Pure Backward Induction Reasoning in Dynamic Games*

EPICENTER Working Paper No. 34 (2025)

Andrés Perea†
Maastricht University

This version: November 2025

Abstract

Properties BI1 and BI2 in Kohlberg and Mertens (1986) entail that the solution of a dynamic game, when restricted to a subgame, should yield the same as applying the solution to the subgame in isolation. Remarkably, none of the existing equilibrium and rationalizability concepts satisfies this pure backward induction principle. This paper offers a action-based rationalizability concept that does. It is based on the epistemic condition that a player, at every history, believes that his opponents', and he himself, will choose optimal actions in the future. Iterating this condition leads to *common belief in future action-based rationality.* Its behavioral consequences are characterized by a computationally convenient elimination procedure, called the *double-utility* procedure. In perfect information games it yields backward induction, whereas it simplifies to a very easy procedure in finitely repeated games.

JEL Classification: C72

Keywords: Backward induction, rationalizability, dynamic games, pure backward induction principle, double-utility games, double-utility procedure

*I thank the audience at the MLSE seminar at Maastricht University for useful comments.

†*E-mail:* a.perea@maastrichtuniversity.nl *Affiliation:* Maastricht University, Department of Quantitative Economics and EpiCenter, The Netherlands. *Web*: www.epicenter.name/Perea/

1 Introduction

In a dynamic game, *backward induction* represents a line of reasoning in which a player, at each of his information sets, bases his belief and behavior solely on the game that lies ahead. When viewed in its purest form, it allows a player to forget about the events that took place in the past, including the players' past behavior, his own past beliefs, and the utilities at terminal histories that he himself avoided. This is essentially the message of the backward induction properties BI1 and BI2 in Kohlberg and Mertens (1986), which together state that the solution of a game, when restricted to a subgame, should coincide with the solution applied to the subgame in isolation.[1]

In particular, a "pure" backward induction concept (a) must not restrict the beliefs of a player at two consecutive information sets by forward consistency[2], that is, by the rules of conditional probability, and (b) must assume that a player cannot commit to future actions. To see why, suppose that a player moves at two information sets where the second follows the first. If we would tie the beliefs held at these information sets by forward consistency, then upon reaching the second information set the player must base his belief on the past belief he held at the first information set, which goes against pure backward induction reasoning. Suppose next that the player considers action a optimal at the first information set because of his belief about the opponents' behavior, and because he is able to commit to action b at the second information set, although another action c is equally reasonable there once the second information set is reached. To guarantee that he will actually play action b at the second information set, he must remember all these considerations at the second information set – something that is precluded by pure backward induction reasoning.

To the best of our knowledge, every rationalizability or equilibrium concept in the existing literature – except for the backward induction procedure which can only be applied to games with perfect information – violates either (a) or (b). Indeed, the concepts of subgame perfect equilibrium (Selten (1965)), sequential equilibrium (Kreps and Wilson (1982)), backwards rationalizability (Penta (2015), Perea (2014), Catonini and Penta (2025)) and sequential rationalizability (Asheim and Perea (2005), Dekel, Fudenberg and Levine (1999, 2002)), all tie the beliefs of a player at the different information sets by forward consistency – either explicitly or implicitly.[3] Moreover, backward dominance (Perea (2014)) and backwards rationalizability assume that players execute *strategies* that are optimal, at each of their information sets, given their condi-

[1] Meier and Perea (2023) introduce the property of *supergame monotonicity* (also called *expansion monotonicity* in later versions) which closely corresponds to BI1 in Kohlberg and Mertens (1986). It states that if a player discovers that the game is actually preceded by earlier moves, then the solution of the "supergame", when restricted to the smaller game, should be a weak refinement of the solution applied to the smaller game in isolation.

[2] This terminology is adopted from Battigalli, Catonini and Manili (2023). In the literature, this condition is often called *Bayesian updating.*

[3] In this paper we only consider dynamic games with observable actions, that is, there may be simultaneous moves but the players always know which actions have been chosen in the past. For such games, subgame perfect and sequential equilibria are equivalent, and they prescribe at every information set a probability distribution over the available actions for every active player. As these probability distributions may be viewed as the common, fixed belief that the players hold about the future actions of their opponents, the two concepts implicitly impose the rules of conditional probability on the players' beliefs. Indeed, if the game moves from a first to a second information set of player, then this player is assumed to preserve his belief about the future actions following the second information set.

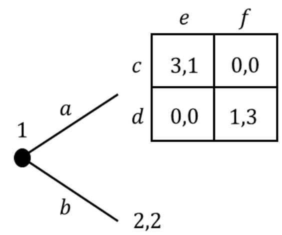

Figure 1: Battle of the sexes with an outside option

tional beliefs about the opponents' strategies. As such, these concepts presuppose that players can commit, at each of their information sets, to the future actions involved in these optimal strategies. Battigalli and de Vito (2021) provide epistemic conditions, different from those in Perea (2014), that also lead to backwards rationalizability. Their model imposes forward consistency but assumes that the players, at each of their information sets, can only control the *current* action at this particular information set, and cannot commit to future actions.[4]

As such, none of the concepts above classifies as a pure backward induction concept. This can be illustrated by the well-known "battle of the sexes game with an outside option" in Figure 1.

It can be verified that all of the concepts above select player 1's strategies (a, c) and b (but not (a, d)) in the whole game, but select player 1's strategies c and d when applied to the subgame following a in isolation. As such, player 1's induced behavior in the whole game, when restricted to the subgame, is that player 1 will choose c there, whereas player 1's induced behavior in the subgame, when this subgame is viewed in isolation, is that player 1 could choose c and d. This violates the backward induction properties BI1 and BI2 in Kohlberg and Mertens (1986).[5]

To see why the concepts rule out strategy (a, d) for player 1 in the full game, recall that each of these

[4]For the concepts of subgame perfect equilibrium, sequential equilibrium and sequential rationalizability, it remains a bit ambiguous whether players can commit to future actions or not. These concepts are all based on (induced) *behavioral strategies,* which assign to every information set of a player a probability distribution over his available actions there. If a player is at a given information set, then his behavioral strategy can either be interpreted as (i) describing his continuation strategy from this information set onwards, in which he can commit to future actions, or (ii) the belief this player has about his own future actions, assuming he cannot commit to these.

[5]Kohlberg and Mertens (1986) argue that sequential equilibrium, and therefore also subgame perfect equilibrium, satisfy the properties BI1 and BI2. I do not agree however: In the game of Figure 1, both concepts select the behavioral strategy profiles (a, c, e), (b, d, f), and $(b, 3/4c + 1/4d, 1/4e + 3/4f)$ when applied to the whole game, and select (c, e), (d, f), and $(3/4c + 1/4d, 1/4e + 3/4f)$ when applied to the subgame in isolation. But then, in my view, player 1's induced behavior in the whole game, when restricted to the subgame, is that player 1 must choose c there, as only the sequential equilibrium (a, c, e) assigns a positive probability to player 1 choosing a. This is different from player 1's induced behavior in the subgame, when viewed in isolation, which allows player 1 to choose c or d. Apparently, Kohlberg and Mertens (1986) hold the view that the sequential equilibria of the whole game, when restricted to the subgame following a, are (c, e), (d, f), and $(3/4c + 1/4d, 1/4e + 3/4f)$, even if the latter two equilibria assign probability zero to the action a.

concepts either (i) imposes forward consistency to tie the beliefs at different information sets, or (ii) assumes that a player can commit to future actions. If (i) holds, and player 1 finds it optimal to choose action a at the beginning, then he must believe with probability at least $2/3$ that player 2 chooses e. By (i), he should still believe so at the second information set, which renders only his action c optimal there. Suppose next that (ii) holds. Then, player 1 will never commit to strategy (a, d) because playing b will always be better at the beginning of the game.

The goal of this paper is to introduce a new rationalizability concept for dynamic games, the *double-utility procedure,* that classifies as a pure backward induction concept, and coincides with the backward induction procedure when applied to a game with perfect information. Like Battigalli and de Vito (2021), we assume that a player can only control the current action at each of his information sets, and cannot commit to future actions. To accomodate for this, a player holds, at each of his information sets, a belief about the opponents' current and future actions, and a belief about his *own* future actions.

We proceed epistemically, by imposing conditions on the players' *belief hierarchies* that formalize a specific form of backward induction reasoning. More precisely, we require that every player believes, at each of his information sets, that (i) his opponents' choose optimal actions now and in the future, and (ii) that he himself will choose optimal actions in the future. Common belief in (i) and (ii) yields the conditions of *common belief in future action-based rationality.*

These conditions are closely related to the conditions of (common full belief in) *optimal planning* and *belief in continuation consistency* in Battigalli and de Vito (2021) which lead to the backwards rationalizability procedure, and the conditions of *common belief in future rationality* in Perea (2014) which underly the backward dominance procedure. The main difference with Battigalli and de Vito (2021) is that we do not impose forward consistency. Perea (2014), on the other hand, assumes that a player *can* commit to future actions, and that he only holds a belief about the opponents' actions, and not about his own future actions, at each of his information sets. The key condition in *common belief in future rationality* is that a player always believes that his opponents will choose optimal (continuation) *strategies* in the future.

Different from any of the other backward induction concepts above, the conditions in common belief in future action-based rationality allow player 1 to play the strategy (a, d) in the game of Figure 1. The reason is that player 1 can rationally choose the action a because he believes at the beginning of the game that player 2 will choose e and that he himself will choose c in the future. Once the subgame is reached, player 1 can rationally choose the action d because he now believes that player 2 will choose f, thus violating forward consistency. These beliefs can be embedded in a belief hierarchy for player 1 that expresses common belief in future action-based rationality. Important is that with the reasoning above, the strategy (a, d) should be interpreted as a pair of "isolated" actions, rather than a plan of "coordinated" actions. In a similar vein, it can be shown that all strategies for players 1 and 2 are compatible with common belief in future action-based rationality.

In our main result, Theorem 4.1, we show that the behavioral consequences of common belief in action-based rationality can be characterized by a recursive elimination procedure which we call the *double-utility procedure.* It proceeds by iteratively eliminating *actions,* rather than strategies, at the various non-terminal histories in the game, as follows.

We start by specifying, at every non-terminal history, and for every profile of actions there, the highest and lowest utility that every player can possibly achieve in the remainder of the game if this action profile is chosen. This gives rise to a so-called *double-utility game* at this history. In the first round we then eliminate, at every non-terminal history, the actions that are *strictly dominated* in this double utility game, meaning that for every opponents' action profile, the highest utility that can be achieved through this action is still lower than the lowest expected utility that can be achieved by some randomized action.

In the second round we start by updating the double-utility games at the various non-terminal histories: For every action profile, the highest and lowest utility for a player are now defined relative to the future actions (by himself and the opponents) that have survived the first round. We subsequently eliminate all actions that are strictly dominated within the updated double-utility games. We continue in this fashion until no further actions can be eliminated in the game.

Theorem 4.1 shows that the actions that can rationally be played under the conditions of common belief in future action-based rationality are precisely the actions that survive the double-utility procedure. By construction, the procedure satisfies the backward induction properties BI1 and BI2 in Kohlberg and Mertens (1986), and hence it may be viewed as a pure backward induction concept. Moreover, the procedure is computationally convenient, as the specification of the double-utility games at the various non-terminal histories only requires us to list the sets of actions, and not the sets of strategies, for each of the players.

We next investigate how the concept behaves in some prominent special classes of games – perfect information games and finitely repeated games. It turns out that for every game with perfect information and without relevant ties[6], the double-utility procedure coincides exactly with the backward induction procedure. As such, the concept constitutes a generalization of backward induction to games with imperfect information.

For every finitely repeated game, in which a given stage game is repeated during finitely many stages, the procedure amounts to the following very simple elimination process: At the last stage we iteratively eliminate strictly dominated actions from the stage game. At the second-to-last stage we compute, for every player i, the difference α_i between the highest and the lowest utility he can obtain in the reduced game at the last stage, and iteratively eliminate, for every player i, actions from the stage game that are strictly dominated by an amount higher than α_i. At the third-to-last stage we compute, for every player i, the difference α_i' between the highest and the lowest utility he can obtain in the reduced game at the second-to-last stage, and iteratively eliminate, for every player i, actions from the stage game that are strictly dominated by an amount higher than $\alpha_i' + \alpha_i$, and so on.

By construction of the procedure, the sets of actions that survive at a given stage become smaller as the repeated game proceeds. Moreover, this procedure can be adapted to the case where the utilities for the players are discounted.

We finally compare our concept to the closely related concept of backward dominance (Perea (2014)). It is shown that the backward dominance procedure is a refinement of the double-utility procedure. As the backwards rationalizability procedure refines the backward dominance procedure, it follows that the former also refines the double-utility procedure.

The paper is organized as follows. In Section 2 we introduce the model of a dynamic game together

[6] That is, every two different actions for a player at a non-terminal history always lead to different utilities for that player.

with related notions such as beliefs, expected utility and optimality of actions. In Section 3 we lay out the epistemic conditions of common belief in future action-based rationality. In Section 4 we define the double-utility procedure and show that it characterizes the actions that can rationally be chosen under the epistemic conditions above. In Section 5 we prove that the order and speed of elimination is irrelevant for the output of the double-utility procedure. This property is used in Section 6 to demonstrate that the procedure reduces to the backward induction procedure when applied to a game with perfect information. In Section 7 we focus on finitely repeated games, and show that the procedure simplifies to a very easy algorithm that can be used to calculate the actions that can rationally be chosen under common belief in future action-based rationality at every stage of the repeated game. In Section 8 we show that the double-utility procedure is more permissive than the backward dominance procedure. In Section 9 we provide some concluding remarks. All proofs can be found in the appendix.

2 Games and Beliefs

2.1 Dynamic Games

In this paper we consider finite dynamic games that allow for simultaneous moves, but where the players always observe which actions have been chosen in the past. Formally, a *finite dynamic game with observable actions* is a tuple $\Gamma = (I, (A_i^*, A_i, u_i)_{i \in I})$, where

(a) I is the finite set of *players*;

(b) A_i^* is the finite set of *actions* for player i. By $\bar{A}$ we denote the set of *action sequences,* containing the empty sequence $\varnothing$, and the finite sequences $\bar{a} = (a^1, ..., a^n)$, where $n \in \mathbb{N}$ and $a^m = (a_i^m)_{i \in I} \in \times_{i \in I} A_i^*$ for every $m \in \{1, ..., n\}$;

(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. By $A(\bar{a}) := \times_{i \in I} A_i(\bar{a})$ we denote the set of action profiles after $\bar{a}$. The action sequence $\bar{a} = (a^1, ..., a^n) \in \bar{A}$ is called a *history* if (i) $a^1 \in A(\varnothing)$, (ii) $a^m \in A(a^1, ..., a^{m-1})$ for every $m \in \{2, ..., n\}$, and (iii) there is at least one active player at $\varnothing$, and at every subsequence $(a^1, ..., a^m)$ with $m \in \{1, ..., n-1\}$. The history $\bar{a}$ is *non-terminal* if there is at least one active player at $\bar{a}$, and is called *terminal* 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;

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

We say that a history h *precedes* a history $h' = (a^1, ..., a^n)$ or, equivalently, h' *follows* h (written as $h' \succ h$) if $h = \varnothing$ or $h = (a^1, ..., a^m)$ for some $m \in \{1, .., n-1\}$. We say that h *weakly precedes* h' or, equivalently, h' *weakly follows* h (written as $h' \succcurlyeq h$), if either h precedes h', or $h = h'$. For a history $h \in H$ and action profile $a \in A(h)$, we denote by (h, a) the history obtained if we add a after h. By $A_{-i}(h) := \times_{j \neq i} A_j(h)$ we denote the set of opponents' action profiles at h.

2.2 Beliefs, Expected Utility, and Optimal Actions

For a player i and history $h \in H$, let

$$A^{i,\succeq h} := (\times_{h' \succ h} A_i(h')) \times (\times_{j \neq i} \times_{h' \succeq h} A_j(h'))$$

be the collection of profiles of actions that i can choose after h, and that i's opponents can choose at and after h_i. If player i chooses an action $a_i \in A_i(h)$ at h, and the players select an action profile $a \in A^{i,\succeq h}$, this leads to a unique terminal history $z(h, a_i, a)$.

For a finite set X, let $\Delta(X)$ be the set of probability distributions on X. Suppose that player i holds at h a *probabilistic belief* $b_i \in \Delta(A^{i,\succeq h})$ about his own future actions[7] and the opponents' current and future actions, and chooses the action $a_i \in A_i(h)$ there. Then, the induced *expected utility* is

$$u_i(a_i, b_i) := \sum_{a \in A^{i,\succeq h}} b_i(a) \cdot u_i(z(h, a_i, a)).$$

The action $a_i \in A_i(h)$ is *optimal* at h for the belief b_i if $u_i(a_i, b_i) \geq u_i(a_i', b_i)$ for all $a_i' \in A_i(h)$.

3 Belief in Future Action-Based Rationality

In this section we formally define the epistemic conditions of *common belief in future action-based rationality.* To do so, we first introduce belief hierarchies and show how these can be encoded by means of types within an epistemic model.

3.1 Belief Hierarchies and Types

We assume that every player holds, at every non-terminal history, a *first-order* probabilistic belief about his own future actions and the opponents' current and future actions, a *second-order* belief about his own future actions and future first-order beliefs and about his opponents' current and future actions and his opponents' current and future first-order beliefs, and so on. Such *belief hierarchies* can be encoded by means of an epistemic model with types, which will be defined next.

An *epistemic model* $M = (T_i(h), b_i)_{i \in I, h \in H}$ prescribes, for every player i, (i) a finite set of types $T_i(h)$ for every history $h \in H$, and (ii) a belief mapping b_i, which assigns to every history $h \in H_i$ and type $t_i \in T_i(h)$ a belief

$$b_i(t_i) \in \Delta(\times_{h' \succ h}(A_i(h') \times T_i(h')) \times (\times_{j \neq i} \times_{h' \succeq h} (A_j(h) \times T_j(h)))).$$

about his own future action-type pairs, and about his opponents' current and future action-type pairs. For every type we can then derive a full belief hierarchy. In particular, the first-order belief induced by type $t_i \in T_i(h)$ is given by $b_i^1(t_i) := \text{marg}_{A^{i,\succeq h}} b_i(t_i)$, which is a probabilistic belief on $A^{i,\succeq h}$. Note that a type $t_i \in T_i(h)$ is, by construction, forward looking, as it only holds beliefs about actions and beliefs at, and after, history h.

[7] In Battigalli and de Vito (2021), player i's belief about his own future behavior is called a *plan.*

From now on we will write

$$T^{i,\succcurlyeq h} := \times_{h' \succ h} T_i(h') \times (\times_{j \neq i} \times_{h' \succcurlyeq h} T_j(h))$$

for every player i and history h. As such, every type $t_i \in T_i(h)$ holds a belief $b_i(t_i) \in \Delta(A^{i,\succcurlyeq h} \times T^{i,\succcurlyeq h})$.

3.2 Common Belief in Future Action-Based Rationality

For a given player i, consider a history $h \in H$, a type $t_i \in T_i(h)$ within an epistemic model, and an action $a_i \in A_i(h)$. Then, the action a_i is *optimal* for type t_i if a_i is optimal for the induced first-order belief $b_i^1(t_i)$. On the basis of this, we can recursively define *k-fold belief in future action-based rationality* for every $k \in \mathbb{N}_0$, where $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. *Common* belief in future action-based rationality is obtained by requiring this condition for every k. In the definition, we say that a type $t_i \in T_i(h)$ assigns a positive probability to an action-type pair (a_j, t'_j) (where j could be i) if $b_i(t_i)$ assigns positive probability to an action-type profile in $A^{i,\succcurlyeq h} \times T^{i,\succcurlyeq h}$ that contains (a_j, t'_j).

Definition 3.1 (Common belief in future action-based rationality) *Within a given epistemic model $M = (T_i(h), b_i)_{i \in I, h \in H}$, every type expresses, by definition, 0-fold belief in future action-based rationality.*

For some $k \in \mathbb{N}$, we say that type $t_i \in T_i(h)$ expresses k-fold belief in future action-based rationality if the belief $b_i(t_i)$ only assigns positive probability to (i) action-type pairs (a_i, t'_i) of player i himself where a_i is optimal for t'_i and t'_i expresses $(k-1)$-fold belief in future action-based rationality, and (ii) action-type pairs (a_j, t_j) of his opponents where a_j is optimal for t_j and t_j expresses $(k-1)$-fold belief in future action-based rationality.

A type expresses common belief in future action-based rationality if it expresses k-fold belief in future action-based rationality for every $k \in \mathbb{N}_0$.

We say that t_i expresses up to k-fold belief in future action-based rationality if it expresses m-fold belief in future action-based rationality for all $m \in \{1, ..., k\}$. For a given player i and history $h^* \in H$, we say that an action $a_i \in A_i(h^*)$ can rationally be chosen under up to k-fold (common) belief in future action-based rationality if there is an epistemic model $M = (T_j(h), b_j)_{j \in I, h \in H}$, and a type $t_i \in T_i(h^*)$, such that (i) t_i expresses up to k-fold (common) belief in future action-based rationality, and (ii) the action a_i is optimal for t_i.

Our condition that a type expresses 1-fold belief in future action-based rationality is similar to the combination of the conditions of *belief in optimal planning* and *belief in continuation consistency* in Battigalli and de Vito (2021). Analogously, our notion of *common* belief in future action-based rationality is closely related to their condition of *common* full belief in optimal planning and in the belief in continuation consistency. For a more detailed comparison, see Section 9.

4 Elimination Procedure

We will now introduce an elimination procedure, called the *double-utility procedure,* and show that it characterizes precisely those actions that can rationally be chosen under common belief in future action-based rationality. As a first step towards this goal we start by defining double-utility games

4.1 Double-Utility Games

A *double-utility game* is a tuple $G = (A_i, u_i^+, u_i^-)_{i \in I}$, where I is the finite set of players, A_i is the finite set of actions for player i, and $u_i^+, u_i^- : \times_{j \in I} A_j \to \mathbb{R}$ are utility functions with $u_i^+(a) \geq u_i^-(a)$ for all action profiles $a \in \times_{j \in I} A_j$. The interpretation is that players may be uncertain about the outcome that results after the action profile $a = (a_j)_{j \in I}$ has been chosen. In this case, $u_i^+(a)$ and $u_i^-(a)$ represent the highest and lowest utility, respectively, that player i expects after the action profile a. We refer to $u_i^+(a)$ and $u_i^-(a)$ as the upper and lower utility, respectively.

Within a double-utility game $G = (A_i, u_i^+, u_i^-)_{i \in I}$, an action $a_i \in A_i$ is *strictly dominated* is there is a randomized action $r_i \in \Delta(A_i)$ such that

$$u_i^+(a_i, a_{-i}) < \sum_{a_i' \in A_i} r_i(a_i') \cdot u_i^-(a_i', a_{-i}) \text{ for all } a_{-i} \in A_{-i}.$$

That is, whatever the opponents do, the upper utility that player i obtains from choosing a_i is still lower than the lower expected utility he obtains from choosing the randomized action r_i instead.[8]

Analogously, an action $a_i \in A_i$ is *optimal* for a probabilistic belief $\beta_i \in \Delta(A_{-i})$ if

$$\sum_{a_{-i} \in A_{-i}} \beta_i(a_{-i}) \cdot u_i^+(a_i, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}} \beta_i(a_{-i}) \cdot u_i^-(a_i', a_{-i}) \text{ for all } a_i' \in A_i.$$

Thus, the upper expected utility that player i obtains from choosing a_i under this belief is at least as high as the lower expected utility he obtains from choosing any other action.

The following result is a generalization of Lemma 3 in Pearce (1984), and provides a connection between strict dominance and optimality for double-utility games.

Lemma 4.1 (Pearce's lemma for double-utility games) *Consider a double-utility game $G = (A_i, u_i^+, u_i^-)_{i \in I}$, a player i, and an action $a_i \in A_i$. Then, action a_i is strictly dominated within G, if and only if, a_i is not optimal in G for any probabilistic belief $\beta_i \in \Delta(A_{-i})$.*

[8] This is related, in spirit, to the notion of an *obviously dominant strategy* in Li (2017). In a dynamic game, a strategy s_i is obviously dominant if for every alternative strategy s_i' and history h where s_i, s_i' differ for the first time, the lowest utility that can be achieved by playing s_i from h onwards is still at least as high as the highest utility that can be achieved by playing s_i' from h onwards.

4.2 Double-Utility Procedure

We now present an elimination procedure, called the *double-utility procedure,* that starts by defining at every non-terminal history a double-utility game, and then proceeds by recursively eliminating strictly dominated actions, and by updating the utility functions, in these double-utility games. We need the following additional notation: For a given terminal history $z \in Z$, non-terminal history $h \in H$ with $z \succcurlyeq h$, and player i, we denote by $a_i(z,h)$ the projection of z on $A_i(h)$. That is, $a_i(z,h)$ is the unique action for player i at h that leads to z.

Definition 4.1 (Double-utility procedure) *(Initial step) For every $h \in H$, define the double-utility game $G^0(h) := (A_i^0(h), u_i^{0+}[h], u_i^{0-}[h])_{i \in I}$ where $A_i^0(h) := A_i(h)$, and for every $a \in \times_{i \in I} A_i^0(h)$ we have*

$$u_i^{0+}[h](a) \ : \ = \max\{u_i(z) \mid z \in Z \text{ and } z \succcurlyeq (h,a)\} \text{ and}$$
$$u_i^{0-}[h](a) \ : \ = \min\{u_i(z) \mid z \in Z \text{ and } z \succcurlyeq (h,a)\}.$$

(Recursive step) Let $k \in \mathbb{N}$ and suppose that the double-utility games $G^{k-1}(h)$ have been defined for every $h \in H$. For every $h \in H$ define the double-utility game $G^k(h) := (A_i^k(h), u_i^{k+}[h], u_i^{k-}[h])_{i \in I}$ where

$$A_i^k(h) := \{a_i \in A_i^{k-1}(h) \mid a_i \text{ not strictly dominated in } G^{k-1}(h)\},$$

and for every $a \in \times_{i \in I} A_i^k(h)$ we have

$$u_i^{k+}[h](a) := \max\{u_i(z) \mid z \in Z, \ z \succcurlyeq (h,a) \text{ and}$$
$$a_j(z,h') \in A_j^k(h') \text{ for all } j \in I \text{ and } h' \in H \text{ with } z \succcurlyeq h' \succcurlyeq (h,a)\} \text{ and}$$
$$u_i^{k-}[h](a) := \min\{u_i(z) \mid z \in Z, \ z \succcurlyeq (h,a) \text{ and}$$
$$a_j(z,h') \in A_j^k(h') \text{ for all } j \in I \text{ and } h' \in H \text{ with } z \succcurlyeq h' \succcurlyeq (h,a)\}.$$

Hence, $u_i^{k+}[h](a)$ is the maximum utility that player i can get if h is reached, the players choose the action profile a at h, and choose actions after (h,a) that have not been eliminated yet by the procedure. Similarly for $u_i^{k-}[h](a)$.

The double-utility procedure is computationally very convenient, since at every history we only need as input the set of *actions* for every player, and not the set of *strategies,* as is the case for most rationalizability procedures such as backward dominance, backwards rationalizability and strong rationalizability (also known as extensive form rationalizability, see Pearce (1984) and Battigalli (1997)). It is well-known that the sets of strategies in a dynamic game can quickly become very large, which makes the latter procedures computationally more demanding for large dynamic games. In contrast, the input size for the double-utility procedure is approximately equal to the size of the dynamic game itself.

Clearly, there will be some $K \in \mathbb{N}_0$ in the double-utility procedure such that $G^K(h) = G^{K+1}(h)$ for all $h \in H$, and the first such number K is the round where the procedure terminates. For a given player i and history $h \in H$, an action $a_i \in A_i(h)$ is said to survive round k of the double-utility procedure if $a_i \in A_i^k(h)$. Action $a_i \in A_i(h)$ is said to *survive the double-utility procedure procedure* if a_i survives every round of the procedure.

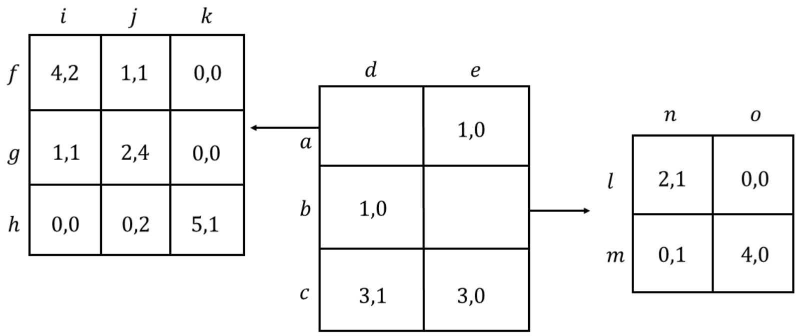

Figure 2: Illustration of action-based backward dominance procedure

Theorem 4.1 (Characterization result) *(a) For every $k \in \mathbb{N}_0$, an action can rationally be chosen under up to k-fold belief in future action-based rationality, if and only if, the action survives round $k+1$ of the double-utility procedure.*

(b) An action can rationally be chosen under common belief in future action-based rationality, if and only if, the action survives the double-utility procedure.

It may be verified that for every history h and player $i \in I$, there is at least one action $a_i \in A_i(h)$ that survives the double-utility procedure. In view of the theorem above, we can then construct a type $t_i \in T_i(h)$ that expresses common belief in future action-based rationality, and for which a_i is optimal. In particular, we see that for every player, and every history, we can always construct a type at that history that expresses common belief in future action-based rationality. The proof of this theorem shows that we can construct a *single* epistemic model that contains such a type for every player. We thus arrive at the following conclusion.

Corollary 4.1 (Existence) *There is an epistemic model $M = (T_i(h), b_i)_{i \in I, h \in H}$ such that for every player i and history $h \in H$ there is a type $t_i \in T_i(h)$ that expresses common belief in future action-based rationality.*

In other words, the epistemic conditions we impose never lead to logical contradictions.

4.3 Example

We will now illustrate the double-utility procedure by means of an example. Consider the dynamic game in Figure 2. In this game there are three non-terminal histories: $\varnothing$ (the beginning) in the middle, (a, d) on the left, and (b, e) on the right. At every non-terminal history, player 1's actions are in the rows, and player 2's actions in the columns. To start the double-utility procedure, we first list the initial double-utility games $G^0(h)$ at the three non-terminal histories h in Table 1. Here, the upper number and the lower number

(a,d)	i	j	k
f	4 2 4 , 2	1 1 1 , 1	0 0 0 , 0
g	1 1 1 , 1	2 4 2 , 4	0 0 0 , 0
h	0 0 0 , 0	0 2 0 , 2	5 1 5 , 1

$\varnothing$	d	e
a	5 4 0 , 0	1 0 1 , 0
b	1 0 1 , 0	4 1 0 , 0
c	3 1 3 , 1	3 0 3 , 0

(b,e)	n	o
l	2 1 2 , 1	0 0 0 , 0
m	0 1 0 , 1	4 0 4 , 0

Table 1: Initial double-utility games in example of Figure 2

(a,d)	i	j
f	4 2 4 , 2	1 1 1 , 1
g	1 1 1 , 1	2 4 2 , 4
h	0 0 0 , 0	0 2 0 , 2

$\varnothing$	d	e
a	**4** 4 0 , 0	1 0 1 , 0
b	1 0 1 , 0	**2** 1 0 , **1**
c	3 1 3 , 1	3 0 3 , 0

(b,e)	n
l	2 1 2 , 1
m	0 1 0 , 1

Table 2: Round 1 double-utility games in example of Figure 2

always denote the upper and lower utility for the respective player. Hence, for the action pair (a,d) at $\varnothing$, the upper and lower utility are 5 and 0 for player 1, and 4 and 0 for player 2, respectively.

Round 1. In $G^0((a,d))$, player 2's action k is strictly dominated by j and can thus be eliminated. In $G^0((b,e))$, player 2's action o is strictly dominated by n and can thus be eliminated. This leads to the double-utility games $G^1(h)$ in Table 2. Note that at $\varnothing$, player 1's upper utility after (a,d) has fallen to 4, that player 1's upper utility after (b,e) has fallen to 2, and that player 2's lower utility after (b,e) has risen to 1.

Round 2. In $G^1((a,d))$, player 1's action h is strictly dominated by g, and can thus be eliminated. In $G^1(\varnothing)$, player 1's action b is strictly dominated by c and can thus be eliminated. Moreover, in $G^1((b,e))$ player 1's action m is strictly dominated by l, and can thus be eliminated. This leads to the double-utility games $G^2(h)$ in Table 3. Note that the lower utility for player 1 and player 2 at $\varnothing$ after (a,d) has risen to 1.

Round 3. In $G^2(\varnothing)$, player 2's action e is strictly dominated by d, and can thus be eliminated. This leads to the final double-utility games $G^3(h)$ in Table 4, from which no further actions can be eliminated. We thus see that at $\varnothing$, the surviving actions are a, c and d, the surviving actions at (a,d) are f, g, i and j, and the surviving actions at (b,e) are l and n. By Theorem 4.1, these are precisely the actions that both players can

(a,d)	i	j
f	4 2 4 , 2	1 1 1 , 1
g	1 1 1 , 1	2 4 2 , 4

$\varnothing$	d	e
a	4 4 **1** , **1**	1 0 1 , 0
c	3 1 3 , 1	3 0 3 , 0

(b,e)	n
l	2 1 2 , 1

Table 3: Round 2 double-utility games in example of Figure 2

(a,d)	i	j
f	4 2 4 , 2	1 1 1 , 1
g	1 1 1 , 1	2 4 2 , 4

$\varnothing$	d
a	4 4 1 , 1
c	3 1 3 , 1

(b,e)	n
l	2 1 2 , 1

Table 4: Final double-utility games in example of Figure 2

rationally choose under common belief in future action-based rationality at the three non-terminal histories.

Consider next the battle of the sexes with an outside option, as depicted in Figure 1. It may be verified that under the double-utility procedure, no actions can be eliminated. As such, Theorem 4.1 guarantees that under common belief in future action-based rationality, both players can rationally play all of their actions at the non-terminal histories.

5 Order Independence

In this section we prove that the order, and speed, by which we eliminate the actions in the double-utility procedure is inessential for its output. We start by introducing reduction operators, after which we define the notion of order independence, provide a sufficient condition for order independence, show that the double-utility procedure can be written as the repeated application of a special reduction operator, and prove that this reduction operator satisfies the sufficient condition.

5.1 Reduction Operators

Consider a finite set A. A *reduction operator* r on A assigns to every non-empty set $D \subseteq A$ a non-empty subset $r(D) \subseteq D$. For a given set $D \subseteq A$ we set $r^0(D) := D$, and for every $k \in \mathbb{N}$ we inductively define $r^k(D) := r(r^{k-1}(D))$ as the k-fold application of the reduction operator r to D.

An *elimination order* for r is a finite sequence $(D^0, D^1, ..., D^K)$ where (a) $D^0 = A$, (b) $r(D^k) \subseteq D^{k+1} \subseteq D^k$ for every $k \in \{0, ..., K-1\}$, and (c) $r(D^K) = D^K$. Here, condition (a) states that we start with the full set A, condition (b) states that in round $k+1$ we eliminate at most as much from D^k as is allowed by r, but possibly less, whereas condition (c) guarantees that D^K cannot be reduced any further, and hence the elimination procedure terminates there. One special elimination order is the *full speed* elimination order $(D^0, D^1, ..., D^K)$, where $D^{k+1} = r(D^k)$ for every $k \in \{0, 1, ..., K-1\}$.

5.2 Order Independence

Some reduction operators have the special property that the final outcome will always be the same, no matter which elimination order is chosen. This property is called *order independence.*

Definition 5.1 (Order independence) *A reduction operator r is* ***order independent*** *if for every two elimination orders $(D^0, D^1, ..., D^K)$ and $(E^0, E^1, ..., E^L)$ we have that $D^K = E^L$.*

We will now introduce a condition, called *monotonicity,* which guarantees that the reduction operator is order independent.

Definition 5.2 (Monotonicity) *A reduction operator r is* ***monotone*** *if for every two sets D, E where $r(E) \subseteq D \subseteq E$, we have that $r(D) \subseteq r(E)$.*

Hence, monotonicity reveals the idea that smaller sets should yield smaller reduced sets. The following result shows that monotonicity implies order independence.

Lemma 5.1 (Monotonicity implies order independence) *Every monotone reduction operator is order independent.*

A proof can be found, for instance, in Perea (2025), proof of Lemma 3.6.2.

5.3 Order Independence of Double-Utility Procedure

The *double-utility procedure* can be formulated as the iterated application of some reduction operator r^{du}, as follows. The set on which r^{du} operates is $A := \times_{i \in I} \times_{h \in H} A_i(h)$, and we only consider product subsets of the form $E = \times_{i \in I} \times_{h \in H} E_i(h)$ where $E_i(h) \subseteq A_i(h)$ for every player i and history $h \in H$. For a given product set $E \subseteq A$, player i and history h, we denote by $E_i(h) \subseteq A_i(h)$ its projection on $A_i(h)$. For every product set $E \subseteq A$ and history $h \in H$, consider the double-utility game $G(h, E) = (E_i(h), u_i^{E+}[h], u_i^{E-}[h])_{i \in I}$ where

$$
\begin{gathered}
u_i^{E+}[h](a) := \max\{u_i(z) \mid z \in Z,\ z \succcurlyeq (h, a) \text{ and} \\
a_j(z, h') \in E_j(h') \text{ for all } j \in I \text{ and all } h' \in H \text{ with } z \succcurlyeq h' \succcurlyeq (h, a)\} \text{ and} \\
u_i^{E-}[h](a) := \min\{u_i(z) \mid z \in Z,\ z \succcurlyeq (h, a) \text{ and} \\
a_j(z, h') \in E_j(h') \text{ for all } j \in I \text{ and all } h' \in H \text{ with } z \succcurlyeq h' \succcurlyeq (h, a)\}
\end{gathered}
$$

for every player i and action profile $a \in \times_{i \in I} E_i(h)$. For every product set $E \subseteq A$, the reduction $D = r^{du}(E)$ is the product set given by

$$D_i(h) := \{a_i \in E_i(h) \mid a_i \text{ not strictly dominated in } G(h, E)\} \tag{5.1}$$

for every player i and history $h \in H$.

For every $k \in \mathbb{N}_0$, let $A^k = \times_{i \in I} \times_{h \in H} A_i^k(h)$ be the profile of action sets obtained in round k of the double-utility procedure. Then, by construction, $A^k = (r^{du})^k(A)$ for every $k \in \mathbb{N}_0$, and therefore the double-utility procedure is induced by the reduction operator r^{du}. It can be shown that the reduction operator r^{du} is monotone.

Lemma 5.2 (Double-utility operator is monotone) *The reduction operator r^{du} is monotone.*

Together with Lemma 5.1 we conclude that the reduction operator r^{du}, and thus the double-utility procedure, is order independent.

Corollary 5.1 (Double-utility procedure is order independent.) *The reduction operator r^{du} is order independent.*

This property will be used in the next section to show that in games with perfect information, the double-utility procedure reduces to the backward induction algorithm.

6 Games with Perfect Information

A dynamic game with observable actions is said to be with *perfect information* if at every non-terminal history h there is only one active player. Say that a dynamic game with perfect information is *without relevant ties* if for every player i, every history $h \in H$, every two distinct actions $a_i, a_i' \in A_i(h)$, and every two terminal histories $z, z' \succcurlyeq h$ such that $a_i = a_i(z, h)$ and $a_i' = a_i(z', h)$ it holds that $u_i(z) \neq u_i(z')$.

Consider a dynamic game Γ with perfect information and without relevant ties, and suppose we apply the double-utility procedure. We have seen in the previous section that this procedure is obtained by the iterated application of the reduction operator r^{du}, and Corollary 5.1 guarantees that the final output does not depend on the particular elimination order.

One particular elimination order is the *backwards elimination order,* where we first perform the eliminations at the last non-terminal histories in the game, then perform the eliminations at the second-to-last non-terminal histories, and so on, until we reach the beginning of the game. It turns out that this elimination order matches precisely the *backward induction procedure.*

To see this, consider a last non-terminal history $h \in H$ which is only followed by terminal histories and where player i is active. In the associated double-utility game at h, the upper and lower utilities for player i are the same. Hence, if we apply the reduction operator r^{du} to this double-utility game, it will only select the unique backward induction action $a^{bi}(h)$ at h.

Consider next a second-to-last non-terminal history $h \in H$ which is only followed by last non-terminal histories and terminal histories, and where player i is active. Suppose we have performed all the eliminations at the last non-terminal histories that follow h. Then, in the associated double-utility game at h, the upper and lower utility for player i after an action $a_i \in A_i(h)$ are the same. Indeed, they both coincide either with the utility $u_i(z)$ if a_i leads immediately to a terminal history z, or they both coincide with the utility obtained after the unique backward induction action following a_i. As such, the reduction operator r^{du} would select the unique backward induction action $a^{bi}(h)$ at h.

By continuing in this fashion, we see that applying the reduction operator r^{du} according to the backwards elimination order coincides precisely with the backward induction procedure. Hence, for every history $h \in H$ it would single out the unique backward induction action $a^{bi}(h)$. We thus arrive at the following conclusion.

Observation 6.1 (Equivalence with backward induction) *In a dynamic game with perfect information and without relevant ties, the only actions that survive the double-utility procedure are the backward induction actions.*

Together with Theorem 4.1, we obtain the following epistemic foundation for backward induction.

Corollary 6.1 (Epistemic foundation for backward induction) *Consider a dynamic game with perfect information and without relevant ties. Then, for every player, the only actions that can rationally be chosen under common belief in future action-based rationality are the backward induction actions.*

This result is related to Theorem 1 in Battigalli and de Vito (2021), which also offers epistemic conditions leading to backward induction in a model where the players choose actions, and not strategies, and where a player holds beliefs about his own future actions. The epistemic conditions in Battigalli and de Vito are *optimal planning* and *belief in continuation consistency*, and common full belief in these two events. One important difference is that Battigalli and de Vito (2021) assume forward consistency whereas we do not. Other differences will be discussed in Section 9.

7 Finitely Repeated Games

In this section we apply the double-utility procedure to finitely repeated games. We first define what a finitely repeated game is, and show that the double-utility procedure reduces to a very simple algorithm for this class of games. It can be used to calculate the sets of actions that the players can rationally choose under common belief in future action-based rationality at every stage of the repeated game. We finally explore some structural properties of these sets of actions at the various stages.

7.1 Definition

A *static game* is a tuple $G = (A_i, v_i)_{i \in I}$ where I is the finite set of players, A_i is the finite set of actions for player i, and $v_i : \times_{j \in I} A_j \to \mathbb{R}$ is player i's utility function. Set $A := \times_{i \in I} A_i$. For every $n \in \mathbb{N}$ and $\delta \in [0, 1]$, let $\Gamma(G,n,\delta)$ be the *n-fold repetition* of G with *discount factor* δ. That is, if the players choose the action profiles $a^1,...,a^n \in A$ at the stages $1,...,n$, then the utility for every player i is $\sum_{m=1}^{n} \delta^{m-1} \cdot v_i(a^m)$. We call $\Gamma(G,n,\delta)$ a *finitely repeated game* with discount factor $\delta \in [0,1]$, and we refer to G as the corresponding *stage game* of Γ. If we set $\delta = 1$, then Γ would be without discounting.

7.2 Strict Dominance by More Than α

Consider a static game $G = (A_i, v_i)_{i \in I}$ and a subset $D_i \subseteq A_i$ for every player i. For a given player i and number $\alpha_i \geq 0$, we say that an action $a_i \in D_i$ is *strictly dominated by more than* α_i on $\times_{j \in I} D_j$ if there is a randomized action $r_i \in \Delta(D_i)$ such that

$$v_i(a_i, a_{-i}) + \alpha_i < \sum_{a_i' \in D_i} r_i(a_i') \cdot v_i(a_i', a_{-i}) \text{ for all } a_{-i} \in D_{-i}.$$

Hence, by choosing $\alpha_i = 0$ we obtain the traditional notion of strict dominance.

Let $(\alpha_i)_{i \in I}$ be a profile of numbers, where $\alpha_i \geq 0$ for every player i. The *iterated elimination of actions that are strictly dominated by more than* $(\alpha_i)_{i \in I}$ is the sequence of action sets $(A^k)_{k \in \mathbb{N}_0}$ where $A^0 := \times_{i \in I} A_i$, and where for every $k \in \mathbb{N}$ and every player i,

$$A_i^k := \{a_i \in A_i^{k-1} \mid a_i \text{ is not strictly dominated by more than } \alpha_i \text{ on } \times_{j \in I} A_j^{k-1}\}.$$

Clearly, there will be some $K \in \mathbb{N}_0$ such that $A^K = A^{K+1}$, and the first such number K is the round where the procedure terminates. By the *iterated elimination of strictly dominated actions* we mean the procedure above where $\alpha_i = 0$ for every player i.

7.3 Procedure for Finitely Repeated Games

Consider the finitely repeated game $\Gamma(G,n,\delta)$. Similarly to what we did for games with perfect information, assume we apply to $\Gamma(G,n,\delta)$ the double-utility procedure with the *backwards elimination order*. What type of procedure do we obtain?

Start by applying the double-utility procedure to a history h at the last stage n. It may be verified that the elimination steps are exactly the ones of the iterated elimination of strictly dominated actions applied to the stage game G. For every player i, let A_i^{*n} be the set of actions that survive this procedure at any of those histories at stage n, and set $A^{*n} := \times_{i \in I} A_i^{*n}$.

Next, turn to a history h at stage $n-1$ of the repeated game. For every player i, let x_i be the total discounted utility collected up to history h. Then, the upper and lower utilities for player i in the double-utility game $G^{du}(h)$ at h are given by

$$\begin{aligned} u_i^+[h](a) &= x_i + \delta^{n-2} \cdot v_i(a) + \delta^{n-1} \cdot \max_{a' \in A^{*n}} v_i(a') \text{ and} \\ u_i^-[h](a) &= x_i + \delta^{n-2} \cdot v_i(a) + \delta^{n-1} \cdot \min_{a' \in A^{*n}} v_i(a'), \end{aligned}$$

respectively. Thus, an action a_i is strictly dominated in the double utility game $G^{du}(h)$ precisely when there is a randomized choice $r_i \in \Delta(A_i)$ such that

$$v_i(a_i, a_{-i}) + \delta \cdot \left(\max_{a' \in A^{n*}} v_i(a') - \min_{a' \in A^{n*}} v_i(a') \right) < \sum_{a'_i \in A_i} r_i(a'_i) \cdot v_i(a'_i, a_{-i}) \text{ for all } a_{-i} \in A_{-i}.$$

That is, precisely when a_i is strictly dominated by more than

$$\alpha_i^{n-1} := \delta \cdot \left(\max_{a' \in A^{n*}} v_i(a') - \min_{a' \in A^{n*}} v_i(a') \right)$$

in the stage game G.

As such, the elimination steps at h correspond exactly to the iterated elimination of actions that are strictly dominated by more than $(\alpha_i^{n-1})_{i \in I}$ in the stage game. For every player i, let A_i^{*n-1} be the set of actions that survive this procedure at any history at stage $n-1$, and set $A^{*n-1} := \times_{i \in I} A_i^{*n-1}$.

At stage $n-2$ we conclude, by a similar argument, that the elimination steps of the double-utility procedure correspond to the iterated elimination of actions that are strictly dominated by more than $(\alpha_i^{n-2})_{i \in I}$ in the stage game, where

$$\alpha_i^{n-2} := \delta \cdot \left(\max_{a' \in A^{*n-1}} v_i(a') - \min_{a' \in A^{*n-1}} v_i(a') + \alpha^{n-1} \right).$$

By continuing in this fashion, we arrive at the following elimination procedure, which we call the *double-utility procedure for finitely repeated games.*

Definition 7.1 (Procedure for finitely repeated games) *Consider a finitely repeated game* $\Gamma(G, n, \delta)$. *The* ***double-utility procedure for finitely repeated games*** *inductively defines, for every stage* $m \in \{1, ..., n\}$, *a profile of numbers* $(\alpha_i^m)_{i \in I}$ *and a set of actions profiles* $A^{*m} \subseteq A$, *as follows.*

At stage n, *define for every player* i *the number* $\alpha_i^n := 0$, *and let* A^{*n} *contain precisely those actions profiles in* A *that survive the iterated elimination of actions that are strictly dominated by more than* $(\alpha_i^n)_{i \in I}$ *in* G.

At every stage $m \in \{1, ..., n-1\}$, *inductively define for every player* i *the number*

$$\alpha_i^m := \delta \cdot \left(\max_{a \in A^{*m+1}} v_i(a) - \min_{a \in A^{*m+1}} v_i(a) + \alpha_i^{m+1} \right),$$

and let A^{*m} *contain precisely those actions profiles in* A *that survive the iterated elimination of actions that are strictly dominated by more than* $(\alpha_i^m)_{i \in I}$ *in* G.

We say that an action $a_i \in A_i$ *survives the double-utility procedure for finitely repeated games* at stage m if $a_i \in A_i^{*m}$.

Based on our arguments above, we see that the double-utility procedure with the backwards elimination order coincides precisely with the double-utility procedure for finitely repeated games. As, by Corollary 5.1, the particular elimination order does not matter for the eventual output in the double-utility procedure, we arrive at the following conclusion.

	e	f	g	h
a	$6,6$	$6,8$	$6,0$	$6,0$
b	$8,6$	$4,4$	$6,0$	$0,0$
c	$0,6$	$0,6$	$5,5$	$8,0$
d	$0,6$	$0,0$	$0,8$	$0,0$

Table 5: Illustration of action-based backward dominance procedure for finitely repeated games

Corollary 7.1 (Procedure for finitely repeated games) *In every finitely repeated game, the actions that survive the double-utility procedure are precisely the actions that survive the double-utility procedure for finitely repeated games at the various stages of the game.*

Together with Theorem 4.1, we thus see that the double-utility procedure for finitely repeated games selects precisely those actions that can rationally be chosen under common belief in future action-based rationality.

7.4 Example

Consider the static game G given by Table 5, which is symmetric between players 1 and 2. Here, player 1 chooses the row and player 2 chooses the column. Consider the induced repeated game $\Gamma(G, 3, 1)$ where $n = 3$ and $\delta = 1$. That is, there are three stages, and there is no discounting. We apply the double-utility procedure for finitely repeated games to determine the actions that the players can rationally play, at each of the three stages, under common belief in future action-based rationality.

Stage 3. By definition, $\alpha_1^3 = \alpha_2^3 = 0$. We thus perform the iterated elimination of actions that are strictly dominated by more than $(0, 0)$. In round 1, player 1's action d is strictly dominated by a by more than 0, and can thus be eliminated. Similarly, we can eliminate h for player 2. In round 2, player 1's action c is strictly dominated by a by more than 0, and can thus be eliminated. Similarly, we can eliminate g for player 2. Then, the eliminations terminate. We are thus left with the actions a, b for player 1, and e, f for player 2. That is, $A^{*3} = \{a, b\} \times \{e, f\}$.

Stage 2. Note that $\alpha_1^2 = \alpha_2^2 = 4$. To see this, observe that the highest and lowest utility for player 1 in A^{*3} are 8 and 4, respectively, and the same for player 2. Thus, by definition,

$$\alpha_1^2 = \delta \cdot ((8 - 4) + \alpha_1^3) = 1 \cdot (4 + 0) = 4,$$

and similarly for player 2. We therefore perform the iterated elimination of actions that are strictly dominated by more than $(4, 4)$. In round 1, player 1's action d is strictly dominated by a by more than 4, and can thus be eliminated. Similarly, we can eliminate h for player 2. Then, the eliminations terminate. Indeed, note that c is not strictly dominated by more than 4 after d and h have been eliminated. To see this, suppose that player 2 chooses g. Then, c yields 5 whereas a and b only yield 1 more. Similarly for g. Hence, the set of remaining action pairs is $A^{*2} = \{a, b, c\} \times \{e, f, g\}$.

Stage 1. Note that $\alpha_1^1 = \alpha_2^1 = 12$. To see this, observe that the highest and lowest utility for player 1 in A^{*2} are 8 and 0, respectively, and the same for player 2. Thus, by definition,

$$\alpha_1^1 = \delta \cdot ((8 - 0) + \alpha_1^2) = 1 \cdot (8 + 4) = 12,$$

and similarly for player 2. We therefore perform the iterated elimination of actions that are strictly dominated by more than $(12, 12)$. It can easily be verified that no action is strictly dominated by more than 12, and thus no action can be eliminated. Hence, the set of remaining action pairs is $A^{*1} = \{a, b, c, d\} \times \{e, f, g, h\}$.

We thus conclude that under common belief in future action-based rationality, the players can rationally play all of their actions in stage 1, can rationally play all of their actions except d and h in stage 2, and can rationally play the actions a, b and e, f in stage 3.

7.5 Monotonicity

Consider a finitely repeated game $\Gamma(G, n, \delta)$. Recall that the double-utility procedure yields, for every stage $m \in \{1, ..., n\}$, a profile $(\alpha_i^m)_{i \in I}$ of numbers, and a set A^{*m} of actions profiles in G. By definition, $\alpha_i^m \geq \alpha_i^{m+1}$ for all $m \in \{1, ..., n-1\}$. It may be verified, by induction on m, that $\alpha_i^m \geq \alpha_i^{m+1}$ and $A^{*m+1} \subseteq A^{*m}$ for every $m \in \{1, ..., n-1\}$ and all players i.

To see this, note first that $\alpha_i^{n-1} \geq \alpha_i^n$ as $\alpha_i^n = 0$ and $\alpha_i^{n-1} \geq 0$. Moreover, $A_i^{*n} \subseteq A_i^{*n-1}$, as A_i^{*n} and A_i^{*n-1} contain the actions that survive the iterated elimination of actions that are strictly dominated by more than $(\alpha_i^n)_{i \in I}$ and $(\alpha_i^{n-1})_{i \in I}$, respectively.

Next, consider some stage $m \in \{1, ..., n-2\}$, and assume that $\alpha_i^{m+1} \geq \alpha_i^{m+2}$ and $A_i^{*m+2} \subseteq A_i^{*m+1}$ for all players i. By definition, for every player i,

$$\alpha_i^m = \delta \cdot \left(\max_{a \in A^{*m+1}} v_i(a) - \min_{a \in A^{*m+1}} v_i(a) + \alpha_i^{m+1} \right).$$

As $A^{*m+2} \subseteq A^{*m+1}$ and $\alpha_i^{m+1} \geq \alpha_i^{m+2}$, we conclude that

$$\alpha_i^m \geq \delta \cdot \left(\max_{a \in A^{*m+2}} v_i(a) - \min_{a \in A^{*m+2}} v_i(a) + \alpha_i^{m+2} \right) = \alpha^{m+1}.$$

Since A_i^{*m} and A_i^{*m+1} contain those actions that survive the iterated elimination of actions that are strictly dominated by more than $(\alpha_i^m)_{i \in I}$ and $(\alpha_i^{m+1})_{i \in I}$, respectively, it follows that $A^{*m+1} \subseteq A^{*m}$. By induction on m, we thus arrive at the following conclusion.

Observation 7.1 (Monotonicity in finitely repeated games) *For a finitely repeated game $\Gamma = (G, n, \delta)$, and every stage $m \in \{1, ..., n\}$, let A^{*m} be the set of action profiles that survive the double-utility procedure at stage m. Then, $A^{*m+1} \subseteq A^{*m}$ for all $m \in \{1, ..., n-1\}$.*

Together with Theorem 4.1 and Corollary 5.1, it follows that common belief in future action-based rationality allows for more, or the same, actions if we move to an earlier stage in the repeated game. This result is rather intuitive: Consider an action a_i that is strictly dominated (by more than 0) in the stage game G, but nevertheless can rationally be chosen under common belief in future action-based rationality at stage $m+1$. This is possible since player i is allowed to believe that after choosing a_i, the future actions after stage $m+1$ will "work in his favour", whereas after choosing any other action, the future actions after stage $m+1$ will "work against him". Such a combination of "optimistic" and "pessimistic" beliefs can apparently make it optimal to choose the strictly dominated action a_i at stage $m+1$.

Now, suppose that player i finds himself at stage m, and considers playing the same action a_i. Since there is now one more future stage, player i can extend the "optimistic" and "pessimistic" beliefs above to make the utility gap between choosing a_i and not choosing a_i even bigger. This combination of beliefs would then certainly make it optimal to choose action a_i at stage m.

But we can say even more: Suppose that after performing the iterated elimination of strictly dominated actions in the stage game, at least two different utilities are possible for player i. Then, player i can rationally play every action at stage 1 of the repeated game, provided the number of stages and the discount factor are large enough.

Proposition 7.1 (When all actions are possible) *For a static game $G = (A_i, v_i)_{i \in I}$, let $A^* \subseteq \times_{i \in I} A_i$ be the set of action profiles that survive the iterated elimination of strictly dominated actions. For a given player i, suppose that there are action profiles $a, a' \in A^*$ with $v_i(a) \neq v_i(a')$. Then, there are $n^* \in \mathbb{N}$ and $\delta^* \in [0, 1)$, such that for every $n \in \mathbb{N}$ with $n \geq n^*$ and every $\delta \in [\delta^*, 1]$, every action $a_i \in A_i$ can rationally be chosen under common belief in future action-based rationality at stage 1 of the repeated game $\Gamma(G, n, \delta)$.*

This result has a flavor similar to the folk theorems for infinitely repeated games, stating that every feasible and individually rational utility profile in the stage game can be achieved under a (subgame perfect) Nash equilibrium in the repeated game if the discount factor is sufficiently large. But the result above is more extreme: It states that, under a certain regularity condition, every action will be possible under common belief in future action-based rationality at the beginning of the repeated game, provided the number of stages and the discount factor are large enough.

The intuition is rather simple: Consider a player i for whom there are action profiles $a, a' \in A^*$ with $v_i(a) > v_i(a')$, and consider an arbitrary action a_i^*. Player i is free to believe, at stage 1, that after choosing a_i^* he will be "rewarded" by the repetition of the action profile a in the future, and that after avoiding the action a_i^* he will be "punished" by the repetition of the action profile a' in the future. It may be verified that such a belief is in accordance with common belief in future action-based rationality. By making the number of stages and the discount factor large enough, it will then be optimal for player i to choose a_i^* at the first stage under such a belief.

8 Relation to Backward Dominance Procedure

In this section we show that the double-utility procedure is weakly more permissive than the backward dominance procedure. Perea (2014) introduced the conditions of *common belief in future (strategy-based) rationality,* and showed that the strategies that can rationally be played under these conditions are precisely those that survive the *backward dominance procedure.* The main difference with our approach is that in Perea (2014), a player is assumed to be able to commit to future actions. This shows in the procedure, which iteratively eliminates *strategies*, rather than actions. But it is also reflected in the epistemic conditions of common belief in future rationality, in which a player only holds beliefs about the opponents' strategies, not about his own future actions, and where a player always believes that his opponents choose *strategies* that are optimal now and in the future.

To formally describe the backward dominance procedure we need some further definitions. A *strategy* for player i is a mapping s_i that assigns to every history $h \in H$ some available action $s_i(h) \in A_i(h)$. Every strategy profile s in $\times_{i \in I} S_i$ induces a unique terminal history $z(s) \in Z$. We say that the strategy profile s *reaches* a history $h \in H$ if $z(s) \succcurlyeq h$. For a given history $h \in H$ and player i we define the sets

$$
\begin{aligned}
S(h) &:= \{s \in \times_{i \in I} S_i \mid s \text{ reaches } h\}, \\
S_i(h) &:= \{s_i \in S_i \mid \text{there is some } s_{-i} \in S_{-i} \text{ such that } (s_i, s_{-i}) \in S(h)\}, \text{ and} \\
S_{-i}(h) &:= \{s_{-i} \in S_{-i} \mid \text{there is some } s_i \in S_i \text{ such that } (s_i, s_{-i}) \in S(h)\}.
\end{aligned}
$$

For a given strategy $s_i \in S_i$ we denote by $H(s_i) := \{h \in H \mid s_i \in S_i(h)\}$ the collection of histories that the strategy s_i allows to be reached.

A *decision problem* at a history $h \in H$ is a profile $(D_i(h))_{i \in I}$ of strategy sets, where $D_i(h) \subseteq S_i(h)$ for all players i. A strategy $s_i \in D_i(h)$ is *strictly dominated* in the decision problem $(D_j(h))_{j \in I}$ if there is some randomized strategy $r_i \in \Delta(D_i(h))$ such that

$$
u_i(z(s_i, s_{-i})) < \sum_{s_i' \in D_i(h)} r_i(s_i') \cdot u_i(z(s_i', s_{-i})) \text{ for all } s_{-i} \in D_{-i}(h).
$$

Definition 8.1 (Backward dominance procedure (Perea (2014))) *(Initial step) For every history $h \in H$ define the decision problem $(S_i^0(h))_{i \in I}$ where $S_i^0(h) := S_i(h)$ for all players i.*

(Recursive step) Let $k \in \mathbb{N}$ and suppose that the decision problems $(S_i^{k-1}(h))_{i \in I}$ have been defined for all histories $h \in H$. Then define, for every history $h \in H$, the decision problem $(S_i^k(h))_{i \in I}$ where

$$
S_i^k(h) := \{s_i \in S_i^{k-1}(h) \mid s_i \text{ not strictly dominated in } (S_j^{k-1}(h'))_{j \in I} \text{ for any } h' \in H(s_i) \text{ with } h' \succcurlyeq h\}
$$

for every player i.

A strategy s_i survives the backward dominance procedure if $s_i \in S_i^k(h)$ for all $k \in \mathbb{N}$ and all histories $h \in H(s_i)$.

The following result shows that the backward dominance procedure is always at least as restrictive as the double-utility procedure. Indeed, every action that is allowed by the backward dominance procedure is also allowed by the double-utility procedure.

Theorem 8.1 (Relation with backward dominance procedure) *For a given player i, consider a strategy s_i that survives the backward dominance procedure. Then, for every history $h \in H(s_i)$, the action $s_i(h)$ survives the double-utility procedure.*

Together with Theorem 5.4 in Perea (2014) and Theorem 4.1 in the present paper, we conclude that every action that is prescribed by a strategy that can rationally be chosen under common belief in future (strategy-based) rationality, can also rationally be chosen under common belief in future action-based rationality.

In each of the examples in this paper, the backward dominance procedure turns out to be strictly more restrictive than the double-utility procedure. Indeed, in the game of Figure 1 we have seen that the backward dominance procedure rules out the strictly dominated strategy (a, d) for player 1, whereas the double-utility procedure allows for the actions a and d. In the game of Figure 2, the backward dominance procedure eliminates the strategy (a, g) for player 1 as it is strictly dominated by c at the beginning of the game. The double-utility procedure, however, allows for the actions a and g.

In the repeated game of Table 5, finally, the backward dominance procedure rules out every strategy for player 1 that selects c in stage 2 and subsequently selects a in stage 3. To see this, recall that at the end of the double-utility procedure, player 1, at stage 2, expects player 2 not to choose h at stages 2 and 3. As the backward dominance procedure is at least as restrictive, player 1 also believes so at stage 2 at the end of the backward dominance procedure. But then, by choosing c at stage 2 and a at stage 3, player 1 can expect at most 11 from the last two stages, whereas always choosing a at stages 2 and 3 yields him 12 for sure. As such, every strategy for player 1 that selects c in stage 2 and a in stage 3 is eliminated by the backward dominance procedure. At the same time, the double-utility procedure allows for the action c at stage 2 and the action a at stage 3.

9 Concluding Remarks

Comparison with Battigalli and de Vito (2021). In terms of modelling and epistemic conditions, our setup comes closest to Battigalli and de Vito (2021) (BdV from now on). Similarly to our approach, BdV assume that players choose actions, not strategies, and that a player, at every history, holds a probabilistic belief about his own actions and the opponents' actions. BdV encode higher-order beliefs for player i by types in an epistemic model that, at every history, hold a probabilistic belief about i's actions at all histories, the opponents' actions at all histories, and the opponents' types.

Differently from our model, BdV assume that a type for player i (a) satisfies forward consistency, (b) holds at every history h a belief about actions that are chosen at histories that do not follow h, (c) holds at every history a belief about his own current action, (d) does not have a belief about his own type, and (e) holds beliefs about his own actions that are independent from his beliefs about the opponents' actions and types. Because of (a) and (b), the types in BdV are not completely forward looking. Another difference is that BdV model a personal state for player i by a pair, consisting of i's "realized" actions, and his type. Our model, on the other hand, does not include such realized actions.

Our epistemic conditions of common belief in future action-based rationality are quite similar, at least in spirit, to the epistemic conditions in BdV. The three basic conditions in BdV are: (i) *optimal planning,* which states that a player believes that he himself chooses optimally at every history, given his beliefs there; (ii) *consistency,* which states that at a personal state for player i, the actions that player i is believed to make at every history according to his type coincide with the realized actions; and (iii) *belief in continuation consistency,* which states that a type believes, at every history h, in the opponents' consistency from h onwards.

As BdV assume that types satisfy forward consistency, they are able to show that every type for player i which satisfies optimal planning believes that player i will choose actions that, when taken together, will constitute optimal strategies. That is, forward consistency serves as a coordination device that turns profiles of individually optimal actions into optimal plans of actions. See their Remark 3. As such, the conditions in BdV rule out strictly dominated strategies. In contrast, we have seen in the various examples of this paper that our conditions of common belief in future action-based rationality allow for strictly dominated strategies.

It may be verified that the conditions of full belief in optimal planning and belief in continuation consistency in BdV jointly imply our condition of 1-fold belief in future action-based rationality. It is stronger, however, since BdV impose forward consistency. Similarly, their conditions of optimal planning, belief in continuation consistency, and common full belief in these two events imply common belief in future action-based rationality. BdV prove, in their Theorem 1, that the former epistemic conditions lead to the backward induction beliefs in games with perfect information without relevant ties. Their result can therefore be viewed as a counterpart to our Corollary 6.1.

For dynamic games with observable actions that violate perfect information, BdV show in Theorem 2 that the behavioral consequences of their epistemic conditions above are characterized by the *backwards rationalizability* procedure (Perea (2014), Penta (2015), Catonini and Penta (2025)). Recall that our epistemic conditions lead to the double-utility procedure which, by Theorem 8.1, is more permissive than backward dominance, and thereby also weaker than backwards rationalizability.

Related models. An important characteristic of our model is that a player, at each of the histories, holds a belief about his *own* future actions. Although this assumption seems quite natural, it is far from standard in the game theory literature. Together with Battigalli and de Vito (2021), the models by Battigalli and Siniscalchi (1999), Battigalli, di Tillio and Samet (2013) and Meier and Perea (2025) are amongst the few to incorporate this assumption.

Battigalli and Siniscalchi (1999) consider an epistemic model in which a type of a player is allowed to hold, at every history, a probabilistic belief about his *own strate*gy, the opponents' strategies and the opponents' types. Battigalli, di Tillio and Samet (2013) provide a variation of this model in which a type, at every history, holds a belief about the *path* that will be realized and the opponents' types. Such a belief induces, in particular, a belief about his own future actions. Meier and Perea (2025) present a model that reflects *cautious* reasoning by the players – a state of mind in which a player never discards any future action by himself or by his opponents. In their model a player holds, at the beginning of the game, a full-support belief with non-standard probabilities (Robinson (1973)) about his *own strategy* and the opponents' strategies, and updates this belief at every history by applying the rules of conditional probability. In particular, this model induces for every player conditional beliefs about his own future actions.

Extensions. In this paper we have restricted attention to *finite* dynamic games with *observable actions*. It remains to be investigated how our model and results can be extended to games with an infinite horizon, games with imperfectly observed past actions, and stochastic games with a finite and infinite horizon in which the transitions between histories are probabilistic rather than deterministic.

10 Appendix

10.1 Proofs of Section 4

Proof of Lemma 4.1. Consider a double-utility game $G = (A_i, u_i^+, u_i^-)_{i \in I}$, a player j, and an action $a_j^* \in A_j$. Let $A_{-j} := \times_{k \neq j} A_k$. Define a "standard" game $\hat{G} = (A_i, u_i)_{i \in I}$ where $u_j : A_j \times A_{-j} \to \mathbb{R}$ is given by

$$u_j(a_j, a_{-j}) := \begin{cases} u_j^+(a_j, a_{-j}), & \text{if } a_j = a_j^* \\ u_j^-(a_j, a_{-j}), & \text{if } a_j \neq a_j^* \end{cases}$$

for all $(a_j, a_{-j}) \in A_j \times A_{-j}$. Then, by construction, a_j^* is strictly dominated in the double-utility game G, if and only if, a_j^* is strictly dominated (in the normal sense) in the standard game $\hat{G}$. Moreover, a_j^* is optimal for some belief about the opponents' action profiles in the double-utility game G, if and only if, a_j^* is optimal (in the normal sense) for some belief about the opponents' action profiles in the standard game $\hat{G}$.

By Lemma 3 in Pearce (1984) we know that a_j^* is strictly dominated in the standard game $\hat{G}$, if and only if, a_j^* is not optimal for any belief about the opponents' action profiles in the standard game $\hat{G}$. Together with the two equivalences above, this completes the proof. ■

To prove Theorem 4.1 we need a preparatory result, which requires some additional notation. For a history $h \in H$ and $k \in \mathbb{N}_0$, let $G^k(h) = (A_i^k(h), u_i^{k+}[h], u_i^{k-}[h])_{i \in I}$ be the double-utility game at round k of the double-utility procedure. Although the utility functions $u_i^{k+}[h], u_i^{k-}[h]$ are only defined for action profiles in $\times_{i \in I} A_i^k(h)$, they can be extended to all action profiles in $\times_{i \in I} A_i(h)$ by setting

$$\begin{gathered} u_i^{k+}[h](a) := \max\{u_i(z) \mid z \in Z,\ z \succcurlyeq (h,a) \text{ and} \\ a_j(z,h') \in A_j^k(h') \text{ for all } j \in I \text{ and all } h' \in H \text{ with } z \succcurlyeq h' \succcurlyeq (h,a)\} \text{ and} \\ u_i^{k-}[h](a) := \min\{u_i(z) \mid z \in Z,\ z \succcurlyeq (h,a) \text{ and} \\ a_j(z,h') \in A_j^k(h') \text{ for all } j \in I \text{ and all } h' \in H \text{ with } z \succcurlyeq h' \succcurlyeq (h,a)\} \end{gathered}$$

for all $a \in \times_{i \in I} A_i(h)$. Also, we denote by $A_{-i}^k(h) := \times_{j \neq i} A_j^k(h)$ the set of opponents' action profiles, and by $A^k(h) := \times_{j \in I} A_j^k(h)$ the set of all players' action profiles, in $G^k(h)$.

Lemma 10.1 (Optimality of actions in procedure) *Consider a history h, a round $k \in \mathbb{N}_0$, a player i and an action $a_i^* \in A_i(h)$. Then, $a_i^* \in A_i^{k+1}(h)$, if and only if, there is a belief $\beta_i \in \Delta(A_{-i}^k(h))$ such that*

$$\sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i(h).$$

Proof. We prove the statement by induction on k.

Induction start. Let $k = 0$. For some history $h \in H$, consider the double-utility game $G^0(h) = (A_i^0(h), u_i^{0+}[h], u_i^{0-}[h])_{i \in I}$, some player i, and an action $a_i^* \in A_i(h)$. Recall that $A_i^0(h) = A_i(h_i)$. By definition, $a_i^* \in A_i^1(h)$ if and only if a_i^* is not strictly dominated in $G^0(h)$. By Lemma 4.1 applied to $G^0(h)$, we conclude that $a_i^* \in A_i^1(h)$, if and only if, there is a belief $\beta_i \in \Delta(A_{-i}^0(h))$ such that

$$\sum_{a_{-i} \in A_{-i}^0(h)} \beta_i(a_{-i}) \cdot u_i^{0+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}^0(h)} \beta_i(a_{-i}) \cdot u_i^{0-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i(h).$$

This concludes the induction start.

Induction step. Now, take some $k \in \mathbb{N}$ and assume that the statement holds for $k-1$. For some history $h \in H$, consider the double-utility game $G^k(h) = (A_i^k(h), u_i^{k+}[h], u_i^{k-}[h])_{i \in I}$, some player i, and an action $a_i^* \in A_i(h)$.

To show the "only if" part, suppose that $a_i^* \in A_i^{k+1}(h)$. By definition, this means that a_i^* is not strictly dominated in $G^k(h)$. Lemma 4.1 applied to $G^k(h)$ then implies that there is a belief $\beta_i \in \Delta(A_{-i}^k(h))$ such that

$$\sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i^k(h). \quad (10.1)$$

Let $\hat{a}_i \in A_i(h)$ be such that

$$\sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](\hat{a}_i, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i(h). \quad (10.2)$$

Then, we have in particular that

$$\sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k+}[h](\hat{a}_i, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i(h). \quad (10.3)$$

As $\beta_i \in \Delta(A_{-i}^k(h))$ and $A_{-i}^k(h) \subseteq A_{-i}^{k-1}(h)$, we have that $\beta_i \in \Delta(A_{-i}^{k-1}(h))$. Moreover, it holds by definition that $u_i^{k-1+}[h](a_i, a_{-i}) \geq u_i^{k+}[h](a_i, a_{-i})$ and $u_i^{k-1-}[h](a_i, a_{-i}) \leq u_i^{k-}[h](a_i, a_{-i})$ for all $(a_i, a_{-i}) \in A(h)$. Together with (10.3) we conclude that

$$\sum_{a_{-i} \in A_{-i}^{k-1}(h)} \beta_i(a_{-i}) \cdot u_i^{k-1+}[h](\hat{a}_i, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}^{k-1}(h)} \beta_i(a_{-i}) \cdot u_i^{k-1-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i(h), \quad (10.4)$$

where $\beta_i \in \Delta(A_{-i}^{k-1}(h))$. By the induction assumption we then know that $\hat{a}_i \in A_i^k(h)$.

By (10.1) and (10.2) it then follows that

$$\sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](\hat{a}_i, a_{-i})$$

$$\geq \sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i(h),$$

where $\beta_i \in \Delta(A_{-i}^k(h))$. This proves the "only if" part.

For the "if" part, suppose that there is a belief $\beta_i \in \Delta(A_{-i}^k(h))$ such that

$$\sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i(h). \quad (10.5)$$

As $\beta_i \in \Delta(A_{-i}^k(h))$ and $A_{-i}^k(h) \subseteq A_{-i}^{k-1}(h)$, we have that $\beta_i \in \Delta(A_{-i}^{k-1}(h))$. Moreover, it holds by definition that $u_i^{k-1+}[h](a_i, a_{-i}) \geq u_i^{k+}[h](a_i, a_{-i})$ and $u_i^{k-1-}[h](a_i, a_{-i}) \leq u_i^{k-}[h](a_i, a_{-i})$ for all $(a_i, a_{-i}) \in A(h)$. Together with (10.5) we conclude that

$$\sum_{a_{-i} \in A_{-i}^{k-1}(h)} \beta_i(a_{-i}) \cdot u_i^{k-1+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}^{k-1}(h)} \beta_i(a_{-i}) \cdot u_i^{k-1-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i(h), \quad (10.6)$$

where $\beta_i \in \Delta(A_{-i}^{k-1}(h))$. By the induction assumption we then know that $a_i^* \in A_i^k(h)$.

Together with (10.5) it thus follows that

$$\sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}^k(h)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i^k(h),$$

where $\beta_i \in \Delta(A_{-i}^k(h))$ and $a_i^* \in A_i^k(h)$. Hence, a_i^* is optimal for some belief in the double-utility game $G^k(h)$. By applying Lemma 4.1 to $G^k(h)$ we conclude that a_i^* is not strictly dominated in $G^k(h)$, and hence $a_i^* \in A_i^{k+1}(h)$. This proves the "if" part, and thereby the proof of the induction step is complete.

By induction on k, the statement of the lemma follows. ■

Proof of Theorem 4.1. We prove the theorem through a sequence of claims. For every player i, history $h \in H$, and $k \in \mathbb{N}_0$, let

$$BR_i^k(h) := \{a_i \in A_i(h) \mid \text{there is an epistemic model } M = (T_j(h), b_j)_{j \in I, h \in H} \text{ and some } t_i \in T_i(h) \text{ such that } t_i \text{ expresses up to } k\text{-fold belief in future action-based rationality, and the action } a_i \text{ is optimal for } t_i\}$$

Remember that $A_i^{k+1}(h)$ is the set of player i actions at h that survives round $k+1$ of the double-utility procedure.

Claim 1. For every player i, history $h \in H$, and $k \in \mathbb{N}_0$, it holds that $BR_i^k(h) \subseteq A_i^{k+1}(h)$.

Proof of claim 1. We proceed by induction on k.

Induction start. Consider first $k = 0$, and some $a_i^* \in BR_i^0(h)$. Then, there is an epistemic model $M = (T_j(h), b_j)_{j \in I, h \in H}$ and some type $t_i \in T_i(h)$ such that the action a_i^* is optimal for t_i. Then, action a_i^* is optimal at h_i for the first-order belief $b_i := b_i^1(t_i) \in \Delta(A^{i, \succcurlyeq h})$. That is,

$$\sum_{a \in A^{i,\succcurlyeq h}} b_i(a) \cdot u_i(z(h, a_i^*, a)) \geq \sum_{a \in A^{i,\succcurlyeq h}} b_i(a) \cdot u_i(z(h, a_i, a)) \tag{10.7}$$

for all $a_i \in A_i(h)$. Define the belief $\beta_i := \text{marg}_{A_{-i}(h)} b_i \in \Delta(A_{-i}(h))$. For every $a \in A^{i,\succcurlyeq h}$, let $a_{-i}(h) \in A_{-i}(h)$ be its projection on $A_{-i}(h)$.

Then, we have that

$$\begin{aligned} \sum_{a \in A^{i,\succcurlyeq h}} b_i(a) \cdot u_i(z(h, a_i^*, a)) &= \sum_{a_{-i} \in A_{-i}(h)} \sum_{a \in A^{i,\succcurlyeq h} : a_{-i}(h) = a_{-i}} b_i(a) \cdot u_i(z(h, a_i^*, a)) \\ &\leq \sum_{a_{-i} \in A_{-i}(h)} \sum_{a \in A^{i,\succcurlyeq h} : a_{-i}(h) = a_{-i}} b_i(a) \cdot u_i^{0+}[h](a_i^*, a_{-i}) \\ &= \sum_{a_{-i} \in A_{-i}(h)} \beta_i(a_{-i}) \cdot u_i^{0+}[h](a_i^*, a_{-i}), \end{aligned} \tag{10.8}$$

where the inequality follows from the fact that $z(h, a_i^*, a) \succcurlyeq (h, (a_i^*, a_{-i}))$ for all $a \in A^{i,\succcurlyeq h}$ with $a_{-i}(h) = a_{-i}$.

Similarly, it can be shown for every $a_i \in A_i(h_i)$ that

$$\sum_{a \in A^{i,\succcurlyeq h}} b_i(a) \cdot u_i(z(h, a_i, a)) \geq \sum_{a_{-i} \in A_{-i}(h)} \beta_i(a_{-i}) \cdot u_i^{0-}[h](a_i, a_{-i}). \tag{10.9}$$

By combining (10.7), (10.8) and (10.9) we conclude that

$$\sum_{a_{-i} \in A_{-i}(h)} \beta_i(a_{-i}) \cdot u_i^{0+}(a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}(h)} \beta_i(a_{-i}) \cdot u_i^{0-}(a_i, a_{-i})$$

for all $a_i \in A_i(h_i)$.

That is, a_i^* is optimal within the double-utility game $G^0(h)$ for the belief $\beta_i \in \Delta(A_{-i}(h))$. By Lemma 4.1 we conclude that a_i^* is not strictly dominated within $G^0(h)$. Therefore, $a_i^* \in A_i^1(h)$, which was to show.

Induction step. Suppose that $k \geq 1$, and that $BR_j^{k-1}(h) \subseteq A_j^k(h)$ for every player j and history $h \in H$. Take some player i and history $h \in H$. We will show that $BR_i^k(h) \subseteq A_i^{k+1}(h)$.

Take some $a_i^* \in BR_i^k(h)$. Then, there is an epistemic model $M = (T_j(h'), b_j)_{j \in I, h' \in H}$ and some type $t_i \in T_i(h)$ that expresses up to k-fold belief in future action-based rationality, such that the action a_i^* is optimal for t_i. Hence, action a_i^* is (i) optimal at h for the first-order belief $b_i := b_i^1(t_i) \in \Delta(A^{i,\succeq h})$, and (ii) $b_i(t_i)$ only assigns positive probability to pairs (a_j, t_j') (allowing for $j = i$) where a_j is optimal for t_j', and t_j' expresses up to $(k-1)$-fold belief in future action-based rationality.

For every action profile $a \in A^{i,\succeq h}$, every history $h' \succeq h$ and every player j, let $a_j(h')$ be the action prescribed for player j at h' if either $j \neq i$, or $j = i$ and $h' \succ h$. By (i) and (ii) above it thus follows that the first-order belief b_i only assigns positive probability to action profiles $a \in A^{i,\succeq h}$ where $a_j(h') \in BR_j^{k-1}(h')$ for all $j \neq i$ and $h' \succeq h$, and $a_i(h') \in BR_i^{k-1}(h')$ for all $h' \succ h$. By the induction assumption, we thus know that b_i only assigns positive probability to action profiles $a \in A^{i,\succeq h}$ where (iii) $a_j(h') \in A_j^k(h')$ for all $j \neq i$ and $h' \succeq h$, and (iv) $a_i(h') \in A_i^k(h')$ for all $h' \succ h$.

Recall that action a_i^* is optimal at h for the first-order belief b_i. That is,

$$\sum_{a \in A^{i,\succeq h}} b_i(a) \cdot u_i(z(h, a_i^*, a)) \geq \sum_{a \in A^{i,\succeq h}} b_i(a) \cdot u_i(z(h, a_i, a)) \tag{10.10}$$

for all $a_i \in A_i(h)$.

Define $\beta_i := \text{marg}_{A_{-i}(h_i)} b_i \in \Delta(A_{-i}(h))$. By property (iii) we know, in fact, that $\beta_i \in \Delta(A_{-i}^k(h))$. We have that

$$\begin{aligned} \sum_{a \in A^{i,\succeq h}} b_i(a) \cdot u_i(z(h, a_i^*, a)) &= \sum_{a_{-i} \in A_{-i}(h)} \sum_{a \in A^{i,\succeq h} : a_{-i}(h) = a_{-i}} b_i(a) \cdot u_i(z(h, a_i^*, a)) \\ &\leq \sum_{a_{-i} \in A_{-i}(h)} \sum_{a \in A^{i,\succeq h} : a_{-i}(h) = a_{-i}} b_i(a) \cdot u_i^{k+}[h](a_i^*, a_{-i}) \\ &= \sum_{a_{-i} \in A_{-i}(h)} \beta_i(a_{-i}) \cdot u_i^{k+}[h](a_i^*, a_{-i}). \end{aligned} \tag{10.11}$$

To see why the inequality holds, take some $a \in A^{i,\succeq h}$ with $a_{-i}(h) = a_{-i}$ and $b_i(a) > 0$. By (iii) and (iv) above, we know that $a_j(h') \in A_j^k(h')$ for all $j \neq i$ and $h' \succeq h$, and $a_i(h') \in A_i^k(h')$ for all $h' \succ h$. Thus, $u_i(z(h, a_i^*, a)) \leq u_i^{k+}[h](a_i^*, a_{-i})$.

Similarly, it can be shown for every $a_i \in A_i(h_i)$ that

$$\sum_{a \in A^{i,\succeq h}} b_i(a) \cdot u_i(z(h, a_i, a)) \geq \sum_{a_{-i} \in A_{-i}(h_i)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i}). \tag{10.12}$$

By combining (10.10), (10.11) and (10.12) we conclude that

$$\sum_{a_{-i} \in A_{-i}(h_i)} \beta_i(a_{-i}) \cdot u_i^{k+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A_{-i}(h_i)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i})$$

for all $a_i \in A_i(h_i)$.

As $\beta_i \in \Delta(A^k_{-i}(h))$ it follows from Lemma 10.1 that $a_i^* \in A_i^{k+1}(h)$, which was to show.

By induction on k we thus conclude that $BR_i^k(h) \subseteq A_i^{k+1}(h)$ for all players i, histories $h \in H$ and $k \in \mathbb{N}_0$. This completes the proof of claim 1. $\Diamond$

For the next claim we need some new notation. For a history h, player i and number $k \in \mathbb{N}_0$, define

$$(A^k)^{i,\succcurlyeq h} := \{a \in A^{i,\succcurlyeq h} \mid a_i(h') \in A_i^k(h') \text{ for all } h' \succ h, \text{ and}$$
$$a_j(h') \in A_j^k(h') \text{ for all } h' \succcurlyeq h \text{ and } j \neq i\}$$

Claim 2. For every player i, history $h \in H$, number $k \in \mathbb{N}_0$ and action $a_i^* \in A_i^{k+1}(h)$, there is some belief $b_i(a_i^*) \in \Delta((A^k)^{i,\succcurlyeq h})$ for which the action a_i^* is optimal at h.

Proof of claim 2. Take some $a_i^* \in A_i^{k+1}(h)$. Then, by Lemma 10.1, there is some belief $\beta_i \in \Delta(A^k_{-i}(h))$ such that

$$\sum_{a_{-i} \in A^k_{-i}(h)} \beta_i(a_{-i}) \cdot u_i^{k+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in A^k_{-i}(h)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i}) \text{ for all } a_i \in A_i(h). \tag{10.13}$$

Take some $a_{-i} \in A^k_{-i}(h)$. Then, by definition of $u_i^{k+}[h](a_i^*, a_{-i})$, there is some $z^+(h, a_i^*, a_{-i}) \in Z$ such that

$$u_i^{k+}[h](a_i^*, a_{-i}) = u_i(z^+(h, a_i^*, a_{-i})) \tag{10.14}$$

and

$$a_j(z^+(h, a_i^*, a_{-i}), h') \in A_j^k(h') \text{ for all } j \in I \text{ and}$$
$$\text{all } h' \in H \text{ with } z^+(h, a_i^*, a_{-i}) \succcurlyeq h' \succcurlyeq (h, (a_i^*, a_{-i})). \tag{10.15}$$

Similarly, for every $a_i \in A_i(h_i) \backslash \{a_i^*\}$ and $a_{-i} \in A^k_{-i}(h)$ there is some $z^-(h, a_i, a_{-i}) \in Z$ such that

$$u_i^{k-}[h](a_i^-, a_{-i}) = u_i(z^-(h, a_i^*, a_{-i})) \tag{10.16}$$

and

$$a_j(z^-(h, a_i, a_{-i}), h') \in A_j^k(h') \text{ for all } j \in I \text{ and}$$
$$\text{all } h' \in H \text{ with } z^-(h, a_i, a_{-i}) \succcurlyeq h' \succcurlyeq (h, (a_i, a_{-i})). \tag{10.17}$$

By (10.15) and (10.17) we can find, for every $a_{-i} \in A^k_{-i}(h)$, an action profile $a[a_{-i}] \in (A^k)^{i,\succcurlyeq h}$ such that

$$a[a_{-i}] \text{ selects at every } h' \text{ with } z^+(h, a_i^*, a_{-i}) \succcurlyeq h' \succcurlyeq (h, (a_i^*, a_{-i})) \tag{10.18}$$
$$\text{the actions } a_j(z^+(h, a_i^*, a_{-i}), h') \text{ for every } j \in I$$

and

$$a[a_{-i}] \text{ selects at every } h' \text{ with } z^{-}(h, a_i, a_{-i}) \succcurlyeq h' \not\succcurlyeq (h, (a_i, a_{-i})) \quad (10.19)$$
$$\text{for some } a_i \in A_i(h)\backslash\{a_i^*\} \text{ the actions } a_j(z^{-}(h, a_i^*, a_{-i}), h') \text{ for every } j \in I.$$

Let $b_i(a_i^*) \in \Delta((A^k)^{i,\succcurlyeq h})$ be the belief that, for every $a_{-i} \in A_{-i}^k(h)$, assigns probability $\beta_i(a_{-i})$ to the action profile $a[a_{-i}]$ defined above, and assigns probability zero to all other action profiles in $A^{i,\succcurlyeq h}$.

On the basis of (10.14), (10.16), (10.18) and (10.19), we then have that

$$u_i(a_i^*, b_i(a_i^*)) = \sum_{a_{-i} \in A_{-i}^k(h_i)} \beta_i(a_{-i}) \cdot u_i^{k+}[h](a_i^*, a_{-i}) \quad (10.20)$$

and, for every $a_i \in A_i^k(h_i)\backslash\{a_i^*\}$,

$$u_i(a_i, b_i(a_i^*)) = \sum_{a_{-i} \in A_{-i}^k(h_i)} \beta_i(a_{-i}) \cdot u_i^{k-}[h](a_i, a_{-i}). \quad (10.21)$$

From (10.20), (10.21) and (10.13) we then conclude that

$$u_i(a_i^*, b_i(a_i^*)) \geq u_i(a_i, b_i(a_i^*)) \text{ for all } a_i \in A_i(h_i).$$

Hence, the action $a_i^* \in A_i^{k+1}(h)$ is optimal at h for the belief $b_i(a_i^*) \in \Delta((A^k)^{i,\succcurlyeq h})$, which was to show. This completes the proof of claim 2. ◊

We will now construct an epistemic model $M = (T_i(h), b_i)_{i \in I, h \in H}$, as follows. For every player i and history h, consider the set of types

$$T_i(h) := \{t_i(h, a_i) \mid a_i \in A_i(h)\}.$$

Take some type $t_i(h, a_i)$ in $T_i(h)$. To define the belief $b_i(t_i(h, a_i)) \in \Delta(A^{i,\succcurlyeq h} \times T^{i,\succcurlyeq h})$ we distinguish three cases. For case 2 and 3, let K be such that $A_j^K(h') = A_j^{K+1}(h')$ for all players j and all $h' \in H$ in the double-utility procedure.

Case 1. Suppose that $a_i \notin A_i^1(h)$. Then, we define $b_i(t_i(h, a_i)) \in \Delta(A^{i,\succcurlyeq h} \times T^{i,\succcurlyeq h})$ arbitrarily.

Case 2. Suppose that $a_i \in A_i^k(h)\backslash A_i^{k+1}(h)$ for some $k \in \{1, ..., K-1\}$. Then, we know from claim 2 that a_i is optimal at h for some belief $b_i(h, a_i) \in \Delta((A^{k-1})^{i,\succcurlyeq h})$. For every action profile $a \in A^{i,\succcurlyeq h}$ let

$$t(a) := ((t_i(h', a_i(h')))_{h' \succ h}, (t_j(h', a_j(h')))_{j \neq i, h' \succcurlyeq h}$$

be the associated type profile in $T^{i,\succcurlyeq h}$. Define the belief $b_i(t_i(h, a_i))$ by

$$b_i(t_i(h, a_i))(a, t) := \begin{cases} b_i(h, a_i)(a), & \text{if } t = t(a) \\ 0, & \text{otherwise} \end{cases} \quad (10.22)$$

for all $(a,t) \in A^{i,\succcurlyeq h} \times T^{i,\succcurlyeq h}$.

Case 3. Suppose that $a_i \in A_i^K(h)$. As $A_i^K(h) = A_i^{K+1}(h)$, it follows that $a_i \in A_i^{K+1}(h)$. Hence, we know from claim 2 that a_i is optimal at h for some belief $b_i(h,a_i) \in \Delta((A^K)^{i,\succcurlyeq h})$. Define the belief $b_i(t_i(h,a_i))$ by

$$b_i(t_i(h,a_i))(a,t) := \begin{cases} b_i(h,a_i)(a), & \text{if } t = t(a) \\ 0, & \text{otherwise} \end{cases} \tag{10.23}$$

for all $(a,t) \in A^{i,\succcurlyeq h} \times T^{i,\succcurlyeq h}$.

Claim 3. For every player i, history $h \in H$ and $k \in \mathbb{N}_0$ such that $a_i \in A_i^{k+1}(h)$, we have that a_i is optimal for $t_i(h,a_i)$, and $t_i(h,a_i)$ expresses up to k-fold belief in future action-based rationality.

Proof of claim 3. By induction on k. Start with $k=0$, and take some $a_i \in A_i^1(h)$. By (10.22) and (10.23) we know that $b_i^1(t_i(h,a_i)) = b_i(h,a_i)$. As, by construction, a_i is optimal at h for the belief $b_i(h,a_i)$, we conclude that a_i is optimal for $t_i(h,a_i)$. Moreover, by definition, $t_i(h,a_i)$ expresses 0-fold belief in future action-based rationality.

Now, let $k \geq 1$, and suppose that the claim holds for $k-1$. Take some $a_i \in A_i^{k+1}(h)$. As $a_i \in A_i^k(h)$ we know, by the induction assumption, that a_i is optimal for $t_i(h,a_i)$. Moreover, by (10.22) and (10.23), the belief $b_i(t_i(h,a_i))$ only assigns positive probability to tuples $(a,t) \in A^{i,\succcurlyeq h} \times T^{i,\succcurlyeq h}$ where $a \in (A^k)^{i,\succcurlyeq h}$. Choose some (a,t) that receives positive probability by $b_i(t_i(h,a_i))$.

Take some $h' \succ h$. Then, by construction, the projection of (a,t) on $A_i(h') \times T_i(h')$ is some $(a_i', t_i(h',a_i'))$ where $a_i' \in A_i^k(h')$. By he induction assumption we thus know that a_i' is optimal for $t_i(h',a_i')$, and that $t_i(h',a_i')$ expresses up to $(k-1)$-fold belief in future action-based rationality.

Next, take some $h' \succcurlyeq h$ and some $j \neq i$. Then, by construction, the projection of (a,t) on $A_j(h') \times T_j(h')$ is some $(a_j', t_j(h',a_j'))$ where $a_j' \in A_j^k(h')$. By the induction assumption we thus know that a_j' is optimal for $t_j(h',a_j')$, and that $t_j(h',a_j')$ expresses up to $(k-1)$-fold belief in future action-based rationality.

By the two insights above, we conclude that $t_i(h,a_i)$ expresses up to k-fold belief in future action-based rationality. By induction on k, the proof of claim 3 is complete. $\diamondsuit$

Claim 4. For every player i, history $h \in H$ and $k \in \mathbb{N}_0$, it holds that $A_i^{k+1}(h) \subseteq BR_i^k(h)$.

Proof of claim 4. Take some $a_i \in A_i^{k+1}(h)$. By claim 3 we conclude that a_i is optimal for the type $t_i(h,a_i)$, and that $t_i(h,a_i)$ expresses up to k-fold belief in future action-based rationality at h_i. Thus, by definition $a_i \in BR_i^k(h)$, which was to show. $\diamondsuit$

By claims 1 and 4 it follows that $BR_i^k(h) = A_i^{k+1}(h)$ for every player i, history $h \in H$, and $k \in \mathbb{N}_0$. With this result we can now prove parts **(a)** and **(b)** of the theorem.

(a) Fix some $k \in \mathbb{N}_0$. Take first an action $a_i \in A_i(h)$ that can rationally be chosen under up to k-fold belief in future action-based rationality. Then, $a_i \in BR_i^k(h)$. As $BR_i^k(h) = A_i^{k+1}(h)$ it follows that $a_i \in A_i^{k+1}(h)$ and hence a_i survives round $k+1$ of the double-utility procedure.

Next, take an action $a_i \in A_i(h)$ that survives round $k+1$ of the double-utility procedure. Hence, $a_i \in A_i^{k+1}(h)$. As $BR_i^k(h) = A_i^{k+1}(h)$ we know that $a_i \in BR_i^k(h)$. That is, a_i can rationally be chosen under up to k-fold belief in future action-based rationality. Together with the insight above, this completes the proof of **(a)**.

(b) To prove part **(b)** of the theorem, we need the following result.

Claim 5. For every player i, history $h \in H_i$ and action $a_i \in A_i^K(h)$, the type $t_i(h, a_i)$ expresses common belief in future action-based rationality.

Proof of claim 5. We prove, by induction on k, that for every player i, history $h \in H$ and action $a_i \in A_i^K(h)$, the type $t_i(h, a_i)$ expresses up to k-fold belief in future action-based rationality.

For $k = 0$ the statement is true by definition. Now, take some $k \geq 1$ and suppose that the statement is true for $k - 1$. Take some player i, history $h \in H$ and action $a_i \in A_i^K(h)$. By (10.23), the belief $b_i(t_i(h, a_i))$ only assigns positive probability to tuples $(a, t) \in A^{i,\succcurlyeq h} \times T^{i,\succcurlyeq h}$ where $a \in (A^K)^{i,\succcurlyeq h}$. Choose some (a, t) that receives positive probability by $b_i(t_i(h, a_i))$.

Take some $h' \succ h$. Then, by construction, the projection of (a, t) on $A_i(h') \times T_i(h')$ is some $(a_i', t_i(h', a_i'))$ where $a_i' \in A_i^K(h')$. By claim 3, the action a_i' is optimal for $t_i(h', a_i')$. Moreover, as $a_i' \in A_i^K(h')$ we know by the induction assumption that $t_i(h', a_i')$ expresses up to $(k-1)$-fold belief in future action-based rationality.

Next, take some $h' \succcurlyeq h$ and some $j \neq i$. Then, by construction, the projection of (a, t) on $A_j(h') \times T_j(h')$ is some $(a_j', t_j(h', a_j'))$ where $a_j' \in A_j^K(h')$. By claim 3, the action a_j' is optimal for $t_j(h', a_j')$. Moreover, as $a_j' \in A_j^K(h')$ we know by the induction assumption that $t_j(h', a_j')$ expresses up to $(k-1)$-fold belief in future action-based rationality.

By the two insights above, we conclude that $t_i(h, a_i)$ expresses up to k-fold belief in future action-based rationality. By induction on k, it follows that $t_i(h, a_i)$ expresses common belief in future action-based rationality. This completes the proof of claim 5. $\Diamond$

We can now prove part **(b)** of the theorem. Take first an action $a_i \in A_i(h)$ that can rationally be chosen under common belief in future action-based rationality. Then, $a_i \in BR_i^k(h)$ for all $k \in \mathbb{N}_0$. As $BR_i^k(h) = A_i^{k+1}(h)$ for all $k \in \mathbb{N}_0$, it follows that $a_i \in A_i^{k+1}(h)$ for all $k \in \mathbb{N}_0$. That is, a_i survives the double-utility procedure.

Next, take an action $a_i \in A_i(h)$ that survives the double-utility procedure. Then, $a_i \in A_i^K(h)$. By claim 3, a_i is optimal for the type $t_i(h, a_i)$. Moreover, by claim 5, type $t_i(h, a_i)$ expresses common belief in future action-base rationality. Altogether, we see that a_i can rationally be chosen under common belief in future action-based rationality. This completes the proof of part **(b)**. ■

10.2 Proof of Section 5

Proof of Lemma 5.2. Take some product sets $D, E \subseteq A$ with $r^{du}(E) \subseteq D \subseteq E$. We show that $r^{du}(D) \subseteq r^{du}(E)$.

Let $D' := r^{du}(D)$ and $E' := r^{du}(E)$. Take some player i and history $h \in H$. We will show that $D_i'(h) \subseteq E_i'(h)$. Consider some $a_i \in D_i'(h)$. As $D' = r^{du}(D)$, it follows by (5.1) that $a_i \in D_i(h)$ and that a_i is not strictly dominated in $G(h, D)$. By Lemma 4.1 we then know that a_i is optimal in $G(h, D)$ for a belief

$b_i \in \Delta(D_{-i}(h))$. That is,

$$\sum_{a_{-i} \in D_{-i}(h)} b_i(a_{-i}) \cdot u_i^{D+}[h](a_i, a_{-i}) \geq \sum_{a_{-i} \in D_{-i}(h)} b_i(a_{-i}) \cdot u_i^{D-}[h](a_i', a_{-i}) \text{ for all } a_i' \in D_i(h). \tag{10.24}$$

As $D \subseteq E$ we have, by construction, that $u_i^{E+}[h](a_i, a_{-i}) \geq u_i^{D+}[h](a_i, a_{-i})$ and $u_i^{E-}[h](a_i, a_{-i}) \leq u_i^{D-}[h](a_i, a_{-i})$ for all $(a_i, a_{-i}) \in D_i(h) \times D_{-i}(h)$. Together with (10.24) and the fact that $D \subseteq E$, we conclude that the belief b_i is in $\Delta(E_{-i}(h))$, and that

$$\sum_{a_{-i} \in E_{-i}(h)} b_i(a_{-i}) \cdot u_i^{E+}[h](a_i, a_{-i}) \geq \sum_{a_{-i} \in E_{-i}(h)} b_i(a_{-i}) \cdot u_i^{E-}[h](a_i', a_{-i}) \text{ for all } a_i' \in D_i(h). \tag{10.25}$$

We now show that (10.25) holds for every $a_i' \in E_i(h)$, and not only for every $a_i' \in D_i(h)$. That is, we show that

$$\sum_{a_{-i} \in E_{-i}(h)} b_i(a_{-i}) \cdot u_i^{E+}[h](a_i, a_{-i}) \geq \sum_{a_{-i} \in E_{-i}(h)} b_i(a_{-i}) \cdot u_i^{E-}[h](a_i', a_{-i}) \text{ for all } a_i' \in E_i(h). \tag{10.26}$$

Let $a_i^* \in E_i(h)$ be such that

$$\sum_{a_{-i} \in E_{-i}(h)} b_i(a_{-i}) \cdot u_i^{E-}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in E_{-i}(h)} b_i(a_{-i}) \cdot u_i^{E-}[h](a_i', a_{-i}) \text{ for all } a_i' \in E_i(h). \tag{10.27}$$

Then, it follows in particular that

$$\sum_{a_{-i} \in E_{-i}(h)} b_i(a_{-i}) \cdot u_i^{E+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i} \in E_{-i}(h)} b_i(a_{-i}) \cdot u_i^{E-}[h](a_i', a_{-i}) \text{ for all } a_i' \in E_i(h).$$

By Lemma 4.1 we conclude that the action a_i^* is not strictly dominated in $G(h, E)$, and hence $a_i^* \in E_i'(h)$. As $E' = r^{du}(E) \subseteq D$ it follows that $a_i^* \in D_i(h)$. Together with (10.25) we then conclude that

$$\sum_{a_{-i} \in E_{-i}(h)} b_i(a_{-i}) \cdot u_i^{E+}[h](a_i, a_{-i}) \geq \sum_{a_{-i} \in E_{-i}(h)} b_i(a_{-i}) \cdot u_i^{E-}[h](a_i^*, a_{-i}).$$

Combining this inequality with (10.27) yields (10.26), which was to show.

By (10.26) and Lemma 4.1, we see that a_i is not strictly dominated in $G(h, E)$, and hence $a_i \in E_i'(h)$. As this holds for every player i, history $h \in H$, and action $a_i \in D_i'(h)$, we conclude that $D' \subseteq E'$, and thus $r^{du}(D) \subseteq r^{du}(E)$. As such, r^{du} is monotone. This completes the proof. ■

10.3 Proof of Section 7

Proof of Proposition 7.1. For every $n \in \mathbb{N}$ with $n \geq 2$, $\delta \in (0,1]$, $m \in \{1,...,n\}$ and player j, let $\alpha_j^m(n,\delta)$ be the α_j^m generated by the double-utility procedure for finitely repeated games at stage m of the repeated game $\Gamma(G,n,\delta)$, and let $A^{*m}(n,\delta) \subseteq \times_{j\in I} A_j$ be the set of action profiles that survive the procedure at stage m of the repeated game $\Gamma(G,n,\delta)$. By construction we have, for a given player i, that $\alpha_i^n(n,\delta) = 0$, and for every stage $m \in \{1,...,n-1\}$ it holds that

$$\alpha_i^m(n,\delta) = \delta \cdot \left(\max_{a\in A^{*m+1}(n,\delta)} v_i(a) - \min_{a\in A^{*m+1}(n,\delta)} v_i(a) + \alpha_i^{m+1}(n,\delta) \right). \tag{10.28}$$

Set

$$\alpha_i^* := \delta \cdot \left(\max_{a\in A^*} v_i(a) - \min_{a\in A^*} v_i(a) \right).$$

Since, by assumption, there are $a, a' \in A^*$ with $v_i(a) \neq v_i(a')$, it follows that $\alpha_i^* > 0$.

As $A^{*n}(n,\delta) = A^*$, we know from (10.28) that $\alpha_i^{n-1}(n,\delta) = \delta \cdot \alpha_i^*$. Moreover, since by Observation 7.1, $A^* = A^{*n}(n,\delta) \subseteq A^{*m}(n,\delta)$ for every $m \in \{1,...,n\}$, we conclude by (10.28) that for every $m \in \{1,...,n-1\}$,

$$\begin{aligned} \alpha_i^m(n,\delta) &\geq \delta \cdot \left(\max_{a\in A^*(n,\delta)} v_i(a) - \min_{a\in A^*(n,\delta)} v_i(a) + \alpha_i^{m+1}(n,\delta) \right) \\ &= \alpha_i^{n-1}(n,\delta) + \delta \cdot \alpha_i^{m+1}(n,\delta) = \delta \cdot \alpha_i^* + \delta \cdot \alpha_i^{m+1}(n,\delta). \end{aligned}$$

This, in turn, implies that

$$\alpha_i^1(n,\delta) \geq \alpha_i^* \cdot \left(\delta + \delta^2 + ... + \delta^{n-1} \right). \tag{10.29}$$

Now, let $\Delta_i := \max_{a\in A} v_i(a) - \min_{a\in A} v_i(a)$. Recall that $\alpha_i^* > 0$. Note that $\delta + \delta^2 + ... + \delta^{n-1}$ can be made arbitrarily large by choosing n large enough and δ sufficiently close to 1, and that this expression is increasing in δ and n. Hence, there are $n^* \in \mathbb{N}$ and $\delta^* \in [0,1)$ such that for every $n \in \mathbb{N}$ with $n \geq n^*$, and every $\delta \in [\delta^*, 1]$, we have that

$$\alpha_i^* \cdot \left(\delta + \delta^2 + ... + \delta^{n-1} \right) \geq \Delta_i.$$

Together with (10.29) we conclude, for every $n \in \mathbb{N}$ with $n \geq n^*$, and every $\delta \in [\delta^*, 1]$, that $\alpha_i^1(n,\delta) \geq \Delta_i$.

But then, by the choice of Δ_i, every action in A_i will survive the iterated elimination of actions that are strictly dominated by more than $(\alpha_j^1(n,\delta))_{j\in I}$, for all $n \geq n^*$ and $\delta \in (\delta^*, 1]$. As such, every action in A_i will survive the double-utility procedure for finitely repeated games at stage 1 of the repeated game $\Gamma(G,n,\delta)$, for all $n \geq n^*$ and $\delta \in [\delta^*, 1]$. Hence, by Corollary 7.1 and Theorem 4.1, we have for every $n \geq n^*$ and $\delta \in [\delta^*, 1]$ that player i can rationally play every action at stage 1 of the repeated game $\Gamma(G,n,\delta)$ under common belief in future action-based rationality. ■

10.4 Proof of Section 8

For the proof of Theorem 8.1 we rely on a result in Perea (2014). To formally state it we need the following definition. Consider a history $h \in H$, a player i, a strategy $s_i \in S_i(h)$, and a belief $b_i \in \Delta(S_{-i}(h))$ about the opponents' strategies. The strategy s_i is *optimal* in $S_i(h)$ for the belief b_i if

$$\sum_{s_{-i} \in S_{-i}} b_i(s_{-i}) \cdot u_i(z(s_i, s_{-i})) \geq \sum_{s_{-i} \in S_{-i}} b_i(s_{-i}) \cdot u_i(z(s_i', s_{-i}))$$

for all $s_i' \in S_i(h)$.

Lemma 10.2 (Optimality of strategies in backward dominance procedure) *For a given history $h \in H$ and $k \in \mathbb{N}$, let $(S_i^k(h))_{i \in I}$ be the decision problem at h in round k of the backward dominance procedure. Then, for every player i, strategy $s_i \in S_i^k(h)$, and history $h' \in H(s_i)$ with $h' \succcurlyeq h$, there is a belief $b_i \in \Delta(S_{-i}^{k-1}(h'))$ such that s_i is optimal in $S_i(h')$ for b_i.*

The proof can be found in Perea (2014), proof of Lemma 9.4.

Proof of Theorem 8.1. For every history $h \in H$ and $k \in \mathbb{N}_0$, let $(S_i^k(h))_{i \in I}$ be the decision problem at h in round k of the backward dominance procedure, and $G^k(h) = (A_i^k(h), u_i^{k+}[h], u_i^{k-}[h])_{i \in I}$ the double-utility game at h in round k of the double-utility procedure. We start by showing the following result.

Claim. For every player i, history $h \in H$, number $k \in \mathbb{N}_0$ and strategy $s_i \in S_i^k(h)$, we have that $s_i(h) \in A_i^k(h)$.

Proof of claim. We show the result by induction on k. For $k = 0$ the statement is trivially true as $A_i^0(h) = A_i(h)$.

Take now some $k \in \mathbb{N}$ and suppose that the statement is true for $k - 1$. Consider, for a given player i and history $h \in H$, a strategy $s_i^* \in S_i^k(h)$. Then, $s_i^* \in S_i^{k-1}(h)$ and, by Lemma 10.2, there is a belief $b_i \in \Delta(S_{-i}^{k-1}(h))$ such that s_i is optimal in $S_i(h)$ for b_i. That is,

$$\sum_{s_{-i} \in S_{-i}^{k-1}(h)} b_i(s_{-i}) \cdot u_i(z(s_i^*, s_{-i})) \geq \sum_{s_{-i} \in S_{-i}^{k-1}(h)} b_i(s_{-i}) \cdot u_i(z(s_i, s_{-i})) \text{ for all } s_i \in S_i(h). \quad (10.30)$$

Let $a_i^* := s_i^*(h)$. As $s_i^* \in S_i^{k-1}(h)$ we know, by the induction assumption, that $a_i^* \in A_i^{k-1}(h)$. For every $j \neq i$ and $a_j \in A_j(h)$, define the set of strategies $S_j(h, a_j) := \{s_j \in S_j(h) \mid s_j(h) = a_j\}$. Define the belief $\beta_i \in \Delta(A_{-i}(h))$ by

$$\beta_i((a_j)_{j \neq i}) := b_i(\times_{j \neq i} S_j(h, a_j)) \quad (10.31)$$

for all $(a_j)_{j \neq i} \in A_{-i}(h)$. We will show that (a) $\beta_i \in \Delta(A_{-i}^{k-1}(h))$, and that (b) a_i^* is optimal in the double-utility game $G^{k-1}(h)$ for the belief β_i.

(a) To see that $\beta_i \in \Delta(A_{-i}^{k-1}(h))$, suppose that $\beta_i((a_j)_{j \neq i}) > 0$ for some $(a_j)_{j \neq i} \in A_{-i}(h)$. By (10.31) there must be some $(s_j)_{j \neq i} \in \times_{j \neq i} S_j(h, a_j)$ with $b_i((s_j)_{j \neq i}) > 0$. As $b_i \in \Delta(S_{-i}^{k-1}(h))$ we must have that

$s_j \in S_j^{k-1}(h)$ for all $j \neq i$. By the induction assumption we thus have that $a_j = s_j(h) \in A_j^{k-1}(h)$ for all $j \neq i$. As such, $\beta_i \in \Delta(A_{-i}^{k-1}(h))$.

(b) We next show that a_i^* is optimal in the double-utility game $G^{k-1}(h)$ for the belief β_i. Let $S_i^{k-1}(\succ h)$ be the set of strategies $s_i \in S_i(h)$ such that $s_i \in S_i^{k-1}(h')$ for all $h' \in H(s_i)$ with $h' \succ h$. Note that $S_i^{k-1}(\succ h)$ is non-empty. Indeed, take a strategy $s_i \in S_i^{k-1}(h)$. Then, by definition of the sets $S_i^{k-1}(h'')$, for every $h' \in H(s_i)$ with $h' \succ h$ we have that $s_i \in S_i^{k-1}(h')$ also, and hence $s_i \in S_i^{k-1}(\succ h)$.

Consider some $s_{-i} = (s_j)_{j \neq i}$ with $b_i(s_{-i}) > 0$, let $a_j := s_j(h)$ for all $j \neq i$, and set $a_{-i} := (a_j)_{j \neq i}$. Also, take some $s_i \in S_i^{k-1}(\succ h)$ and some history $h' \in H$ with $z(s_i, s_{-i}) \succcurlyeq h' \succcurlyeq (h, (s_i(h), a_{-i}))$. We will show that (i) $s_i(h') \in A_i^{k-1}(h')$ and (ii) $s_j(h') \in A_j^{k-1}(h')$ for all $j \neq i$.

(i) Since $z(s_i, s_{-i}) \succcurlyeq h'$ we conclude that $h' \in H(s_i)$. As $h' \succ h$ and $s_i \in S_i^{k-1}(\succ h)$, it follows that $s_i \in S_i^{k-1}(h')$. By the induction assumption we conclude that $s_i(h') \in A_i^{k-1}(h')$.

(ii) Take some $j \neq i$. As $z(s_i, s_{-i}) \succcurlyeq h'$ we conclude that $h' \in H(s_j)$. Since $b_i(s_{-i}) > 0$ and $b_i \in \Delta(S_{-i}^{k-1}(h))$ it follows that $s_j \in S_j^{k-1}(h)$. As $h' \in H(s_j)$ and $h' \succ h$, it follows by the definitions of $S_j^{k-1}(h)$ and $S_j^{k-1}(h')$ that $s_j \in S_j^{k-1}(h')$ also. By the induction assumption we conclude that $s_j(h') \in A_j^{k-1}(h')$.

Since (i) and (ii) hold for every history h' with $z(s_i, s_{-i}) \succcurlyeq h' \succcurlyeq (h, (a_i, a_{-i}))$, it follows that

$$u_i^{k-1-}[h](a_i, a_{-i}) \leq u_i(z(s_i, s_{-i})) \leq u_i^{k-1+}[h](a_i, a_{-i}). \tag{10.32}$$

Recall that $a_i^* = s_i^*(h)$ with $s_i^* \in S_i^{k-1}(h)$. Since $S_i^{k-1}(h) \subseteq S_i^{k-1}(\succ h)$ it follows that $s_i^* \in S_i^{k-1}(\succ h)$. As (10.32) applies to every $s_i \in S_i^{k-1}(\succ h)$ and every $s_{-i} \in S_{-i}^{k-1}(h)$ with $b_i(s_{-i}) > 0$, and $\beta_i \in \Delta(A_{-i}^{k-1}(h))$, we conclude on the basis of (10.31) that

$$\sum_{s_{-i} \in S_{-i}^{k-1}(h)} b_i(s_{-i}) \cdot u_i(z(s_i^*, s_{-i})) \leq \sum_{a_{-i} \in A_{-i}^{k-1}(h)} \beta_i(a_{-i}) \cdot u_i^{k-1+}[h](a_i^*, a_{-i}) \tag{10.33}$$

and

$$\sum_{s_{-i} \in S_{-i}^{k-1}(h)} b_i(s_{-i}) \cdot u_i(z(s_i, s_{-i})) \geq \sum_{a_{-i} \in A_{-i}^{k-1}(h)} \beta_i(a_{-i}) \cdot u_i^{k-1-}[h](a_i, a_{-i}) \tag{10.34}$$

for every $s_i \in S_i^{k-1}(\succ h)$ with $s_i(h) = a_i$.

Note that for every $a_i \in A_i^{k-1}(h)$ there is some $s_i \in S_i^{k-1}(\succ h)$ with $s_i(h) = a_i$. To see this, let

$$H^{first}(h) := \{h' \in H \mid h' \succ h \text{ and } \nexists h'' \in H \text{ s.t. } h' \succ h'' \succ h\}$$

be the first non-terminal histories that follow h. For every $h' \in H^{first}(h)$ take some $\hat{s}_i[h'] \in S_i^{k-1}(h')$, and let $s_i \in S_i(h)$ be a strategy such that $s_i(h) = a_i$, and $s_i(h'') = \hat{s}_i[h'](h'')$ for all $h' \in H^{first}(h)$ and all $h'' \succcurlyeq h'$. As, by definition, $S_i^{k-1}(h') \cap S_i(h'') \subseteq S_i^{k-1}(h'')$ for all $h' \in H^{first}(h)$ and all $h'' \succcurlyeq h'$ it follows that $s_i \in S_i^{k-1}(h'')$ for all $h' \in H^{first}(h)$ and all $h'' \succcurlyeq h'$. Hence, $s_i(h) = a_i$ and $s_i \in S_i^{k-1}(\succ h)$. Therefore, it is indeed the case that for every $a_i \in A_i^{k-1}(h)$ there is some $s_i \in S_i^{k-1}(\succ h)$ with $s_i(h) = a_i$.

Based on this insight, it follows from (10.30), (10.33) and (10.34) that

$$\sum_{a_{-i}\in A_{-i}^{k-1}(h)} \beta_i(a_{-i}) \cdot u_i^{k-1+}[h](a_i^*, a_{-i}) \geq \sum_{a_{-i}\in A_{-i}^{k-1}(h)} \beta_i(a_{-i}) \cdot u_i^{k-1-}[h](a_i, a_{-i})$$

for all $a_i \in A_i^{k-1}(h)$. That is, $a_i^* \in A_i^{k-1}(h)$ is optimal for the belief $\beta_i \in \Delta(A_{-i}^{k-1}(h))$ in the double-utility game $G^{k-1}(h)$. This completes the proof of (b).

By Lemma 4.1 we conclude that $a_i^* \in A_i^{k-1}(h)$ is not strictly dominated in the double-utility game $G^{k-1}(h)$. Hence, $a_i^* \in A_i^k(h)$. Thus, we have shown that $s_i^*(h) = a_i^* \in A_i^k(h)$ for every history h, every player i, and every strategy $s_i^* \in S_i^k(h)$. By induction on k, the proof of the claim is complete. $\Diamond$

To prove the theorem, choose a player i and a strategy s_i that survives the backward dominance procedure. Then, $s_i \in S_i^k(h)$ for all $h \in H(s_i)$ and all $k \in \mathbb{N}$. By the claim, we then conclude for every history $h \in H(s_i)$ that $s_i(h) \in A_i^k(h)$ for all k. That is, for every history $h \in H(s_i)$, the action $s_i(h)$ survives the double-utility procedure. This completes the proof. ■

References

[1] Asheim, G.B. and A. Perea (2005), Sequential and quasi-perfect rationalizability in extensive games, *Games and Economic Behavior* **53**, 15–42.

[2] Battigalli, P. (1997), On rationalizability in extensive games, *Journal of Economic Theory* **74,** 40–61.

[3] Battigalli, P., Catonini, E. and J. Manili (2023), Belief change, rationality, and strategic reasoning in sequential games, *Games and Economic Behavior* **142,** 527–551.

[4] Battigalli, P. and N. De Vito (2021), Beliefs, plans, and perceived intentions in dynamic games, *Journal of Economic Theory* **195,** https://doi.org/10.1016/j.jet.2021.105283.

[5] Battigalli, P., Di Tillio, A. and D. Samet (2013), Strategies and interactive beliefs in dynamic games, in D. Acemoglu, M. Arellano and E. Dekel (eds.), *Advances in Economics and Econometrics* (Cambridge University Press), pp. 391–422.

[6] Battigalli, P. and M. Siniscalchi (1999), Hierarchies of conditional beliefs and interactive epistemology in dynamic games, *Journal of Economic Theory* **88,** 188–230.

[7] Catonini, E. and A. Penta (2025), Backward induction reasoning beyond backward induction, Forthcoming in *American Economic Journal: Microeconomics.*

[8] Dekel, E., Fudenberg, D. and D.K. Levine (1999), Payoff information and self-confirming equilibrium, *Journal of Economic Theory* **89**, 165–185.

[9] Dekel, E., Fudenberg, D. and D.K. Levine (2002), Subjective uncertainty over behavior strategies: A correction, *Journal of Economic Theory* **104,** 473–478.

[10] Kohlberg, E. and J.-F. Mertens (1986), On the strategic stability of equilibria, *Econometrica* **54,** 1003–1037.

[11] Kreps, D.M. and R. Wilson (1982), Sequential equilibria, *Econometrica* **50**, 863–894.

[12] Li, S. (2017), Obviously strategy-proof mechanisms, *Americal Economic Review* **107,** 3257–3287.

[13] Meier, M. and A. Perea (2023), Forward induction in a backward inductive manner, *Epicenter Working Paper No. 27.*

[14] Meier, M. and A. Perea (2025), Reasoning about your own future mistakes, *Games and Economic Behavior* **154,** 302–328.

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

[16] Penta, A. (2015), Robust dynamic implementation, *Journal of Economic Theory* **160,** 280–316.

[17] Perea, A. (2014), Belief in the opponents' future rationality, *Games and Economic Behavior* **83,** 231–254.

[18] Perea, A. (2025), *From Decision Theory to Game Theory*: *Reasoning about the Decisions of Others,* Cambridge University Press.

[19] Robinson, A. (1973), Function theory on some nonarchimedean fields, *The American Mathematical Monthly* **80,** 87–109.

[20] Selten, R. (1965), Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit, *Zeitschrift für die Gesammte Staatswissenschaft* **121,** 301–324, 667–689.