hereby complete.  $\blacksquare$

Recall the definition of  $D_{-i}^+(h_i, f)$  from Section 3.3. We define

$$D_{-i}^{+*}(h_i, f) := \{(s_j)_{j \neq i} \in S_{-i}(h_i) \mid \text{for all } j \neq i, s_j \in D_j(h_j) \text{ for all } h_j \in f_{ij}(h_i) \cap H_j(s_j)\}.$$

$$\text{Hence, } D_{-i}^+(h_i, f) = D_{-i}^{+*}(h_i, f) \cap D_{-i}(h_i).$$

**Lemma 8.2 (Opponents' strategies in rationalizability procedure)** *Consider two focus functions  $f, g$  with  $g \subseteq f$  and let  $D = (rf^k \circ rg^m)(D^{full})$  for some  $k, m \geq 0$ . Then, for every player  $i$  and information set  $h_i \in H_i$  we have that  $D_{-i}^{+*}(h_i, f) = D_{-i}^+(h_i, f)$ , and  $D_{-i}^{+*}(h_i, g) = D_{-i}^+(h_i, g)$  whenever  $k = 0$ .*

**Proof.** We show the statement by induction on  $k + m$ .

**Induction start.** For  $k + m = 0$  the statement is true since in this case  $D_{-i}(h_i) = S_{-i}(h_i)$ .

**Induction step.** Assume next that  $k + m \geq 1$ , and that the statement is true for  $k + m - 1$ .

**Case 1.** Suppose that  $k \geq 1$ . Note that  $D = rf(E)$ , where  $E := (rf^{k-1} \circ rg^m)(D^{full})$ , and we know by the induction assumption that  $E_{-i}^{+*}(h_i, f) = E_{-i}^+(h_i, f)$ . We will show that  $D_{-i}^{+*}(h_i, f) = D_{-i}^+(h_i, f)$ . Since, by construction,  $D_{-i}^+(h_i, f) \subseteq D_{-i}^{+*}(h_i, f)$  it only remains to show that  $D_{-i}^{+*}(h_i, f) \subseteq D_{-i}^+(h_i, f)$ . To show this, take some  $s_{-i} = (s_j)_{j \neq i} \in D_{-i}^{+*}(h_i, f)$ . Then,  $s_{-i} \in S_{-i}(h_i)$  and for all  $j \neq i$  we have that  $s_j \in D_j(h_j)$  for all  $h_j \in f_{ij}(h_i) \cap H_j(s_j)$ . As  $D \subseteq E$ , we conclude that  $s_j \in E_j(h_j)$  for all  $h_j \in f_{ij}(h_i) \cap H_j(s_j)$ . This implies that  $s_{-i} \in E_{-i}^{+*}(h_i, f)$ . By the induction assumption we then know that  $s_{-i} \in E_{-i}^+(h_i, f)$ . As  $D = rf(E)$ , we conclude that  $s_{-i} \in D_{-i}(h_i)$ . Since, by assumption,  $s_{-i} \in D_{-i}^{+*}(h_i, f)$ , it follows that  $s_{-i} \in D_{-i}^+(h_i, f)$ . Thus,  $D_{-i}^{+*}(h_i, f) \subseteq D_{-i}^+(h_i, f)$ , and hence  $D_{-i}^{+*}(h_i, f) = D_{-i}^+(h_i, f)$ , which was to show.

**Case 2.** Suppose that  $k = 0$ . Then,  $m \geq 1$  and  $D = rg(E)$  where  $E := rg^{m-1}(D^{full})$ . By the induction assumption we thus know that  $E_{-i}^{+*}(h_i, g) = E_{-i}^+(h_i, g)$ .

We first show that  $D_{-i}^{+*}(h_i, f) = D_{-i}^+(h_i, f)$ . Again, it only remains to show that  $D_{-i}^{+*}(h_i, f) \subseteq D_{-i}^+(h_i, f)$ . Take some  $s_{-i} = (s_j)_{j \neq i} \in D_{-i}^{+*}(h_i, f)$ . Then,  $s_{-i} \in S_{-i}(h_i)$  and for all  $j \neq i$  we have that  $s_j \in D_j(h_j)$  for all  $h_j \in f_{ij}(h_i) \cap H_j(s_j)$ . As  $D \subseteq E$  and  $g \subseteq f$  we conclude that  $s_j \in E_j(h_j)$  for all  $h_j \in g_{ij}(h_i) \cap H_j(s_j)$ , which implies that  $s_{-i} \in E_{-i}^{+*}(h_i, g)$ . By the induction assumption we then know that  $s_{-i} \in E_{-i}^+(h_i, g)$ . As  $D = rg(E)$ , we then know that  $s_{-i} \in D_{-i}(h_i)$ . Since, by assumption,  $s_{-i} \in D_{-i}^{+*}(h_i, f)$ , it follows that  $s_{-i} \in D_{-i}^+(h_i, f)$ . Thus,  $D_{-i}^{+*}(h_i, f) \subseteq D_{-i}^+(h_i, f)$ , and hence  $D_{-i}^{+*}(h_i, f) = D_{-i}^+(h_i, f)$ , which was to show.

We next show that  $D_{-i}^{+*}(h_i, g) = D_{-i}^+(h_i, g)$ . This can be done in the same way as for Case 1, interchanging the roles of  $f$  and  $g$ . The proof then follows by induction on  $k + m$ .  $\blacksquare$

Consider a belief  $b_i \in \Delta(S_{-i}(h_i))$ , and some  $h'_i \in H_i$  following  $h_i$  with  $b_i(S_{-i}(h'_i)) > 0$ . Then, the belief  $b'_i \in \Delta(S_{-i}(h'_i))$  given by  $b'_i(s_{-i}) := b_i(s_{-i})/b_i(S_{-i}(h'_i))$  for all  $s_{-i} \in S_{-i}(h'_i)$  is called the *forward consistent update* of  $b_i$  at  $h'_i$ .

**Lemma 8.3 (Preservation of optimality)** Consider a belief  $b_i \in \Delta(S_{-i}(h_i))$ , a strategy  $s_i$  that is optimal in  $S_i(h_i)$  for  $b_i$ , and an information set  $h'_i \in H_i(s_i)$  following  $h_i$  with  $b_i(S_{-i}(h'_i)) > 0$ . Let  $b'_i$  be the forward consistent update of  $b_i$  at  $h'_i$ . Then,  $s_i$  is optimal in  $S_i(h'_i)$  for  $b'_i$ .

The proof can be found in Perea (2012), proof of Lemma 8.14.9.

We say that a decision problem map  $D$  is *dynamically consistent* if for every player  $i$  and every two information sets  $h_i, h'_i \in H_i$  where  $h_i$  precedes  $h'_i$ , it holds that  $D_{-i}(h_i) \cap S_{-i}(h'_i) \subseteq D_{-i}(h'_i)$ . For a given strategy  $s_i \in S_i$  and belief  $b_i \in \Delta(S_{-i})$  we denote by

$$H(s_i, b_i) := \{h \in H \mid s_i \in S_i(h) \text{ and } b_i(S_{-i}(h)) > 0\}$$

the collection of information sets that  $i$  believes to be reached with positive probability under  $(s_i, b_i)$ . For a given strategy  $s_i \in S_i$  and a collection of information sets  $\hat{H} \subseteq H$ , let  $s_i|_{\hat{H}} := (s_i(h_i))_{h_i \in \hat{H} \cap H_i}$  be its restriction to information sets in  $\hat{H}$ .

**Lemma 8.4 (Perfectionating a strategy)** Consider a dynamically consistent decision problem map  $D$ , a strategy  $s_i \in S_i$ , and let  $H_i^* = \{h_i \in H_i(s_i) \mid s_i \in D_i^*(h_i)\}$ . Then, for every  $h^* \in H$  with  $s_i \in S_i(h^*)$  the following holds:

- (a) There is a strategy  $s_i^* \in S_i(h^*)$  such that  $s_i^* \in D_i^*(h_i)$  for all  $h_i \in (H_i^* \cap H_i^-(h^*)) \cup (H_i(s_i^*) \setminus H_i^-(h^*))$ .
- (b) If  $h^* \in H_i^*$ , then there is a belief  $b_i \in \Delta(D_{-i}(h^*))$  for which the strategy  $s_i$  is optimal in  $S_i(h^*)$ , and for which the strategy  $s_i^*$  in (a) can be constructed such that  $s_i^*|_{H(s_i, b_i)} = s_i|_{H(s_i, b_i)}$ .
- (c) If  $h^* \in H_i \setminus H_i^*$ , then for every belief  $b_i \in \Delta(D_{-i}(h^*))$  the strategy  $s_i^*$  in (a) can be constructed such that  $s_i^*$  is optimal in  $S_i(h^*)$  for  $b_i$ .
- (d) There is a strategy  $s_i^* \in S_i(h^*)$  such that  $s_i^*|_{H(h^*)} = s_i|_{H(h^*)}$  and  $s_i^* \in D_i^*(h_i)$  for all  $h_i \in (H_i^* \cap H(h^*)) \cup (H_i(s_i^*) \setminus H(h^*))$ .

**Proof.** Consider an information set  $h^* \in H$  with  $s_i \in S_i(h^*)$ .

(a) Let  $h_i^1$  be the first information set, if any, in  $H_i^* \cap H_i^-(h^*)$ . By assumption,  $s_i \in D_i^*(h_i^1)$ . Hence, by definition,  $s_i$  is not strictly dominated in  $(S_i(h_i^1), D_{-i}(h_i^1))$ . By Lemma 2.1 there is a belief  $b_i(h_i^1) \in \Delta(D_{-i}(h_i^1))$  such that  $s_i$  is optimal in  $S_i(h_i^1)$  for  $b_i(h_i^1)$ . Let

$$H_i^1 := \{h_i \in H_i^+(h_i^1) \mid b_i(h_i^1)(S_{-i}(h_i)) > 0\}.$$

For every  $h_i \in H_i^1 \setminus \{h_i^1\}$  let  $b_i(h_i)$  be the forward consistent update of  $b_i(h_i^1)$ . Then, we know from Lemma 8.3 that for every  $h_i \in H_i^1 \cap H_i(s_i)$  the strategy  $s_i$  is optimal in  $S_i(h_i)$  for  $b_i(h_i)$ . Moreover, as  $b_i(h_i^1) \in \Delta(D_{-i}(h_i^1))$  and  $D$  is dynamically consistent, it can be verified that  $b_i(h_i) \in \Delta(D_{-i}(h_i))$  for every  $h_i \in H_i^1 \setminus \{h_i^1\}$ .

Let  $h_i^2$  be the first information set, if any, in  $(H_i^* \cap H_i^-(h^*)) \setminus H_i^1$ . By assumption,  $s_i \in D_i^*(h_i^2)$ . Hence, by a similar argument as above, there is a belief  $b_i(h_i^2) \in \Delta(D_{-i}(h_i^2))$  such that  $s_i$  is optimal in  $S_i(h_i^2)$  for  $b_i(h_i^2)$ . Set

$$H_i^2 := \{h_i \in H_i^+(h_i^2) \mid b_i(h_i^2)(S_{-i}(h_i)) > 0\}.$$

For every  $h_i \in H_i^2 \setminus \{h_i^2\}$  let  $b_i(h_i)$  be the forward consistent update of  $b_i(h_i^2)$  at  $h_i$ . Then, similarly as above, we conclude that  $s_i$  is optimal in  $S_i(h_i)$  for  $b_i(h_i)$  and that  $b_i(h_i) \in \Delta(D_{-i}(h_i))$ .

We would then select  $h_i^3$  as the first information set, if any, in  $(H_i^* \cap H_i^-(h^*)) \setminus (H_i^1 \cup H_i^2)$ , and so on. We continue this construction until we arrive at a collection  $H_i^K$  where  $H_i^* \cap H_i^-(h^*) \subseteq H_i^1 \cup \dots \cup H_i^K$ . We then set

$$s_i^*(h_i) := s_i(h_i) \text{ for all } h_i \in (H_i^1 \cup \dots \cup H_i^K) \cap H_i(s_i). \quad (8.2)$$

Moreover, we define

$$s_i^*(h_i) := s_i(h_i) \text{ for all } h_i \in H_i^-(h^*) \setminus (H_i^1 \cup \dots \cup H_i^K). \quad (8.3)$$

We then set

$$\hat{H}_i^1 := (H_i \setminus (H_i^-(h^*) \cup H_i^1 \cup \dots \cup H_i^K))^{first}.$$

For every  $h_i^1 \in \hat{H}_i^1$  we select a belief  $b_i(h_i^1) \in \Delta(D_{-i}(h_i^1))$ , and let  $s_i[h_i^1] \in S_i(h_i^1)$  be a strategy that is optimal in  $S_i(h_i^1)$  for  $b_i(h_i^1)$ . We define

$$\hat{H}_i^{1+} := \{h_i \in H_i \mid h_i \text{ weakly follows some } h_i^1 \in \hat{H}_i^1 \text{ with } b_i(h_i^1)(S_{-i}(h_i)) > 0\}.$$

Take some  $h_i \in \hat{H}_i^{1+} \setminus \hat{H}_i^1$ , and let  $h_i^1 \in \hat{H}_i^1$  be the unique information set in  $\hat{H}_i^1$  that precedes  $h_i$ . Define  $b_i(h_i)$  to be the forward consistent update of  $b_i(h_i^1)$  at  $h_i$ . Then, we know from Lemma 8.3 that strategy  $s_i[h_i^1]$  is optimal in  $S_i(h_i)$  for  $b_i(h_i)$  whenever  $h_i \in H_i(s_i[h_i^1])$ . Moreover, similarly as above, we know that  $b_i(h_i) \in \Delta(D_{-i}(h_i))$ .

We then set

$$\hat{H}_i^2 := (H_i \setminus (\hat{H}_i^{1+} \cup H_i^-(h^*) \cup H_i^1 \cup \dots \cup H_i^K))^{first},$$

and so on. We continue this construction until we reach a set  $\hat{H}_i^M$  such that

$$H_i = \hat{H}_i^{1+} \cup \dots \cup \hat{H}_i^{M+} \cup H_i^-(h^*) \cup H_i^1 \cup \dots \cup H_i^K.$$

On  $\hat{H}_i^{1+} \cup \dots \cup \hat{H}_i^{M+}$  define the strategy  $s_i^*$  such that for every  $m \in \{1, \dots, M\}$ , every  $h_i^m \in \hat{H}_i^m$ , and every  $h_i \in \hat{H}_i^{m+}$  weakly following  $h_i^m$ ,

$$s_i^*(h_i) := (s_i[h_i^m])(h_i). \quad (8.4)$$

This completes the construction of the strategy  $s_i^*$ .

By (8.2) and (8.3) it follows that  $s_i^*|_{H_i^-(h^*)} = s_i|_{H_i^-(h^*)}$ . As  $s_i \in S_i(h^*)$  we conclude that  $s_i^* \in S_i(h^*)$  as well. We will now show that  $s_i^* \in D_i^*(h_i)$  for all  $h_i \in (H_i^* \cap H_i^-(h^*)) \cup (H_i(s_i^*) \setminus H_i^-(h^*))$ . We separate two cases.

Suppose first that  $h_i \in H_i^k \cap H_i(s_i^*)$  for some  $k \in \{1, \dots, K\}$ . Then, by (8.2),  $s_i^*$  coincides with  $s_i$  on  $H_i^k \cap H_i(s_i^*)$ . Recall from above that strategy  $s_i$  is optimal in  $S_i(h_i)$  for  $b_i(h_i)$ . As, by construction of  $H_i^k$ , the expected utility  $u_i(s_i^*, b_i(h_i))$  only depends on the behavior of  $s_i^*$  at information sets in  $H_i^k$ , it follows that  $s_i^*$  is optimal in  $S_i(h_i)$  for  $b_i(h_i)$  as well. Since we have seen above that  $b_i(h_i) \in \Delta(D_{-i}(h_i))$ , we conclude by Lemma 2.1 that  $s_i^*$  is not strictly dominated in  $(S_i(h_i), D_{-i}(h_i))$ . That is,  $s_i^* \in D_i^*(h_i)$ , which was to show.

Suppose next that  $h_i \in \hat{H}_i^{m+} \cap H_i(s_i^*)$  for some  $m \in \{1, \dots, M\}$ . In a similar way as above, it can be shown that  $s_i^* \in D_i^*(h_i)$ .

(b) Suppose now that  $h^* \in H_i^*$ . Recall from (a) that  $K$  is such that  $H_i^* \cap H_i^-(h^*) \subseteq H_i^1 \cup \dots \cup H_i^K$ . We distinguish two cases: (1)  $h^* \in H_i^K$  and (2)  $h^* \notin H_i^K$ .

**Case 1.** Suppose that  $h^* \in H_i^K$ . Let  $h_i^K$  be the unique information set in  $(H_i^-(h^*) \cap H_i^K)^{first}$ , let  $b_i(h_i^K) \in \Delta(D_{-i}(h_i^K))$  be the associated belief at  $h_i^K$ , and let  $b_i(h^*)$  be the forward consistent update of  $b_i(h_i^K)$  at  $h^*$ . By definition of the forward consistent update we know that  $H(s_i, b_i(h^*)) \subseteq H(s_i, b_i(h_i^K))$ . As, by the construction in (a),  $s_i^*|_{H(s_i, b_i(h_i^K))} = s_i|_{H(s_i, b_i(h_i^K))}$ , it follows that  $s_i^*|_{H(s_i, b_i(h^*))} = s_i|_{H(s_i, b_i(h^*))}$ . Moreover, we have shown in (a) that  $b_i(h^*) \in \Delta(D_{-i}(h^*))$  and that  $s_i$  is optimal for  $b_i(h^*)$  in  $S_i(h^*)$ . Hence we can set  $b_i := b_i(h^*)$ .

**Case 2.** Suppose that  $h^* \notin H_i^K$ . Then, by construction in (a),  $h^* \in \hat{H}_i^1$ . As  $h^* \in H_i^*$  we know that  $s_i \in D_i^*(h^*)$ , and hence  $s_i$  is not strictly dominated in  $(S_i(h^*), D_{-i}(h^*))$ . By Lemma 2.1, there is thus a belief  $b_i \in \Delta(D_{-i}(h^*))$  for which  $s_i$  is optimal in  $S_i(h^*)$ . By the construction in (a) using the belief  $b_i$  at  $h^*$ , we know that  $s_i^*|_{H(s_i, b_i)} = s_i|_{H(s_i, b_i)}$ .

(c) Suppose next that  $h^* \in H_i \setminus H_i^*$ . Recall from (a) that  $K$  is such that  $H_i^* \cap H_i^-(h^*) \subseteq H_i^1 \cup \dots \cup H_i^K$ . Let  $h_i^K$  be the unique information set in  $(H_i^-(h^*) \cap H_i^K)^{first}$ , and let  $b_i(h_i^K) \in \Delta(D_{-i}(h_i^K))$  be the associated belief constructed in (a) for which  $s_i$  is optimal in  $S_i(h_i^K)$ . We show that  $h^* \notin H_i^K$ .

Suppose, contrary to what we want to show, that  $h^* \in H_i^K$ . Then, by definition,  $b_i(h_i^K)(S_{-i}(h^*)) > 0$ . Let  $b_i(h^*)$  be the forward consistent update of  $b_i(h_i^K)$  at  $h^*$ . By a similar argument as in (a) we conclude that  $b_i(h^*) \in \Delta(D_{-i}(h^*))$ . As  $s_i$  is optimal for  $b_i(h_i^K)$  in  $S_i(h_i^K)$ , it follows by Lemma 8.3 that  $s_i$  is optimal for  $b_i(h^*)$  in  $S_i(h^*)$ . Since  $b_i(h^*) \in \Delta(D_{-i}(h^*))$  it follows by Lemma 2.1 that  $s_i$  is not strictly dominated in  $(S_i(h^*), D_{-i}(h^*))$ , and hence  $s_i \in D_i^*(h^*)$ . That is,  $h^* \in H_i^*$ . This, however, is a contradiction as we assume that  $h^* \notin H_i^*$ . Hence, we conclude that  $h^* \notin H_i^K$ .

This means, however, that  $h^*$  will be included in  $\hat{H}_i^1$ . For every  $b_i \in \Delta(D_{-i}(h^*))$  we can thus choose the belief  $b_i(h^*) := b_i$  in the construction in (a). Since we have shown in (a) that  $s_i^*$  is optimal for  $b_i(h^*)$  in  $S_i(h^*)$ , it follows that  $s_i^*$  is optimal for  $b_i$  in  $S_i(h^*)$ .  
**(d)** This proof is very similar to the proof for (a), and is therefore left to the reader. ■

**Lemma 8.5 (Optimal planning)** *Consider a dynamically consistent decision problem map  $D$ . Then, for every player  $i$ ,*

- (a) *there is a strategy  $s_i$  such that  $s_i \in D_i^*(h_i)$  for every  $h_i \in H_i(s_i)$ , and*
- (b) *for every information set  $h_i \in H_i$  there is a strategy  $s_i \in S_i(h_i)$  such that  $s_i \in D_i^*(h'_i)$  for every  $h'_i \in H_i(s_i) \setminus H_i^-(h_i)$ .*

**Proof.** **(a)** By Lemma 8.4 (a) applied to the information set  $h^* = \emptyset$  marking the beginning of the game, there is a strategy  $s_i^*$  such that  $s_i^* \in D_i^*(h_i)$  for all  $h_i \in H_i(s_i^*)$ .  
**(b)** Consider an information set  $h_i \in H_i$  and a strategy  $s_i \in S_i(h_i)$ . Then, by Lemma 8.4 (a) applied to the information set  $h^* = h_i$  there is some strategy  $s_i^* \in S_i(h_i)$  such that  $s_i^* \in D_i^*(h'_i)$  for all  $h'_i \in H_i(s_i^*) \setminus H_i^-(h_i)$ . ■

**Lemma 8.6 (Sufficient conditions for dynamic consistency)** *Consider two focus functions  $f, g$  with  $g \subseteq f$  that are individually forward decreasing and individually preserving focus on past information sets, and let  $D = (rf^k \circ rg^m)(D^{full})$  for some  $k, m \geq 0$ . Then,  $D$  is dynamically consistent.*

**Proof.** We show the statement by induction on  $k + m$ .

**Induction start.** For  $k = 0$  and  $m = 0$  the statement is trivially true.

**Induction step.** Now take some  $k, m$  with  $k + m \geq 1$  and assume that the statement is true for  $k + m - 1$ .

**Case 1.** Suppose that  $k \geq 1$ . Then,  $D = rf(E)$  where  $E := (rf^{k-1} \circ rg^m)(D^{full})$  is dynamically consistent by the induction assumption. Take some information sets  $h_i, h'_i \in H_i$  where  $h_i$  precedes  $h'_i$ . We will show that  $D_{-i}(h_i) \cap S_{-i}(h'_i) \subseteq D_{-i}(h'_i)$ , distinguishing two cases.

**Case 1.1.** Suppose that  $E_{-i}^+(h_i, f) \neq \emptyset$ . Take some  $(s_j)_{j \neq i} \in D_{-i}(h_i) \cap S_{-i}(h'_i)$ . Then, in particular,  $(s_j)_{j \neq i} \in E_{-i}(h_i) \cap S_{-i}(h'_i)$ . Since, by the induction assumption,  $E$  is dynamically consistent, we conclude that  $(s_j)_{j \neq i} \in E_{-i}(h'_i)$ . As  $E_{-i}^+(h_i, f) \neq \emptyset$  we have that

$D_{-i}(h_i) = E_{-i}^+(h_i, f)$ , and hence  $(s_j)_{j \neq i} \in E_{-i}^+(h_i, f) \cap S_{-i}(h'_i)$ . Thus,  $s_j \in E_j(h_j)$  for all  $h_j \in f_{ij}(h_i) \cap H_j(s_j)$  for all  $j \neq i$ . As  $f$  is individually forward decreasing we know that  $f_{ij}(h'_i) \subseteq f_{ij}(h_i)$ , and hence  $s_j \in E_j(h_j)$  for all  $h_j \in f_{ij}(h'_i) \cap H_j(s_j)$  for all  $j \neq i$ . Since we have seen above that  $(s_j)_{j \neq i} \in E_{-i}(h'_i)$ , and  $D = rf(E)$ , it follows that  $(s_j)_{j \neq i} \in D_{-i}(h'_i)$ , which was to show.

**Case 1.2.** Suppose next that  $E_{-i}^+(h_i, f) = \emptyset$ . We show that  $E_{-i}^+(h'_i, f) = \emptyset$  as well. Suppose not. Then, there is some  $(s_j)_{j \neq i} \in E_{-i}^+(h'_i, f)$ . Hence,  $s_j \in E_j(h_j)$  for all  $h_j \in f_{ij}(h'_i) \cap H_j(s_j)$  for all  $j \neq i$ . Fix a player  $j \neq i$ . As  $f$  individually preserves focus on past information sets,  $f_{ij}(h_i) \cap H_j^-(h_i) \subseteq f_{ij}(h'_i)$ . Therefore,  $s_j \in E_j(h_j)$  for all  $h_j \in f_{ij}(h_i) \cap H_j^-(h_i)$ . By Lemma 8.1 we know that  $E_j(h_j) = E_j^*(h_j)$ , and hence it follows that  $s_j \in E_j^*(h_j)$  for all  $h_j \in f_{ij}(h_i) \cap H_j^-(h_i)$ . Moreover, by the induction assumption, we know that  $E$  is dynamically consistent. Using Lemma 8.4 (a) for  $h^* = h_i$ , we then know that there is some  $s_j^* \in S_j(h_i)$  such that  $s_j^* \in E_j^*(h_j)$  for all  $h_j \in f_{ij}(h_i) \cap H_j(s_j^*)$ . Since by Lemma 8.1 we know that  $E_j(h_j) = E_j^*(h_j)$ , it follows that

$$s_j^* \in E_j(h_j) \text{ for all } h_j \in f_{ij}(h_i) \cap H_j(s_j^*). \quad (8.5)$$

As  $s_j^* \in S_j(h_i)$  for all  $j \neq i$  we know that  $(s_j^*)_{j \neq i} \in S_{-i}(h_i)$ . Together with (8.5) we conclude that  $(s_j^*)_{j \neq i} \in E_{-i}^+(h_i, f)$ . By Lemma 8.2 we have that  $E_{-i}^{+*}(h_i, f) = E_{-i}^+(h_i, f)$ , and thus  $(s_j^*)_{j \neq i} \in E_{-i}^+(h_i, f)$ . This, however, is a contradiction, since  $E_{-i}^+(h_i, f) = \emptyset$ .

We must therefore conclude that  $E_{-i}^+(h'_i, f) = \emptyset$  as well. As  $D = rf(E)$  we must have, by definition, that  $D_{-i}(h_i) = E_{-i}(h_i)$  and  $D_{-i}(h'_i) = E_{-i}(h'_i)$ . Hence,

$$D_{-i}(h_i) \cap S_{-i}(h'_i) = E_{-i}(h_i) \cap S_{-i}(h'_i) \subseteq E_{-i}(h'_i) = D_{-i}(h'_i),$$

where the set inclusion follows from the induction assumption that  $E$  is dynamically consistent. By combining Case 1.1 and Case 1.2 we conclude that  $D$  is dynamically consistent.

**Case 2.** Suppose that  $k = 0$ . Then,  $m \geq 1$  and  $D = rg(E)$  where  $E := rg^{m-1}(D^{full})$  is dynamically consistent by the induction assumption. In the same way as in Case 1, interchanging the roles of  $f$  and  $g$ , it can be shown that  $D$  is dynamically consistent. By induction on  $k + m$ , the proof is complete.  $\blacksquare$

## 8.2 Proof of Theorem 4.1

**Proof of Theorem 4.1.** Let  $(D^k)_{k=0}^\infty$  be the  $f$ -rationalizability procedure starting at  $D^{full}$ , and let  $K$  be such that  $D^{K+1} = D^K$ . By Lemma 8.6 we know that  $D^K$  is dynamically consistent. As such, Lemma 8.5 (a) guarantees that for every player  $i$  there is a strategy  $s_i$  such that  $s_i \in D_i^{K*}(h_i)$  for all  $h_i \in H_i(s_i)$ . Since, by Lemma 8.1,  $D_i^{K*}(h_i) = D_i^K(h_i)$  for all players  $i$  and all  $h_i \in H_i$ , we conclude that  $s_i \in D_i^K(h_i)$  for all  $h_i \in H_i(s_i)$ . That is,  $s_i$  is  $f$ -rationalizable. This completes the proof.  $\blacksquare$

## 8.3 Preparatory Results for Theorem 5.1

Our first result shows that the set of rationalizable strategies will not change if we restrict the reasoning at  $h$  to information sets that either weakly follow, or precede,  $h$ .

**Lemma 8.7 (Focus on future and past)** *Let  $f, g$  be two focus mappings with  $g \subseteq f$  that are individually forward decreasing and individually preserving focus on past information sets. Let  $\hat{f}, \hat{g}$  be the focus mappings where  $\hat{f}_{ij}(h) := f_{ij}(h) \cap H(h)$  and  $\hat{g}_{ij}(h) := g_{ij}(h) \cap H(h)$  for all information sets  $h \in H$ , and all players  $i \neq j$ . Then, for every player the sets of  $(g, f)$ - and  $(\hat{g}, \hat{f})$ -rationalizable strategies are the same, the sets of  $f$ - and  $\hat{f}$ -rationalizable strategies are the same, and the sets of  $g$ - and  $\hat{g}$ -rationalizable strategies are the same.*

**Proof.** For every  $k, m \geq 0$ , let  $D^{k,m} := (rf^k \circ rg^m)(D^{full})$  and  $\hat{D}^{k,m} := (r\hat{f}^k \circ r\hat{g}^m)(D^{full})$ . Define

$$D_{-i}^{k,m}(h_i)|_{H(h_i)} := \{(s_j|_{H(h_i)})_{j \neq i} \mid (s_j)_{j \neq i} \in D_{-i}^{k,m}(h_i)\},$$

and similarly for  $\hat{D}_{-i}^{k,m}(h_i)|_{H(h_i)}$ .

*Claim.* For every  $k, m \geq 0$ , player  $i$  and information set  $h_i \in H_i$ , we have that  $D_{-i}^{k,m}(h_i)|_{H(h_i)} = \hat{D}_{-i}^{k,m}(h_i)|_{H(h_i)}$  and  $D_i^{k,m}(h_i) = \hat{D}_i^{k,m}(h_i)$ .

*Proof of claim.* We show the statements by induction on  $k + m$ .

**Induction start.** If  $k + m = 0$ , then  $D^{0,0} = \hat{D}^{0,0} = D^{full}$ , and the statements trivially hold.

**Induction step.** Suppose that  $k + m \geq 1$ , and that the statements hold for  $k + m - 1$ .

We first show that  $D_{-i}^{k,m}(h_i)|_{H(h_i)} = \hat{D}_{-i}^{k,m}(h_i)|_{H(h_i)}$ . We distinguish two cases: (1)  $k + m = 1$ , and (2)  $k + m \geq 2$ .

**Case 1.** Suppose that  $k + m = 1$ . Then, either (1.1)  $k = 0$  and  $m = 1$ , or (1.2)  $k = 1$  and  $m = 0$ .

**Case 1.1.** If  $k = 0$  and  $m = 1$ , then  $D^{0,1} = rg(D^{full})$  and  $\hat{D}^{0,1} = r\hat{g}(D^{full})$ . Hence, by construction,  $D_{-i}^{0,1}(h_i) = \hat{D}_{-i}^{0,1}(h_i) = S_{-i}(h_i)$ , which implies that  $D_{-i}^{0,1}(h_i)|_{H(h_i)} = \hat{D}_{-i}^{0,1}(h_i)|_{H(h_i)}$  for all players  $i$  and  $h_i \in H_i$ .

**Case 1.2.** If  $k = 1$  and  $m = 0$ , then  $D^{1,0} = rf(D^{full})$  and  $\hat{D}^{1,0} = r\hat{f}(D^{full})$ . In a similar way as above, it then follows that  $D_{-i}^{1,0}(h_i)|_{H(h_i)} = \hat{D}_{-i}^{1,0}(h_i)|_{H(h_i)}$  for all players  $i$  and  $h_i \in H_i$ .

**Case 2.** Suppose that  $k + m \geq 2$ . We distinguish four subcases: (2.1)  $k \geq 1$  and  $D_{-i}^{k-1,m+}(h_i, f) \neq \emptyset$ , (2.2)  $k \geq 1$  and  $D_{-i}^{k-1,m+}(h_i, f) = \emptyset$ , (2.3)  $k = 0$  and  $D_{-i}^{0,m-1+}(h_i, g) \neq \emptyset$ , and (2.4)  $k = 0$  and  $D_{-i}^{0,m-1+}(h_i, g) = \emptyset$ .

**Case 2.1.** Assume that  $k \geq 1$  and  $D_{-i}^{k-1,m+}(h_i, f) \neq \emptyset$ . As, by the induction assumption,  $D_j^{k-1,m}(h_j) = \hat{D}_j^{k-1,m}(h_j)$  for all  $j \neq i$  and  $h_j \in H_j$ , and  $\hat{f} \subseteq f$ , it follows that  $D_{-i}^{k-1,m+}(h_i, f) \subseteq \hat{D}_{-i}^{k-1,m+}(h_i, \hat{f})$ . Therefore,  $\hat{D}_{-i}^{k-1,m+}(h_i, \hat{f}) \neq \emptyset$  also. We thus conclude that  $D_{-i}^{k,m}(h_i) = D_{-i}^{k-1,m+}(h_i, f) \subseteq \hat{D}_{-i}^{k-1,m+}(h_i, \hat{f}) = \hat{D}_{-i}^{k,m}(h_i)$ . In particular,  $D_{-i}^{k,m}(h_i)|_{H(h_i)} \subseteq \hat{D}_{-i}^{k,m}(h_i)|_{H(h_i)}$ .

To show that  $\hat{D}_{-i}^{k,m}(h_i)|_{H(h_i)} \subseteq D_{-i}^{k,m}(h_i)|_{H(h_i)}$ , take some  $(s_j)_{j \neq i} \in \hat{D}_{-i}^{k,m}(h_i)$ . Fix a player  $j \neq i$ . As  $\hat{D}_{-i}^{k,m}(h_i) = \hat{D}_{-i}^{k-1,m+}(h_i, \hat{f})$ , we have that  $s_j \in S_j(h_i)$  and  $s_j \in \hat{D}_j^{k-1,m}(h_j)$  for all  $h_j \in \hat{f}_{ij}(h_i) \cap H_j(s_j)$ . By the induction assumption,  $\hat{D}_j^{k-1,m}(h_j) = D_j^{k-1,m}(h_j)$ , and hence  $s_j \in D_j^{k-1,m}(h_j)$  for all  $h_j \in \hat{f}_{ij}(h_i) \cap H_j(s_j)$ . Moreover, as  $k + m \geq 2$  we know by Lemma 8.1 that  $D_j^{k-1,m}(h_j) = D_j^{k-1,m*}(h_j)$ , and thus  $s_j \in D_j^{k-1,m*}(h_j)$  for all  $h_j \in \hat{f}_{ij}(h_i) \cap H_j(s_j)$ . By Lemma 8.6 we know that  $D^{k-1,m}$  is dynamically consistent. Hence, we can use Lemma 8.4 (d), choosing  $h^* = h_i$ , to show that there is a strategy  $s_j^* \in S_j(h_i)$  with  $s_j^*|_{H(h_i)} = s_j|_{H(h_i)}$  such that

$$s_j^* \in D_j^{k-1,m*}(h_j) \text{ for all } h_j \in (\hat{f}_{ij}(h_i) \cap H_j(s_j) \cap H(h_i)) \cup (H_j(s_j^*) \setminus H(h_i)). \quad (8.6)$$

Note that, by definition,  $\hat{f}_{ij}(h_i) \cap H(h_i) = f_{ij}(h_i) \cap H(h_i)$ . Moreover, as  $s_j^*|_{H(h_i)} = s_j|_{H(h_i)}$  it follows that  $H_j(s_j) \cap H(h_i) = H_j(s_j^*) \cap H(h_i)$ . Together with (8.6) we conclude that  $s_j^* \in D_j^{k-1,m*}(h_j)$  for all  $h_j \in f_{ij}(h_i) \cap H_j(s_j^*)$ . Since, by Lemma 8.1,  $D_j^{k-1,m*}(h_j) = D_j^{k-1,m}(h_j)$  for all  $h_j \in H_j$ , we see that  $s_j^* \in D_j^{k-1,m}(h_j)$  for all  $h_j \in f_{ij}(h_i) \cap H_j(s_j^*)$ . As this holds for all  $j \neq i$ , and  $s_j^* \in S_j(h_i)$  for all  $j \neq i$ , we conclude that  $(s_j^*)_{j \neq i} \in D_{-i}^{k-1,m+*}(h_i, f)$ . By Lemma 8.2 we know that  $D_{-i}^{k-1,m+*}(h_i, f) = D_{-i}^{k-1,m+}(h_i, f)$ , and hence  $(s_j^*)_{j \neq i} \in D_{-i}^{k-1,m+}(h_i, f) = D_{-i}^{k,m}(h_i)$ . Moreover, since  $s_j^*|_{H(h_i)} = s_j|_{H(h_i)}$  for all

$j \neq i$ , it follows that  $(s_j|_{H(h_i)})_{j \neq i} = (s_j^*|_{H(h_i)})_{j \neq i} \in D_{-i}^{k,m}(h_i)|_{H(h_i)}$ . As this holds for every  $(s_j)_{j \neq i} \in \hat{D}_{-i}^{k,m}(h_i)$ , we conclude that  $\hat{D}_{-i}^{k,m}(h_i)|_{H(h_i)} \subseteq D_{-i}^{k,m}(h_i)|_{H(h_i)}$ . Altogether, we see that  $D_{-i}^{k,m}(h_i)|_{H(h_i)} = \hat{D}_{-i}^{k,m}(h_i)|_{H(h_i)}$ .

**Case 2.2.** Suppose that  $k \geq 1$  and  $D_{-i}^{k-1,m+}(h_i, f) = \emptyset$ . We show that  $\hat{D}_{-i}^{k-1,m+}(h_i, \hat{f}) = \emptyset$ . Assume, contrary to what we want to show, that  $\hat{D}_{-i}^{k-1,m+}(h_i, \hat{f}) \neq \emptyset$ . Then, there is some  $(s_j)_{j \neq i} \in \hat{D}_{-i}^{k-1,m+}(h_i, \hat{f})$ . Following the steps in Case 2.1, we can then construct a strategy profile  $(s_j^*)_{j \neq i} \in D_{-i}^{k-1,m+}(h_i, f)$ , which would contradict our assumption that  $D_{-i}^{k-1,m+}(h_i, f) = \emptyset$ . Hence,  $\hat{D}_{-i}^{k-1,m+}(h_i, \hat{f}) = \emptyset$ . This means that  $D_{-i}^{k,m}(h_i) = D_{-i}^{k-1,m}(h_i)$  and  $\hat{D}_{-i}^{k,m}(h_i) = \hat{D}_{-i}^{k-1,m}(h_i)$ . As, by the induction assumption,  $D_{-i}^{k-1,m}(h_i)|_{H(h_i)} = \hat{D}_{-i}^{k-1,m}(h_i)|_{H(h_i)}$ , it follows that  $D_{-i}^{k,m}(h_i)|_{H(h_i)} = \hat{D}_{-i}^{k,m}(h_i)|_{H(h_i)}$ .

Cases 2.3 and 2.4 can be shown in a similar fashion. By Cases 1.1, 1.2, and 2.1–2.4, it follows that  $D_{-i}^{k,m}(h_i)|_{H(h_i)} = \hat{D}_{-i}^{k,m}(h_i)|_{H(h_i)}$ .

We next show that  $D_i^{k,m}(h_i) = \hat{D}_i^{k,m}(h_i)$ . Note that every strategy profile  $(s_i, s_{-i}) \in S_i(h_i) \times S_{-i}(h_i)$  only reaches histories in  $H(h_i)$ . As  $D_{-i}^{k,m}(h_i), \hat{D}_{-i}^{k,m}(h_i) \subseteq S_{-i}(h_i)$  and  $D_{-i}^{k,m}(h_i)|_{H(h_i)} = \hat{D}_{-i}^{k,m}(h_i)|_{H(h_i)}$ , it follows that a strategy  $s_i \in S_i(h_i)$  is strictly dominated in  $(S_i(h_i), D_{-i}^{k,m}(h_i))$  if and only if it is strictly dominated in  $(S_i(h_i), \hat{D}_{-i}^{k,m}(h_i))$ . Hence,  $D_i^{k,m*}(h_i) = \hat{D}_i^{k,m*}(h_i)$ . Since  $k + m \geq 1$ , it follows by Lemma 8.1 that  $D_i^{k,m*}(h_i) = D_i^{k,m}(h_i)$  and  $\hat{D}_i^{k,m*}(h_i) = \hat{D}_i^{k,m}(h_i)$ , and hence we conclude that  $D_i^{k,m}(h_i) = \hat{D}_i^{k,m}(h_i)$ . By induction on  $k + m$ , the proof of the claim is complete.  $\diamond$

On the basis of the claim, we see that the sets of  $(g, f)$ - and  $(\hat{g}, \hat{f})$ -rationalizable strategies coincide, and that the same holds for  $f$ - versus  $\hat{f}$ -rationalizability, and  $g$ - versus  $\hat{g}$ -rationalizability.  $\blacksquare$

We say that a focus mapping  $f$  is *past and future oriented* if  $f_{ij}(h) \subseteq H(h)$  for all information sets  $h \in H$  and all players  $i \neq j$ . By Lemma 8.7 we can thus assume, without loss of generality, that  $f$  is past and future oriented if we are only interested in the associated rationalizable strategies.

For two focus functions  $f, g$  with  $g \subseteq f$ , a decision problem map  $D$  is called  $(g, f)$ -generated if  $D = (rf^k \circ rg^m)(D^{full})$  for some  $k, m \geq 0$ . For a  $(g, f)$ -generated decision problem map  $D$ , player  $i$  and information set  $h \in H$ , we define the set

$$D_i^+(h, f) := \{s_i \in S_i(h) \mid s_i \in D_i(h'_i) \text{ for some } j \neq i \text{ and all } h'_i \in f_{ji}(h) \cap H_i(s_i)\}.$$

If  $h \in H_i$  then we additionally define the set

$$D_i^-(h, f) := \{s_i \in D_i(h) \mid s_i \in D_i(h'_i) \text{ for some } j \neq i \text{ and all } h'_i \in f_{ji}(h) \cap H_i^-(h)\}.$$

Note that  $f_{ji}(h) = f_{ki}(h)$  for all  $j, k \neq i$  if  $f$  is forward decreasing, and hence it does not matter which opponent  $j \neq i$  we consider in the definitions of  $D_i^+(h, f)$  and  $D_i^-(h, f)$ . As the dynamic game is with observable actions, it follows by Lemma 8.2 that for every player  $i$  and information set  $h_i \in H_i$  we have  $D_{-i}^+(h_i, f) = \times_{j \neq i} D_j^+(h_i, f)$ .

By

$$H^e(D, f) := \{h \in H \mid D_i^+(h, f) \neq \emptyset \text{ for all } i\}$$

we denote the collection of *explicable* information sets at the decision problem map  $D$  and focus function  $f$ . For a set of strategy combinations  $\hat{S} \subseteq \times_{i \in I} S_i$  we define  $H(\hat{S}) := \{h \in H \mid \text{there is some } s \in \hat{S} \text{ that reaches } h\}$ . In the lemma below, we set  $D^{rf} := rf(D)$ .

**Lemma 8.8 (Explicability is closed under optimal behavior)** *Consider two focus functions  $f, g$  with  $g \subseteq f$  that satisfy the conditions in Theorem 5.1, and let  $D$  be a  $(g, f)$ -generated decision problem map. Then,*

- (a) *for every information set  $h \in H^e(D, f)$  we have  $H(\times_{i \in I} D_i^+(h, f)) \subseteq H^e(D, f)$ ,*
- (b) *the set  $H^e(D, f)$  is closed under predecessors, and*
- (c) *for every player  $i$ , information set  $h_i \in H^e(D, f) \cap H_i$  and strategy  $s_i \in D_i^{rf-}(h_i, f)$ , there is a belief  $b_i \in \Delta(D_{-i}^{rf}(h_i))$  such that  $s_i$  is optimal for  $b_i$  in  $S_i(h_i)$  and  $H(s_i, b_i) \subseteq H^e(D, f)$ .*

**Proof.** (a) Take some  $h \in H^e(D, f)$ , and some  $h' \in H(\times_{i \in I} D_i^+(h, f))$ . Hence, there is some  $(s_i)_{i \in I} \in \times_{i \in I} D_i^+(h, f)$  such that  $(s_i)_{i \in I}$  reaches  $h'$ . Moreover,  $(s_i)_{i \in I}$  reaches  $h$  as well since  $D_i^+(h, f) \subseteq S_i(h)$  for all  $i$ . Then, either  $h'$  weakly follows  $h$  or  $h'$  precedes  $h$ . We will show that  $h'$  is reached by some strategy profile in  $\times_{i \in I} D_i^+(h', f)$ , thereby showing that  $h' \in H^e(D, f)$ . To do so we distinguish two cases: (1)  $h'$  weakly follows  $h$  and (2)  $h'$  precedes  $h$ .

**Case 1.** Suppose that  $h'$  weakly follows  $h$ . For an arbitrary player  $i$  we show that  $s_i \in D_i^+(h', f)$ . Since  $s_i \in D_i^+(h, f)$  we know, by definition, that  $s_i \in D_i(h''_i)$  for some  $j \neq i$  and all  $h''_i \in f_{ji}(h) \cap H_i(s_i)$ . As  $h'$  weakly follows  $h$  and  $f$  is forward decreasing we know that  $f_{ji}(h') \subseteq f_{ji}(h)$ . As such,  $s_i \in D_i(h''_i)$  for all  $h''_i \in f_{ji}(h') \cap H_i(s_i)$ . Moreover,

$s_i \in S_i(h')$  as  $(s_i)_{i \in I}$  reaches  $h'$ . Therefore,  $s_i \in D_i^+(h', f)$ . Thus, we see that  $s_i \in D_i^+(h', f)$  for all  $i$ , and hence  $h' \in H^e(D, f)$ .

**Case 2.** Suppose that  $h'$  precedes  $h$ . For an arbitrary player  $i$  we show that we can construct some  $s_i^* \in D_i^+(h', f)$ . Since  $s_i \in D_i^+(h, f)$  we know, by definition, that  $s_i \in D_i(h_i'')$  for some  $j \neq i$  and all  $h_i'' \in f_{ji}(h) \cap H_i(s_i)$ . As we know, by Lemma 8.1, that  $D_i(h_i'') = D_i^*(h_i'')$ , we see for this  $j \neq i$  that

$$s_i \in D_i^*(h_i'') \text{ for all } h_i'' \in f_{ji}(h) \cap H_i(s_i). \quad (8.7)$$

By Lemma 8.6 we know that  $D$  is dynamically consistent. But then, it follows from (8.7) and Lemma 8.4 (a), setting  $h^* = h'$ , that there is some  $s_i^* \in S_i(h')$  such that  $s_i^* \in D_i^*(h_i'')$  for every  $h_i'' \in (f_{ji}(h) \cap H_i^-(h')) \cup (H_i(s_i^*) \setminus H_i^-(h'))$ . By Lemma 8.1 we know that  $D_i^*(h_i'') = D_i(h_i'')$ , and hence  $s_i^* \in D_i(h_i'')$  for every  $h_i'' \in (f_{ji}(h) \cap H_i^-(h')) \cup (H_i(s_i^*) \setminus H_i^-(h'))$ .

As  $h'$  precedes  $h$  and  $f$  preserves focus on past information sets, we conclude that  $f_{ji}(h') \cap H_i^-(h') \subseteq f_{ji}(h)$ , and hence  $s_i^* \in D_i(h_i'')$  for every  $h_i'' \in (f_{ji}(h') \cap H_i^-(h')) \cup (H_i(s_i^*) \setminus H_i^-(h'))$ . In particular,  $s_i^* \in D_i(h_i'')$  for every  $h_i'' \in f_{ji}(h') \cap H_i(s_i^*)$ . As  $s_i^* \in S_i(h')$  we conclude that  $s_i^* \in D_i^+(h', f)$ . We thus see that  $s_i^* \in D_i^+(h', f)$  for all  $i$ , and hence  $h' \in H^e(D, f)$ . By Cases 1 and 2, the proof of (a) is complete.

**(b)** Take an information set  $h \in H^e(D, f)$  and a predecessor  $h'$  of  $h$ . Take some  $(s_i)_{i \in I} \in \times_{i \in I} D_i^+(h, f)$ . Since  $s_i \in S_i(h)$  for all  $i$ , we conclude that  $(s_i)_{i \in I}$  reaches  $h$ . As  $h'$  precedes  $h$ , the strategy combination  $(s_i)_{i \in I}$  must reach  $h'$  as well. Recall that  $(s_i)_{i \in I} \in \times_{i \in I} D_i^+(h, f)$ , which implies that  $h' \in H(\times_{i \in I} D_i^+(h, f))$ . By (a) we then conclude that  $h' \in H^e(D, f)$ . As such,  $H^e(D, f)$  is closed under predecessors.

**(c)** Take an information set  $h_i^* \in H^e(D, f) \cap H_i$  and a strategy  $s_i \in D_i^{rf-}(h_i^*, f)$ . Let  $H_i^* := \{h_i \in H_i(s_i) \mid s_i \in D_i^{rf}(h_i)\}$ . By Lemma 8.1 we know that  $D_i^{rf}(h_i) = D_i^{rf*}(h_i)$  for all  $h_i \in H_i$ , and hence  $H_i^* = \{h_i \in H_i(s_i) \mid s_i \in D_i^{rf*}(h_i)\}$ . As  $s_i \in D_i^{rf-}(h_i^*, f)$  we know that  $s_i \in D_i^{rf}(h_i^*)$  and hence  $h_i^* \in H_i^*$ . Moreover, it follows by Lemma 8.6 that  $D^{rf}$  is dynamically consistent. By Lemma 8.4 (a) and (b) applied to  $D^{rf}$  there is thus a strategy  $s_i^*$  and a belief  $b_i \in \Delta(D_{-i}^{rf}(h_i^*))$  such that (i)  $s_i$  is optimal for  $b_i$  in  $S_i(h_i^*)$ , (ii)  $s_i^* \in D_i^{rf*}(h_i)$  for all  $h_i \in (H_i^* \cap H_i^-(h^*)) \cup (H_i(s_i^*) \setminus H_i^-(h^*))$ , and (iii)  $s_i^*|_{H(s_i, b_i)} = s_i|_{H(s_i, b_i)}$ .

We first show that  $s_i^* \in S_i(h_i^*)$ . Recall that  $s_i \in S_i(h_i^*)$  and  $b_i \in \Delta(S_{-i}(h_i^*))$ , which implies that  $H_i^-(h_i^*) \subseteq H(s_i, b_i)$ . As  $s_i^*|_{H(s_i, b_i)} = s_i|_{H(s_i, b_i)}$  it follows that  $s_i^* \in S_i(h_i^*)$  also.

We next show that  $s_i^* \in D_i^+(h_i^*, f)$ . Recall that  $s_i \in D_i^{rf-}(h_i^*, f)$ , which means that  $s_i \in D_i^{rf}(h_i^*)$  and  $s_i \in D_i^{rf}(h_i)$  for all  $h_i \in f_{ji}(h_i^*) \cap H_i^-(h_i^*)$  and some  $j \neq i$ . Hence,

$f_{ji}(h_i^*) \cap H_i^-(h_i^*) \subseteq H_i^*$ . By (ii) above it follows that  $s_i^* \in D_i^{rf*}(h_i) = D_i^{rf}(h_i)$  for all  $h_i \in f_{ji}(h_i^*) \cap H_i(s_i^*)$ . Since we have seen that  $s_i^* \in S_i(h_i^*)$ , it follows that  $s_i^* \in D_i^{rf+}(h_i^*, f)$ . As  $D^{rf} \subseteq D$  we conclude, in particular, that  $s_i^* \in D_i^+(h_i^*, f)$ .

We finally prove that  $H(s_i, b_i) \subseteq H^e(D, f)$ . As, by (iii) above,  $s_i^*|_{H(s_i, b_i)} = s_i|_{H(s_i, b_i)}$ , it follows that  $H(s_i, b_i) = H(s_i^*, b_i)$ . Recall from above that  $s_i^* \in D_i^+(h_i^*, f)$ . Moreover, as  $h_i^* \in H^e(D, f)$  we know that  $D_{-i}^{rf}(h_i^*) = D_{-i}^+(h_i^*, f) = \times_{j \neq i} D_j^+(h_i^*, f)$ . Since  $b_i \in \Delta(D_{-i}^{rf}(h_i^*))$  it follows that  $H(s_i^*, b_i) \subseteq H(\times_{j \in I} D_j^+(h_i^*, f))$ . Together with (a) we then conclude that  $H(s_i^*, b_i) \subseteq H^e(D, f)$ . Thus,  $H(s_i, b_i) = H(s_i^*, b_i) \subseteq H^e(D, f)$ . This completes the proof.  $\blacksquare$

**Lemma 8.9 (Forward decreasing sets of strategies)** *Consider a focus function  $f$  that is forward decreasing and individually preserving focus on past information sets, and a decision problem map  $D$ . Then, for every two players  $i \neq j$  and information set  $h_i \in H_i$ , we have that  $D_i^-(h_i, f) \subseteq D_i^-(h'_i, f)$  for all  $h'_i \in f_{ji}(h_i) \cap H_i^-(h_i)$ .*

**Proof.** For some arbitrary player  $j \neq i$ , let  $h'_i \in f_{ji}(h_i) \cap H_i^-(h_i)$ , and take some  $s_i \in D_i^-(h_i, f)$ . As  $f$  is forward decreasing we know that  $f_{ki}(h_i) = f_{ji}(h_i)$  for all  $k \neq i, j$ . Hence, we conclude by definition of  $D_i^-(h_i, f)$  that  $s_i \in D_i(h'_i)$ . Now, take some  $h''_i \in f_{ji}(h'_i) \cap H_i^-(h'_i)$ . As  $h'_i$  precedes  $h_i$ , and  $f$  individually preserves focus on past information sets, we know that  $f_{ji}(h'_i) \cap H_i^-(h'_i) \subseteq f_{ji}(h_i)$ , and therefore  $h''_i \in f_{ji}(h_i) \cap H_i^-(h_i)$ . Since  $s_i \in D_i^-(h_i, f)$  we conclude that  $s_i \in D_i(h''_i)$ . As this holds for all  $h''_i \in f_{ji}(h'_i) \cap H_i^-(h'_i)$ , and since we have seen above that  $s_i \in D_i(h'_i)$ , it follows that  $s_i \in D_i^-(h'_i, f)$ . Thus, we see that  $D_i^-(h_i, f) \subseteq D_i^-(h'_i, f)$ .  $\blacksquare$

Consider a focus function  $f$ , two decision problem maps  $D, E$  and a profile of transformation mappings  $\sigma = (\sigma_i)_{i \in I}$  where  $\sigma_i : S_i \rightarrow S_i$  for every player  $i$ . For some decision problem map  $D'$  we write  $D \subseteq_{\sigma, f, D'} E$  if either  $D' = D^{full}$  or  $D = rf(D')$ , and if for every player  $i$  and every strategy  $s_i \in S_i$  the following three properties hold:

- (T1)  $\sigma_i(s_i)|_{H^e(D', f)} = s_i|_{H^e(D', f)}$ ,
- (T2)  $\sigma_i(s_i) \in E_i(h_i)$  whenever  $h_i \in H^e(D', f) \cap H_i$  and  $s_i \in D_i^-(h_i, f)$ , and
- (T3)  $\sigma_i(s_i) \in E_i(h_i)$  whenever  $h_i \in H_i(\sigma_i(s_i)) \setminus H^e(D', f)$ .

Intuitively,  $D \subseteq_{\sigma, f, D'} E$  entails that every strategy “allowed by  $D$ ” can be transformed into a strategy “allowed by  $E$ ” while leaving its behavior on the explicable information sets under  $D'$  and  $f$  unchanged.

**Lemma 8.10 (Properties of transformation mappings)** *Consider two focus functions  $f, g$  with  $g \subseteq f$  that satisfy the conditions in Theorem 5.1, and assume that  $f$  is past and future oriented. Take two decision problem maps  $D, E$  such that  $D \subseteq_{\sigma, f, D'} E$  for some  $\sigma$  and  $D'$ , and where  $D'$  is  $(g, f)$ -generated. Then, for every player  $i$  and strategy  $s_i$ ,*

- (a)  $\sigma_i(s_i) \in E_i^-(h_i, f)$  whenever  $h_i \in H^e(D', f) \cap H_i$  and  $s_i \in D_i^-(h_i, f)$ , and
- (b)  $\sigma_i(s_i) \in E_i^+(h, f)$  whenever  $h \in H^e(D', f)$  and  $s_i \in D_i^+(h, f)$ .

**Proof.** (a) Suppose that  $h_i \in H^e(D', f) \cap H_i$  and  $s_i \in D_i^-(h_i, f)$ . We show that  $\sigma_i(s_i) \in E_i^-(h_i, f)$ . As  $s_i \in D_i^-(h_i, f)$  we know by property (T2) that  $\sigma_i(s_i) \in E_i(h_i)$ . Consider now some  $j \neq i$  and some  $h'_i \in f_{ji}(h_i) \cap H_i^-(h_i)$ . By Lemma 8.9 we know that  $D_i^-(h_i, f) \subseteq D_i^-(h'_i, f)$ , and hence  $s_i \in D_i^-(h'_i, f)$ . Moreover, since  $h'_i$  precedes  $h_i$  and  $h_i \in H^e(D', f)$ , we conclude by Lemma 8.8 (b) that  $h'_i \in H^e(D', f)$  as well. By (T2) it then follows that  $\sigma_i(s_i) \in E_i(h'_i)$ . As this holds for all  $h'_i \in f_{ji}(h_i) \cap H_i^-(h_i)$ , and since we have seen above that  $\sigma_i(s_i) \in E_i(h_i)$ , we conclude that  $\sigma_i(s_i) \in E_i^-(h_i, f)$ , which was to show.

(b) Suppose that  $h \in H^e(D', f)$  and  $s_i \in D_i^+(h, f)$ . We show that  $\sigma_i(s_i) \in E_i^+(h, f)$ . Take some  $j \neq i$  and some  $h_i \in f_{ji}(h) \cap H_i(\sigma_i(s_i))$ . If  $h_i \notin H^e(D', f)$ , then it follows by (T3) that  $\sigma_i(s_i) \in E_i(h_i)$ .

Suppose next that  $h_i \in H^e(D', f)$ . As, by Lemma 8.8 (b),  $H^e(D', f)$  is closed under predecessors, it follows that  $H_i^-(h_i) \subseteq H^e(D', f)$ . By (T1) we thus know that  $s_i|_{H_i^-(h_i)} = \sigma_i(s_i)|_{H_i^-(h_i)}$ . Since  $h_i \in H_i(\sigma_i(s_i))$  we conclude that  $h_i \in H_i(s_i)$  as well. Hence,  $h_i \in f_{ji}(h) \cap H_i(s_i)$ .

We will now show that  $s_i \in D_i^-(h_i, f)$ . As  $h_i \in f_{ji}(h) \cap H_i(s_i)$  and  $s_i \in D_i^+(h, f)$ , we conclude that  $s_i \in D_i(h_i)$ . Consider some  $h'_i \in f_{ji}(h_i) \cap H_i^-(h_i)$ . We will show that  $h'_i \in f_{ji}(h)$ . Since  $h_i \in f_{ji}(h)$  and  $f$  is past and future oriented, either  $h_i$  weakly follows  $h$ , or  $h_i$  precedes  $h$ .

Suppose first that  $h_i$  weakly follows  $h$ . As  $f$  is individually forward decreasing we know that  $f_{ji}(h_i) \subseteq f_{ji}(h)$ . Since  $h'_i \in f_{ji}(h_i) \cap H_i^-(h_i)$  it follows that  $h'_i \in f_{ji}(h)$ . Assume next that  $h_i$  precedes  $h$ . As  $f$  is individually preserving focus on past information sets we have that  $f_{ji}(h_i) \cap H_i^-(h_i) \subseteq f_{ji}(h)$ . Since  $h'_i \in f_{ji}(h_i) \cap H_i^-(h_i)$ , we conclude that  $h'_i \in f_{ji}(h)$ .

We thus know that  $h'_i \in f_{ji}(h)$ . Since we assumed that  $h'_i \in f_{ji}(h_i) \cap H_i^-(h_i)$ , we conclude that  $h'_i \in f_{ji}(h) \cap H_i^-(h_i)$ . Moreover, as  $s_i \in S_i(h_i)$  and  $h'_i \in H_i^-(h_i)$  it follows that  $h'_i \in H_i(s_i)$ . We thus see that  $h'_i \in f_{ji}(h) \cap H_i(s_i)$ . Since  $s_i \in D_i^+(h, f)$  we conclude

that  $s_i \in D_i(h'_i)$ . As this holds for every  $h'_i \in f_{ji}(h_i) \cap H_i^-(h_i)$ , and since we have seen that  $s_i \in D_i(h_i)$ , we see that  $s_i \in D_i^-(h_i, f)$ . Since, by the assumption above,  $h_i \in H^e(D', f)$ , it follows by (T2) that  $\sigma_i(s_i) \in E_i(h_i)$ .

Altogether, we conclude that  $\sigma_i(s_i) \in E_i(h_i)$  for all  $h_i \in f_{ji}(h) \cap H_i(\sigma_i(s_i))$ , and hence  $\sigma_i(s_i) \in E_i^+(h, f)$ . ■

**Lemma 8.11 (Explicability is monotone)** *Consider two focus functions  $f, g$  with  $g \subseteq f$  that satisfy the conditions in Theorem 5.1, and assume that  $f$  is past and future oriented. Take two decision problem maps  $D, E$  such that  $D \subseteq_{\sigma, f, D'} E$  for some  $\sigma$  and  $D'$ , and where  $D'$  is  $(g, f)$ -generated. Then,  $H^e(D, f) \subseteq H^e(E, f)$ .*

**Proof.** Take an information set  $h \in H^e(D, f)$  and a strategy combination  $(s_i)_{i \in I} \in \times_{i \in I} D_i^+(h, f)$ . As  $D \subseteq D'$  we have that  $H^e(D, f) \subseteq H^e(D', f)$ , and hence  $h \in H^e(D', f)$ . By Lemma 8.10 (b) it thus follows that  $(\sigma_i(s_i))_{i \in I} \in \times_{i \in I} E_i^+(h, f)$ . In particular,  $E_i^+(h, f) \neq \emptyset$  for all players  $i$ , and hence  $h \in H^e(E, f)$ . Thus,  $H^e(D, f) \subseteq H^e(E, f)$ . ■

Consider two profiles of transformation mappings  $\tau = (\tau_i)_{i \in I}$  and  $\rho = (\rho_i)_{i \in I}$  where  $\tau_i, \rho_i : S_i \rightarrow S_i$  for every player  $i$ . Then,  $\rho \circ \tau$  denotes the profile of transformation mappings that defines for every player  $i$  the transformation mapping  $\rho_i \circ \tau_i : S_i \rightarrow S_i$ .

**Lemma 8.12 (Transitivity of inclusion operator)** *Consider two focus functions  $f, g$  with  $g \subseteq f$  that satisfy the conditions in Theorem 5.1, and assume that  $f$  is past and future oriented. Consider three decision problem maps  $D, E, F$  where  $D$  is  $(g, f)$ -generated, such that  $D \subseteq_{\tau, f, D'} E \subseteq_{\rho, f, E'} F$  for some  $\tau, \rho$  and  $D', E'$  with  $H^e(D', f) = H^e(E', f)$ , and where  $D', E'$  are  $(g, f)$ -generated. Then,  $D \subseteq_{\rho \circ \tau, f, D'} F$ .*

**Proof.** By relying on Lemma 8.10 (a) the proof is rather elementary, and is therefore left to the reader. ■

For two focus functions  $f, g$ , and two decision problem maps  $D, E$  where  $E = (rf^k \circ rg^m)(D^{full})$  for some  $k, m \geq 0$ , we say that  $D$  is a  $(g, f)$ -semi reduction of  $E$  if either  $D = (rf^k \circ rg^{m+1})(D^{full})$ , or  $D = (rf^{k+1} \circ rg^{m-1})(D^{full})$  (provided  $m \geq 1$ ).

**Lemma 8.13 (Monotonicity on explicable information sets)** *Consider two focus functions  $f, g$  with  $g \subseteq f$  that satisfy the conditions in Theorem 5.1 and assume, in addition,*

that  $f$  and  $g$  are past and future oriented. Suppose that  $E = (rf^k \circ rg^m)(D^{full})$  for some  $k, m \geq 0$ , that  $D$  is a  $(g, f)$ -semi reduction of  $E$ , and let

$$rf(E) \subseteq_{\tau, f, E} D \subseteq_{\rho, f, D'} E$$

for some  $\tau, \rho$  and  $D'$  that is  $(g, f)$ -generated. Then,

$$rf(D) \subseteq_{\sigma, f, D} rf(E)$$

for some  $\sigma$ .

**Proof.** Set  $D^{rf} := rf(D)$  and  $E^{rf} := rf(E)$ . Hence, we must show that  $D^{rf} \subseteq_{\sigma, f, D} E^{rf}$  for some  $\sigma$ . By Lemma 8.6 we know that  $E^{rf}$  is dynamically consistent.

For every player  $i$  we construct the mapping  $\sigma_i : S_i \rightarrow S_i$  as follows. Since  $E^{rf}$  is dynamically consistent, we know by Lemma 8.5 (b) that there is for every information set  $h_i \in (H_i \setminus H^e(D, f))^{first}$  a strategy  $s_i^E[h_i]$  such that  $s_i^E[h_i] \in E_i^{rf*}(h'_i)$  for all  $h'_i \in H_i(s_i^E[h_i]) \setminus H_i^-(h_i)$ . Hence, in particular,  $s_i^E[h_i] \in E_i^{rf*}(h'_i)$  for all  $h'_i \in H_i(s_i^E[h_i]) \cap H_i^+(h_i)$ . Since we know from Lemma 8.1 that  $E_i^{rf*}(h'_i) = E_i^{rf}(h'_i)$  for all  $h'_i \in H_i$ , it follows that  $s_i^E[h_i] \in E_i^{rf}(h'_i)$  for all  $h'_i \in H_i(s_i^E[h_i]) \cap H_i^+(h_i)$ .

For every strategy  $s_i \in S_i$  we define  $\sigma_i(s_i)$  to be the unique strategy where

$$\sigma_i(s_i)|_{H^e(D, f)} = s_i|_{H^e(D, f)} \text{ and} \tag{8.8}$$

$$\sigma_i(s_i)|_{H_i^+(h_i)} = s_i^E[h_i]|_{H_i^+(h_i)} \text{ for all } h_i \in (H_i \setminus H^e(D, f))^{first}. \tag{8.9}$$

This construction is well-defined since, by Lemma 8.8 (b), the collection  $H^e(D, f)$  is closed under predecessors, and hence the collection  $H \setminus H^e(D, f)$  is closed under weak followers.

In view of (8.8) and (8.9) we conclude that  $\sigma$  satisfies properties (T1) and (T3). It remains to prove property (T2).

**(T2)** We say that (T2) holds for player  $i$  at information set  $h_i \in H^e(D, f) \cap H_i$  if for every  $s_i \in D_i^{rf-}(h_i, f)$  we have that  $\sigma_i(s_i) \in E_i^{rf}(h_i)$ . Suppose, contrary to what we want to show, that (T2) does not hold for some player  $i$  and information set  $h_i \in H^e(D, f) \cap H_i$ . Since  $H^e(D, f)$  is closed under predecessors, there is some  $h_i^* \in H^e(D, f) \cap H_i$  such that (T2) does not hold at  $h_i^*$ , but holds at every  $h_i \in H_i^-(h_i^*)$ .

Since (T2) does not hold at  $h_i^*$ , there is a strategy  $s_i \in D_i^{rf-}(h_i^*, f)$  with  $\sigma_i(s_i) \notin E_i^{rf}(h_i^*)$ . However, for some  $j \neq i$  and every  $h_i \in f_{ji}(h_i^*) \cap H_i^-(h_i^*)$ , we have that  $\sigma_i(s_i) \in$

$E_i^{rf}(h_i)$ . To see this, take some  $h_i \in f_{ji}(h_i^*) \cap H_i^-(h_i^*)$ . Then, by Lemma 8.9,  $D_i^{rf-}(h_i^*, f) \subseteq D_i^{rf-}(h_i, f)$ . Since  $s_i \in D_i^{rf-}(h_i^*, f)$  we know that  $s_i \in D_i^{rf-}(h_i, f)$ . As we assume that (T2) holds at  $h_i$ , it follows that  $\sigma_i(s_i) \in E_i^{rf}(h_i)$ . We thus see that  $\sigma_i(s_i) \notin E_i^{rf}(h_i^*)$  but  $\sigma_i(s_i) \in E_i^{rf}(h_i)$  for all  $h_i \in f_{ji}(h_i^*) \cap H_i^-(h_i^*)$ .

*Claim 1.* There is a belief  $b_i^D \in \Delta(D_{-i}^{rf}(h_i^*))$  such that  $s_i$  is optimal in  $S_i(h_i^*)$  for  $b_i^D$ , and  $H(s_i, b_i^D) \subseteq H^e(D, f)$ .

*Proof of Claim 1.* Since  $h_i^* \in H^e(D, f) \cap H_i$  and  $s_i \in D_i^{rf-}(h_i^*, f)$ , it follows by Lemma 8.8 (c) that there is some  $b_i^D \in \Delta(D_{-i}^{rf}(h_i^*))$  such that  $s_i$  is optimal in  $S_i(h_i^*)$  for  $b_i^D$ , and  $H(s_i, b_i^D) \subseteq H^e(D, f)$ .  $\diamond$

Recall from above that  $h_i^* \in H^e(D, f)$ . Hence, we know that  $D_{-i}^+(h_i^*, f) = \times_{j \neq i} D_j^+(h_i^*, f) \neq \emptyset$  and thus, by definition,  $D_{-i}^{rf}(h_i^*) = D_{-i}^+(h_i^*, f)$ . Therefore, the belief  $b_i^D$  in Claim 1 is in  $\Delta(D_{-i}^+(h_i^*, f))$ . We know from above that  $D \subseteq_{\rho, f, D'} E$  for some  $\rho$  and  $D'$ , where  $\rho = (\rho_j)_{j \in I}$ . For every opponents' strategy combination  $s_{-i} = (s_j)_{j \neq i}$  we define  $\rho_{-i}(s_{-i}) := (\rho_j(s_j))_{j \neq i}$ . Define the belief  $b_i^E$  by

$$b_i^E(s_{-i}) := \sum_{s'_{-i} \in S_{-i} : \rho_{-i}(s'_{-i}) = s_{-i}} b_i^D(s'_{-i}). \quad (8.10)$$

*Claim 2.* It holds that  $b_i^E \in \Delta(E_{-i}^{rf}(h_i^*))$ .

*Proof of Claim 2.* Take some  $s_{-i} \in S_{-i}$  with  $b_i^E(s_{-i}) > 0$ . By (8.10) there is some  $s'_{-i} \in S_{-i}$  with  $\rho_{-i}(s'_{-i}) = s_{-i}$  such that  $b_i^D(s'_{-i}) > 0$ . Since  $b_i^D \in \Delta(D_{-i}^+(h_i^*, f))$  it must be that  $s'_{-i} \in D_{-i}^+(h_i^*, f)$ . That is,  $s'_j \in D_j^+(h_i^*, f)$  for all  $j \neq i$ . As  $h_i^* \in H^e(D, f)$ , we know from Lemma 8.10 (b) applied to  $\rho$  that  $\rho_j(s'_j) \in E_j^+(h_i^*, f)$  for all  $j \neq i$ , and hence  $s_{-i} = \rho_{-i}(s'_{-i}) \in E_{-i}^+(h_i^*, f)$ . As this holds for every  $s_{-i} \in S_{-i}$  with  $b_i^E(s_{-i}) > 0$ , we see that  $b_i^E \in \Delta(E_{-i}^+(h_i^*, f))$ . Hence,  $b_i^E \in \Delta(E_{-i}^{rf}(h_i^*))$ .  $\diamond$

*Claim 3.* We have that  $u_i(s_i, b_i^D) = u_i(\sigma_i(s_i), b_i^E)$ .

*Proof of Claim 3.* By definition,

$$\begin{aligned} u_i(\sigma_i(s_i), b_i^E) &= \sum_{s_{-i} \in S_{-i}} b_i^E(s_{-i}) \cdot u_i(z(\sigma_i(s_i), s_{-i})) \\ &= \sum_{s_{-i} \in S_{-i}} \sum_{s'_{-i} \in S_{-i} : \rho_{-i}(s'_{-i}) = s_{-i}} b_i^D(s'_{-i}) \cdot u_i(z(\sigma_i(s_i), s_{-i})), \end{aligned} \quad (8.11)$$

where the second equality follows from the definition of  $b_i^E$  in (8.10). Take some  $s_{-i}, s'_{-i} \in S_{-i}$  with  $\rho_{-i}(s'_{-i}) = s_{-i}$  and  $b_i^D(s'_{-i}) > 0$ . Then, by definition,  $(s_i, s'_{-i})$  only reaches

information sets in  $H(s_i, b_i^D)$ . Together with Claim 1 we conclude that

$$(s_i, s'_{-i}) \text{ only reaches information sets in } H^e(D, f). \quad (8.12)$$

By property (8.8) we know that

$$\sigma_i(s_i)|_{H^e(D, f)} = s_i|_{H^e(D, f)}. \quad (8.13)$$

Recall that  $D \subseteq_{\rho, f, D'} E$ . Let  $s'_{-i} = (s'_j)_{j \neq i}$ . By property (T1) of  $\rho$  it follows that  $\rho_j(s'_j)|_{H^e(D', f)} = s'_j|_{H^e(D', f)}$  for every  $j \neq i$ . Recall that  $D = rf(D')$  or  $D' = D^{full}$ . In either case we have that  $D \subseteq D'$ , which implies that  $H^e(D, f) \subseteq H^e(D', f)$ , and hence

$$\rho_j(s'_j)|_{H^e(D, f)} = s'_j|_{H^e(D, f)} \text{ for every } j \neq i. \quad (8.14)$$

By combining (8.12), (8.13) and (8.14), we see that

$$z(\sigma_i(s_i), s_{-i}) = z(s_i, s'_{-i}) \text{ whenever } \rho_{-i}(s'_{-i}) = s_{-i} \text{ and } b_i^D(s'_{-i}) > 0. \quad (8.15)$$

By combining (8.11) and (8.15) we get

$$\begin{aligned} u_i(\sigma_i(s_i), b_i^E) &= \sum_{s_{-i} \in S_{-i}} \sum_{s'_{-i} \in S_{-i}: \rho_{-i}(s'_{-i}) = s_{-i}} b_i^D(s'_{-i}) \cdot u_i(z(\sigma_i(s_i), s_{-i})) \\ &= \sum_{s'_{-i} \in S_{-i}} b_i^D(s'_{-i}) \cdot u_i(z(s_i, s'_{-i})) = u_i(s_i, b_i^D), \end{aligned}$$

which was to show.  $\diamond$

*Claim 4.* There is some  $s_i^* \in E_i^{rf+}(h_i^*, f)$  that is optimal for  $b_i^E$  in  $S_i(h_i^*)$ .

*Proof of Claim 4.* Recall our assumptions above that (i)  $\sigma_i(s_i) \notin E_i^{rf}(h_i^*) = E_i^{rf*}(h_i^*)$ , and (ii)  $\sigma_i(s_i) \in E_i^{rf}(h_i) = E_i^{rf*}(h_i)$  for all  $h_i \in f_{ji}(h_i^*) \cap H_i^-(h_i^*)$ . Let  $H_i^{**} := \{h_i \in H_i(\sigma_i(s_i)) \mid \sigma_i(s_i) \in E_i^{rf*}(h_i)\}$  and hence  $h_i^* \in H_i \setminus H_i^{**}$ . Recall that  $E^{rf}$  is dynamically consistent, and that  $b_i^E \in \Delta(E_{-i}^{rf}(h_i^*))$ . It then follows by Lemma 8.4 (a) and (c), setting  $h^* = h_i^*$ , that there is some  $s_i^* \in S_i(h_i^*)$  such that  $s_i^* \in E_i^{rf*}(h_i)$  for all  $h_i \in (H_i^{**} \cap H_i^-(h_i^*)) \cup (H_i(s_i^*) \setminus H_i^-(h_i^*))$ , and  $s_i^*$  is optimal for  $b_i^E$  in  $S_i(h_i^*)$ . By our assumption (ii) above we know that  $f_{ji}(h_i^*) \cap H_i^-(h_i^*) \subseteq H_i^{**}$ , and hence it follows that  $s_i^* \in E_i^{rf*}(h_i) = E_i^{rf}(h_i)$  for all  $h_i \in f_{ji}(h_i^*) \cap H_i(s_i^*)$ . We thus conclude that  $s_i^* \in E_i^{rf+}(h_i^*, f)$ .  $\diamond$

*Claim 5.* We have that  $\tau_i(s_i^*) \in D_i^+(h_i^*, f)$ .

*Proof of Claim 5.* Recall that  $D \subseteq_{\rho, f, D'} E$ . By Lemma 8.11 it then follows that  $H^e(D, f) \subseteq H^e(E, f)$ . Since we have seen above that  $h_i^* \in H^e(D, f)$  we conclude that  $h_i^* \in H^e(E, f)$ . Recall also that  $E^{rf} \subseteq_{\tau, f, E} D$ . As  $h_i^* \in H^e(E, f)$  and, by Claim 4,  $s_i^* \in E_i^{rf+}(h_i^*, f)$ , it follows by Lemma 8.10 (b) applied to  $\tau$  that  $\tau_i(s_i^*) \in D_i^+(h_i^*, f)$ .  $\diamond$

*Claim 6.* It holds that  $u_i(s_i^*, b_i^E) = u_i(\tau_i(s_i^*), b_i^D)$ .

*Proof of Claim 6.* By definition,

$$\begin{aligned} u_i(s_i^*, b_i^E) &= \sum_{s_{-i} \in S_{-i}} b_i^E(s_{-i}) \cdot u_i(z(s_i^*, s_{-i})) \\ &= \sum_{s_{-i} \in S_{-i}} \sum_{s'_{-i} \in S_{-i} : \rho_{-i}(s'_{-i}) = s_{-i}} b_i^D(s'_{-i}) \cdot u_i(z(s_i^*, s_{-i})), \end{aligned}$$

where the second equality follows from the definition of  $b_i^E$  in (8.10). Take some  $s_{-i}, s'_{-i} \in S_{-i}$  with  $\rho_{-i}(s'_{-i}) = s_{-i}$  and  $b_i^D(s'_{-i}) > 0$ . Since  $b_i^D \in \Delta(D_{-i}^+(h_i^*, f))$  we know that  $s'_{-i} \in D_{-i}^+(h_i^*, f) = \times_{j \neq i} D_j^+(h_i^*, f)$ . Moreover, we have seen in Claim 5 that  $\tau_i(s_i^*) \in D_i^+(h_i^*, f)$ . Hence,  $(\tau_i(s_i^*), s'_{-i}) \in \times_{j \in I} D_j^+(h_i^*, f)$ . As  $h_i^* \in H^e(D, f)$  it follows from Lemma 8.8 (a) that  $H(\times_{j \in I} D_j^+(h_i^*, f)) \subseteq H^e(D, f)$ , and thus

$$(\tau_i(s_i^*), s'_{-i}) \text{ only reaches information sets in } H^e(D, f).$$

The rest of the proof is very similar to the proof of Claim 3, and is therefore left to the reader.  $\diamond$

*Claim 7.* The strategy  $\sigma_i(s_i)$  is optimal in  $S_i(h_i^*)$  for  $b_i^E$ .

*Proof of Claim 7.* By Claim 4, there is a strategy  $s_i^* \in E_i^{rf+}(h_i^*, f)$  that is optimal in  $S_i(h_i^*)$  for  $b_i^E$ . Now suppose, contrary to what we want to show, that  $\sigma_i(s_i)$  is not optimal in  $S_i(h_i^*)$  for  $b_i^E$ . Then, we conclude that  $u_i(\sigma_i(s_i), b_i^E) < u_i(s_i^*, b_i^E)$ . If we combine this with Claims 3 and 6 we obtain

$$u_i(s_i, b_i^D) = u_i(\sigma_i(s_i), b_i^E) < u_i(s_i^*, b_i^E) = u_i(\tau_i(s_i^*), b_i^D).$$

However, this contradicts Claim 1 which states that  $s_i$  is optimal in  $S_i(h_i^*)$  for the belief  $b_i^D$ . Hence, we must conclude that  $\sigma_i(s_i)$  is optimal in  $S_i(h_i^*)$  for  $b_i^E$ .  $\diamond$

As we know from Claim 2 that  $b_i^E \in \Delta(E_{-i}^{rf}(h_i^*))$ , it follows from Claim 7 and Lemma 2.1 that  $\sigma_i(s_i)$  is not strictly dominated in  $(S_i(h_i^*), E_{-i}^{rf}(h_i^*))$ , and thus  $\sigma_i(s_i) \in E_i^{rf*}(h_i^*) =$

$E_i^{rf}(h_i^*)$ . This, however, contradicts our assumption that  $\sigma_i(s_i) \notin E_i^{rf}(h_i^*)$ . Hence, we must conclude that (T2) holds. As such, (T1), (T2) and (T3) hold, which implies that  $D^{rf} \subseteq_{\sigma,f,D} E^{rf}$ . This completes the proof.  $\blacksquare$

**Lemma 8.14 (Sandwich lemma)** *Consider two focus functions  $f, g$  with  $g \subseteq f$  that satisfy the conditions in Theorem 5.1 and assume, in addition, that  $f$  and  $g$  are past and future oriented. For a given  $m \geq 0$  let  $(E^k)_{k=0}^\infty$  be the  $f$ -rationalizability procedure that starts at  $rg^m(D^{full})$ , and  $(D^k)_{k=0}^\infty$  the  $f$ -rationalizability procedure that starts at  $rg^{m+1}(D^{full})$ . Then,*

$$E^1 \subseteq_{\tau,f,E^0} D^0 \subseteq_{\rho,f,D^{full}} E^0$$

for some  $\tau$  and  $\rho$ , and for every  $k \geq 1$  we have that

$$E^{k+1} \subseteq_{\tilde{\sigma},f,E^k} D^k \subseteq_{\sigma,f,D^{k-1}} E^k$$

for some  $\sigma$  and  $\tilde{\sigma}$ .

**Proof.** (a) We start by showing that  $D^0 \subseteq_{\rho,f,D^{full}} E^0$  for some  $\rho$ . Note that  $D^0 = rg(E^0) \subseteq E^0$ . We can thus use, for every player  $i$ , the identity mapping  $\rho_i : S_i \rightarrow S_i$  with  $\rho_i(s_i) := s_i$  for every strategy  $s_i$ . Then, for every strategy  $s_i \in S_i$  we have that properties (T1), (T2) and (T3) hold. In fact, (T3) is trivially satisfied since  $H^e(D^{full}, f) = H$ , and hence  $H_i(\rho_i(s_i)) \setminus H^e(D^{full}, f)$  is always empty. Thus,  $D^0 \subseteq_{\rho,f,D^{full}} E^0$ , which was to show.

(b) We next show that  $E^1 \subseteq_{\tau,f,E^0} D^0$  for some  $\tau$ . Recall that  $E^1 = rf(E^0)$  and  $D^0 = rg(E^0)$ . Set  $E := E^0$ ,  $E^{rf} := rf(E)$  and  $E^{rg} := rg(E)$ . Hence, we must show that  $E^{rf} \subseteq_{\tau,f,E} E^{rg}$ .

By Lemma 8.6 we know that  $E^{rg}$  is dynamically consistent. For every player  $i$  we construct the mapping  $\tau_i : S_i \rightarrow S_i$  as follows. Since  $E^{rg}$  is dynamically consistent we know by Lemma 8.5 (b) that there is for every information set  $h_i \in (H_i \setminus H^e(E, f))^{first}$  a strategy  $s_i[h_i]$  such that  $s_i[h_i] \in E_i^{rg*}(h'_i) = E_i^{rg}(h'_i)$  for all  $h'_i \in H_i(s_i[h_i]) \setminus H_i^-(h_i)$ . Hence, in particular,  $s_i[h_i] \in E_i^{rg}(h'_i)$  for all  $h'_i \in H_i(s_i[h_i]) \cap H_i^+(h_i)$ .

For every strategy  $s_i \in S_i$  we define  $\tau_i(s_i)$  to be the unique strategy where

$$\tau_i(s_i)|_{H^e(E,f)} = s_i|_{H^e(E,f)} \text{ and} \tag{8.16}$$

$$\tau_i(s_i)|_{H_i^+(h_i)} = s_i[h_i]|_{H_i^+(h_i)} \text{ for all } h_i \in (H_i \setminus H^e(E, f))^{first}. \tag{8.17}$$

In view of (8.16) and (8.17) we conclude that  $\tau$  satisfies properties (T1) and (T3). We will now prove (T2).

**(T2)** Take some strategy  $s_i$  and some information set  $h_i^* \in H^e(E, f) \cap H_i$  with  $s_i \in E_i^{rf-}(h_i^*, f)$ . We must show that  $\tau_i(s_i) \in E_i^{rg}(h_i^*)$ .

*Claim 1.* There is a belief  $b_i \in \Delta(E_{-i}^{rf}(h_i^*))$  such that  $s_i$  is optimal for  $b_i$  in  $S_i(h_i^*)$  and  $H(s_i, b_i) \subseteq H^e(E, f)$ .

*Proof of Claim 1.* As  $h_i^* \in H^e(E, f) \cap H_i$  and  $s_i \in E_i^{rf-}(h_i^*, f)$ , we conclude by Lemma 8.8 (c) that there is some  $b_i \in \Delta(E_{-i}^{rf}(h_i^*))$  such that  $s_i$  is optimal for  $b_i$  in  $S_i(h_i^*)$  and  $H(s_i, b_i) \subseteq H^e(E, f)$ .  $\diamond$

*Claim 2.* We have that  $b_i \in \Delta(E_{-i}^{rg}(h_i^*))$ .

*Proof of Claim 2.* Recall that  $b_i \in \Delta(E_{-i}^{rf}(h_i^*))$ . To show this claim we prove that  $E_{-i}^{rf}(h_i^*) \subseteq E_{-i}^{rg}(h_i^*)$ . Take some  $(s_j)_{j \neq i} \in E_{-i}^{rf}(h_i^*)$ . As we have seen above that  $h_i^* \in H^e(E, f)$  we know that  $E_{-i}^+(h_i^*, f) = \times_{j \neq i} E_j^+(h_i^*, f) \neq \emptyset$ , which means that  $E_{-i}^{rf}(h_i^*) = E_{-i}^+(h_i^*, f)$ . Thus,  $(s_j)_{j \neq i} \in E_{-i}^+(h_i^*, f)$ , which means that  $(s_j)_{j \neq i} \in E_{-i}(h_i^*)$ , and for every  $j \neq i$  it holds that  $s_j \in E_j(h_j)$  for all  $h_j \in f_{ij}(h_i^*) \cap H_j(s_j)$ . Since  $g \subseteq f$  it follows, for every  $j \neq i$ , that  $s_j \in E_j(h_j)$  for all  $h_j \in g_{ij}(h_i^*) \cap H_j(s_j)$ . But then,  $(s_j)_{j \neq i} \in E_{-i}^+(h_i^*, g) = E_{-i}^{rg}(h_i^*)$ . Hence, it follows that  $E_{-i}^{rf}(h_i^*) \subseteq E_{-i}^{rg}(h_i^*)$ . Since  $b_i \in \Delta(E_{-i}^{rf}(h_i^*))$  we conclude that  $b_i \in \Delta(E_{-i}^{rg}(h_i^*))$ .  $\diamond$

*Claim 3.* It holds that  $u_i(\tau_i(s_i), b_i) = u_i(s_i, b_i)$ .

*Proof of Claim 3.* This proof is very similar to the proof of Claim 3 in the proof of Lemma 8.13, and is therefore left to the reader.  $\diamond$

*Claim 4.* We have that  $\tau_i(s_i)$  is optimal in  $S_i(h_i^*)$  for  $b_i$ .

*Proof of Claim 4.* As, by Claim 1,  $s_i$  is optimal in  $S_i(h_i^*)$  for  $b_i$ , it follows by Claim 3 that also  $\tau_i(s_i)$  is optimal in  $S_i(h_i^*)$  for  $b_i$ .  $\diamond$

Since we know from Claim 2 that  $b_i \in \Delta(E_{-i}^{rg}(h_i^*))$ , it follows from Claim 4 and Lemma 2.1 that  $\tau_i(s_i)$  is not strictly dominated in  $(S_i(h_i^*), E_{-i}^{rg}(h_i^*))$ . Hence,  $\tau_i(s_i) \in E_i^{rg*}(h_i^*) = E_i^{rg}(h_i^*)$ . Since this holds for every player  $i$ , every strategy  $s_i$ , and every information set  $h_i^* \in H^e(E, f) \cap H_i$  with  $s_i \in E_i^{rf-}(h_i^*, f)$ , we see that (T2) is satisfied. By the properties (T1), (T2) and (T3) above we thus conclude that  $E^{rf} \subseteq_{\tau, f, E} E^{rg}$ , and hence  $E^1 \subseteq_{\tau, f, E^0} D^0$ , which was to show.

(c) By using (a) and (b), and a repeated application of Lemma 8.13, it can be shown by induction on  $k$  that (k.1)  $D^k \subseteq_{\sigma, f, D^{k-1}} E^k$  and (k.2)  $E^{k+1} \subseteq_{\tilde{\sigma}, f, E^k} D^k$  hold for every  $k \geq 1$ . We leave this to the reader. This completes the proof. ■

## 8.4 Proof of Theorem 5.1

Let  $f, g$  be two focus functions with  $g \subseteq f$  such that that  $f$  is forward decreasing and preserving focus on past information sets, and  $g$  is individually forward decreasing and individually preserving focus on past information sets. Define the past and future oriented focus mappings  $\hat{f}, \hat{g}$  by  $\hat{f}_{ij}(h) := f_{ij}(h) \cap H(h)$  and  $\hat{g}_{ij}(h) := g_{ij}(h) \cap H(h)$  for all information sets  $h \in H$  and all players  $i \neq j$ .

By Lemma 8.7 we know that the  $(g, f)$ - and the  $(\hat{g}, \hat{f})$ -rationalizable strategies are the same, the  $f$ - and  $\hat{f}$ -rationalizable strategies are the same, and similarly for  $g$  and  $\hat{g}$ . Moreover, it may be verified that  $\hat{g} \subseteq \hat{f}$ , the focus mapping  $\hat{f}$  is forward decreasing and preserving focus on past information sets, and  $\hat{g}$  is individually forward decreasing and individually preserving focus on past information sets. As Theorem 5.1 is about rationalizable strategies and outcomes only we may assume, without loss of generality, that  $f$  and  $g$  are past and future oriented.

(a) Let  $M$  be such that  $rg^M(D^{full}) = rg^{M+1}(D^{full})$ . For every  $m \in \{0, \dots, M\}$ , let  $(D^{m,k})_{k=0}^\infty$  be the  $f$ -rationalizability procedure that starts at  $rg^m(D^{full})$ . Let  $K$  be such that  $D^{m,K} = D^{m,K+1}$  for all  $m \in \{0, \dots, M\}$ . Then, by construction,  $D^{m,K} = (f^\infty \circ g^m)(D^{full})$  for every  $m \in \{0, \dots, M\}$  and  $D^{M,K} = (f^\infty \circ g^\infty)(D^{full})$ .

By Lemma 8.14 we know, for every  $m \in \{0, \dots, M-1\}$ , that  $D^{m,K+1} \subseteq_{\sigma^m, f, D^{m,K}} D^{m+1,K}$  and  $D^{m+1,K+1} \subseteq_{\tau^m, f, D^{m+1,K}} D^{m,K+1}$  for some  $\sigma^m, \tau^m$ . As  $D^{m,K+1} = D^{m,K}$  and  $D^{m+1,K+1} = D^{m+1,K}$  it follows that

$$D^{m,K} \subseteq_{\sigma^m, f, D^{m,K}} D^{m+1,K} \subseteq_{\tau^m, f, D^{m+1,K}} D^{m,K} \quad (8.18)$$

for some  $\sigma^m, \tau^m$ . By Lemma 8.11 we then know that  $H^e(D^{m,K}) \subseteq H^e(D^{m+1,K}) \subseteq H^e(D^{m,K})$ . That is,  $H^e(D^{m,K}) = H^e(D^{m+1,K})$  for all  $m \in \{0, 1, \dots, M-1\}$  and all players  $i$ . By (8.18) and a repeated application of Lemma 8.12 we then conclude that

$$D^{0,K} \subseteq_{(\sigma^{M-1} \circ \dots \circ \sigma^0), f, D^{0,K}} D^{M,K} \text{ and } D^{M,K} \subseteq_{(\tau^0 \circ \dots \circ \tau^{M-1}), f, D^{0,K}} D^{0,K}. \quad (8.19)$$

Moreover, we have seen in the proof of Lemma 8.13 that we can construct these transformation mappings  $\sigma^0, \dots, \sigma^{M-1}$  and  $\tau^0, \dots, \tau^{M-1}$  on the basis of (8.8) and (8.9). Suppose that the transformation mappings above have this specific construction.

Set  $\sigma^* := \sigma^{M-1} \circ \dots \circ \sigma^0$  and  $\tau^* := \tau^0 \circ \dots \circ \tau^{M-1}$ . We will show that  $\sigma^*$  transforms every combination  $(s_i)_{i \in I}$  of  $f$ -rationalizable strategies into a combination  $(\sigma_i^*(s_i))_{i \in I}$  of  $(g, f)$ -rationalizable strategies yielding the same outcome. Let  $\sigma^* = (\sigma_i^*)_{i \in I}$ . Then, by construction of the transformation mappings  $\sigma^0, \dots, \sigma^{M-1}$ , for every player  $i$  the mapping  $\sigma_i^* : S_i \rightarrow S_i$  transforms every strategy  $s_i$  into the unique strategy  $\sigma_i^*(s_i)$  that satisfies

$$\sigma_i^*(s_i)|_{H^e(D^{0,K}, f)} = s_i|_{H^e(D^{0,K}, f)} \text{ and} \quad (8.20)$$

$$\sigma_i^*(s_i)|_{H_i^+(h_i)} = s_i^{M,K}[h_i]|_{H_i^+(h_i)} \text{ for all } h_i \in (H_i \setminus H^e(D^{0,K}, f))^{first}, \quad (8.21)$$

where  $s_i^{M,K}[h_i] \in D_i^{M,K}(h_i')$  for all  $h_i' \in H_i^+(h_i) \cap H_i(s_i^{M,K}[h_i])$ .

Now, take for every player  $i$  an  $f$ -rationalizable strategy  $s_i$ . As  $rf^\infty(D^{full}) = D^{0,K}$  we must have, for every player  $i$ , that  $s_i \in D_i^{0,K}(h_i)$  for every  $h_i \in H_i(s_i)$ . Then, every information set that is reached by  $(s_i)_{i \in I}$  is in  $H^e(D^{0,K}, f)$ . To see this, consider an information set  $h^* \in H$  that is reached by  $(s_i)_{i \in I}$ . As for every player  $i$  we know that  $s_i \in D_i^{0,K}(h_i)$  for every  $h_i \in H_i(s_i)$ , we see that  $s_i \in D_i^{0,K^+}(h^*)$ , and thus  $D_i^{0,K^+}(h^*, f) \neq \emptyset$  for every player  $i$ . As such,  $h^* \in H^e(D^{0,K}, f)$ . Hence, every information set that is reached by  $(s_i)_{i \in I}$  is in  $H^e(D^{0,K}, f)$ . By (8.20) it then follows that  $z((s_i)_{i \in I}) = z((\sigma_i^*(s_i))_{i \in I})$ .

We next show that  $\sigma_i^*(s_i)$  is  $(g, f)$ -rationalizable for every player  $i$ . Take some player  $i$ . As  $s_i$  is  $f$ -rationalizable and  $rf^\infty(D^{full}) = D^{0,K}$ , we conclude that  $s_i \in D_i^{0,K}(h_i)$  for every  $h_i \in H_i(s_i)$ . This implies that  $s_i \in D_i^{0,K^-}(h_i, f)$  for all  $h_i \in H_i(s_i)$ . Since, by (8.19),  $D^{0,K} \subseteq_{\sigma^*, f, D^{0,K}} D^{M,K}$ , we know that  $\sigma^*$  satisfies conditions (T2) and (T3). As  $s_i \in D_i^{0,K^-}(h_i, f)$  for all  $h_i \in H_i(s_i)$  it follows by (T2) and (T3) that  $\sigma_i^*(s_i) \in D_i^{M,K}(h_i)$  for all  $h_i \in H_i(\sigma_i^*(s_i))$ . Since  $D^{M,K} = (f^\infty \circ g^\infty)(D^{full})$  we conclude that  $\sigma_i^*(s_i)$  is  $(g, f)$ -rationalizable.

We thus see that  $\sigma^*$  transforms every combination of  $f$ -rationalizable strategies  $(s_i)_{i \in I}$  into a combination of  $(g, f)$ -rationalizable strategies  $(\sigma_i^*(s_i))_{i \in I}$  such that  $z((s_i)_{i \in I}) = z((\sigma_i^*(s_i))_{i \in I})$ . That is, both combinations of strategies yield the same outcome. In a similar fashion it can be shown that  $\tau^*$  transforms every combination of  $(g, f)$ -rationalizable strategies  $(s_i)_{i \in I}$  into a combination of  $f$ -rationalizable strategies  $(\tau_i^*(s_i))_{i \in I}$  yielding the same outcome. As such, the set of  $f$ -rationalizable outcomes is the same as the set of  $(g, f)$ -rationalizable outcomes.

(b) Take an  $f$ -rationalizable outcome  $z$ . Then, by (a), the outcome  $z$  is also  $(g, f)$ -rationalizable and hence, in particular,  $g$ -rationalizable.

(c) Suppose that  $f$  is future oriented. Then,  $f_{ij}(h) \subseteq H_j^+(h)$  for every information set  $h \in H$  and players  $i \neq j$ . By relying on Lemma 8.4 (a) it can then be proven that  $H^e(D^{0,K}, f) = H$ . But then, it follows from the construction of  $\sigma_i^*$  in (8.20) and (8.21) that  $\sigma_i^*$  is the identity mapping from  $S_i$  to  $S_i$ . Now, take some  $f$ -rationalizable strategy  $s_i$  for player  $i$ . By our proof of (a) we know that  $s_i = \sigma_i^*(s_i)$  is  $(g, f)$ -rationalizable and hence, in particular,  $s_i$  is  $g$ -rationalizable. That is, every  $f$ -rationalizable strategy is  $g$ -rationalizable. This completes the proof. ■

## References

- [1] Battigalli, Pierpaolo (1997), On rationalizability in extensive games, *Journal of Economic Theory* **74**, 40–61.
- [2] Battigalli, Pierpaolo (2003), Rationalizability in infinite, dynamic games of incomplete information, *Research in Economics* **57**, 1–38.
- [3] Battigalli, Pierpaolo and Emiliano Catonini (2026), Monotonicity and robust implementation under forward-induction reasoning, *Econometrica* **94**, 505–536.
- [4] Battigalli, Pierpaolo, Catonini, Emiliano and Julien Manili (2023), Belief change, rationality, and strategic reasoning in sequential games, *Games and Economic Behavior* **142**, 527–551.
- [5] Ben-Porath, Elchanan (1997), Rationality, Nash equilibrium and backwards induction in perfect-information games, *Review of Economic Studies* **64**, 23–46.
- [6] Catonini, Emiliano (2020), On non-monotonic strategic reasoning, *Games and Economic Behavior* **120**, 209–224.
- [7] Catonini, Emiliano and Antonio Penta (2026), Backward induction reasoning beyond backward induction, *American Economic Journal: Microeconomics* **18**, 30–58.
- [8] Chen, Jing and Silvio Micali (2013), The order independence of iterated dominance in extensive games, *Theoretical Economics* **8**, 125–163.

- [9] Heifetz, Aviad and Andrés Perea (2015), On the outcome equivalence of backward induction and extensive form rationalizability, *International Journal of Game Theory* **44**, 37–59.
- [10] Meier, Martin and Andrés Perea (2023), Forward induction in a backward inductive manner, *Epicenter Working Paper No. 27*.
- [11] Pearce, David G. (1984), Rationalizable strategic behavior and the problem of perfection, *Econometrica* **52**, 1029–1050.
- [12] Penta, Antonio (2015), Robust dynamic implementation, *Journal of Economic Theory* **160**, 280–316.
- [13] Perea, Andrés (2012), *Epistemic Game Theory: Reasoning and Choice*, Cambridge University Press.
- [14] Perea, Andrés (2014), Belief in the opponents’ future rationality, *Games and Economic Behavior* **83**, 231–254.
- [15] Perea, Andrés (2017), Order independence in dynamic games, *Epicenter Working Paper No. 8*.
- [16] Perea, Andrés (2018), Why forward induction leads to the backward induction outcome: A new proof for Battigalli’s theorem, *Games and Economic Behavior* **110**, 120–138.
- [17] Perea, Andrés (2024), More reasoning, less outcomes: A monotonicity result for reasoning in dynamic games, *Epicenter Working Paper No. 32*, <https://www.epicenter.name/wp-content/uploads/2024/12/Monotonicity-WP.pdf>
- [18] Perea, Andrés and Elias Tsakas (2019), Limited focus in dynamic games, *International Journal of Game Theory* **48**, 571–607.
- [19] Reny, Philip (1992), Backward induction, normal form perfection and explicable equilibria, *Econometrica* **60**, 627–649.
- [20] Rubinstein, Ariel (1991), Comments on the interpretation of game theory, *Econometrica* **59**, 909–924.

- [21] Shimoji, Makoto and Joel Watson (1998), Conditional dominance, rationalizability, and game forms, *Journal of Economic Theory* **83**, 161–195.