Dynamic Unawareness and Rationalizable Behavior

Heifetz, Aviad and Meier, Martin and Schipper, Burkhard C

The Open University of Israel, Institute for Economic Analysis, Barcelona, University of California, Davis

4 May 2009

Dynamic Unawareness and Rationalizable Behavior*

Aviad Heifetz† Martin Meier‡ Burkhard C. Schipper§

May 4, 2009

Abstract

We define generalized extensive-form games which allow for mutual unawareness of actions. We extend Pearce's (1984) notion of extensive-form (correlated) rationalizability to this setting, explore its properties and prove existence. We define also a new variant of this solution concept, prudent rationalizability, which refines the set of outcomes induced by extensive-form rationalizable strategies. Finally, we define the normal form of a generalized extensive-form game, and characterize in it extensive-form rationalizability by iterative conditional dominance.

Keywords: Unawareness, extensive-form games, extensive-form rationalizability, prudent rationalizability, iterative conditional dominance.

JEL-Classifications: C70, C72, D80, D82.

*We are grateful to Pierpaolo Battigalli for numerous insightful comments andsuggestions, as well as to seminar participants at Barcelona, Bocconi, Caltech, Maryland, Maastricht, Pittsburgh, USC, UC Davis, Tel Aviv, Stony Brook 2007, LOFT 2008 and Games 2008. Aviad is grateful for financial support from the Open University of Israel's Research Fund grant no. 46106. Martin was supported by the Spanish Ministerio de Educación y Ciencia via a Ramon y Cajal Fellowship (IAE-CSIC) and a Research Grant (SEJ 2006-02079), and Barcelona Graduate School of Economics. Burkhard is grateful for financial support from the NSF SES-0647811.

†The Economics and Management Department, The Open University of Israel. Email: aviadhe@openu.ac.il

‡Instituto de Análisis Económico - CSIC, Barcelona, and Institut für Höhere Studien, Wien. Email: martin.meier@iae.csic.es

§Department of Economics, University of California, Davis. Email: bcschipper@ucdavis.edu

1 Introduction

In real-life dynamic interactions, unawareness of players regarding the relevant actions available to them is at least as prevalent as uncertainty regarding other players' strategies, payoffs or moves of nature. Players frequently become aware of actions they (or other players) could have taken in retrospect, when they can only re-evaluate the past actions chosen by partners or rivals who *were* aware of those actions from the start, and hence re-assess their likely future behavior. Yet, while uncertainty can be captured within the standard framework of extensive-form games with imperfect information, unawareness and mutual uncertainty regarding awareness require an extension of this framework. Such an extension is the first task of the current paper.

At first, one may wonder why the standard framework would not suffice. After all, if a player is unaware of an action which is actually available to her, then for all practical purposes she cannot choose it. Why wouldn't it be enough simply to truncate from the tree all the paths starting with such an action?

The reason is that the strategic implications of unawareness of an action are distinct from the unavailability of the same action. To see this, consider the following standard "battle-of-the-sexes"game (where Bach and Stravinsky concerts are the two available choices for each player)

		II	
		B	S
I	B	3, 1	0, 0
	S	0, 0	1, 3

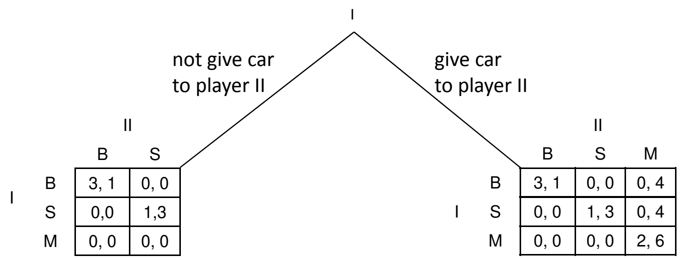
augmented by a dominant Mozart concert for player II:

		II		
		B	S	M
	B	3, 1	0, 0	0, 4
I	S	0, 0	1, 3	0, 4
	M	0, 0	0, 0	2, 6

The new game is dominance solvable, and (M,M) is the unique Nash equilibrium.

Suppose that the Mozart concert is in a distant town, and II can go there only if player I gives him her car in the first place: Here, if player I doesn't give the car to player II,

Figure 1:



player II may conclude by forward induction that player I would go to the Bach concert with the hope of getting the payoff 3 (because by giving the car to II, player I could have achieved the payoff 2). The best reply of player II is to follow suit and attend the Bach concert as well. Hence, in the unique rationalizable outcome, player I is not to give the car to player II and to go to the Bach concert.[1]

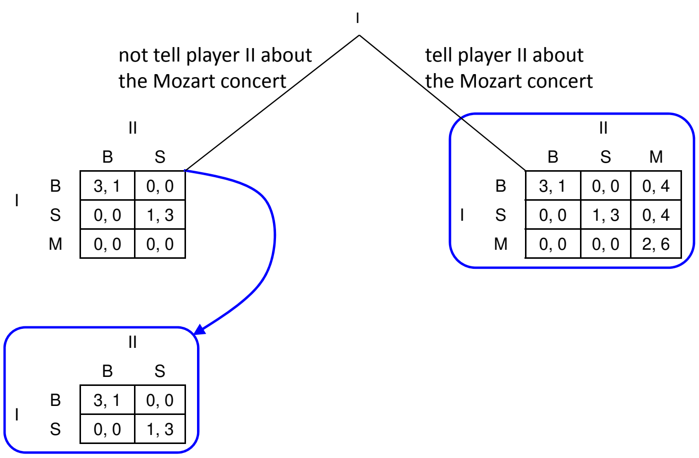
But what if, instead, the Mozart concert is in town but player II is initially unaware of the Mozart concert, while player I can enable player II to go to the concert simply by telling him about it? If player II remains unaware of the Mozart concert, then neither does he conceive that player I could have told him about the Mozart concert, and in particular he cannot carry out any forward-induction calculation. For him, the game is a standard battle-of-the-sexes game, where both actions of player I are rationalizable. This strategic situation is depicted in Figure 2.

The strategic situation cannot be described by a standard extensive-form game. If player I chooses not to tell player II about the Mozart concert, then player II's information set (depicted in blue) consists of a node in a simpler game –namely the one-shot battle-of-the-sexes with no preceding move by player I.

This is a simple example of the general novel framework that we define in Section 2 for dynamic interaction with possibly mutual unawareness of actions, generalizing standard extensive-form games. The framework will not only allow modeling of situations in which one player is certain that another player is unaware of portions of the game tree, as in the above example, but also of situations in which a player is uncertain regarding the way another player views the game tree, as well as situations in which the player is uncertain

[1]For a discussion of forward induction in battle-of-the-sexes games see van Damme (1989).

Figure 2:



regarding the uncertainties of the other player about yet other players' views of the game tree, and so forth.

In fact, this framework allows not just for unawareness but also for other forms of misconception about the structure of the game. Section 6 specifies further properties obtaining in generalized extensive-form games where the only source of players 'misconception' is unawareness and mutual unawareness of available actions and paths in the game. Since we focus on this type of unawareness, most of the examples in the paper satisfy the further properties specified in Section 6. Nevertheless, modeling *awareness of unawareness* does require the general framework in Section 2, as explained at its end.

In this new framework, for each information set of a player her strategy specifies – from the point of view of the modeler – what the player would do if and when that information set of hers is ever reached. In this sense, a player does not necessarily 'own' her full strategy at the beginning of the game, because she might not be initially aware of all of her information sets. That's why a sensible generalization of Pearce's (1984) notion of extensive-form rationalizability is non-trivial.

In Section 3 we put forward a modified definition, prove existence, and show the sense in which it coincides with extensive-form rationalizability in standard extensive-form games.

We focus here on a rationalizability solution concept rather than on some notion of equilibrium. While an equilibrium is ideally interpreted as a rest-point of some dynamic learning or adaptation process, or alternatively as a pre-meditated agreement or expectation, we find it difficult to carry over such interpretations to a setting in which every increase of awareness is by definition a shock or a surprise. Once a player's view of the game itself is challenged in the course of play, it is hard to justify the idea that a convention or an agreement for the continuation of the game are readily available.

We chose to focus on extensive-form rationalizability because it embodies forward induction reasoning. If an opponent makes a player aware of some relevant aspect of reality, it is implausible to dismiss the increased level of awareness as an unintended consequence of the opponent's behavior. Rather, the player should try to rationalize the opponent's choice, re-interpret the opponent's past behavior, and try to infer from it the opponent's future moves. Extensive-form rationalizability indeed captures a 'best rationalization principle' (Battigalli 1997).

In Section 4 we introduce a related solution concept, prudent rationalizability, which is the direct generalization of iterated admissibility to dynamic games with unawareness. Unlike in normal-form games, this generalization is surprisingly *not* always a refinement of extensive-form rationalizability (even for standard extensive-form games). However, we prove that prudent rationalizable strategies do refine the set of *outcomes* obtainable by extensive-form rationalizable strategies. We show how prudent rationalizability is effective in ruling out less plausible rationalizable outcomes in examples due to Pearce (1984) and Ozbay (2007).

In standard game theory, the extensive form has been considered as a more complete description of the strategic situation than the normal form. This has been questioned by Kohlberg and Mertens (1986) who argued that the normal form contains all strategically relevant information. For standard extensive-form games, Shimoji and Watson (1998) showed how extensive-form reasoning embodied in extensive-form rationalizability can be carried out in the normal form. Arguably generalized extensive-form games contain more "time relevant" structure than standard extensive-form games since they also formalize changes in the awareness of players. It is therefore an intriguing question whether a solution to generalized extensive-form games can be found when the analysis is carried out in the appropriately defined normal form associated to a generalized extensive-form game. In Section 5 we define the normal form associated to general extensive-form games. We extend Shimoji and Watson' characterization of extensive-form rationalizability by iterated conditional strict dominance to games with unawareness. In some applications, it may be more practical to apply iterated conditional strict dominance in the normal form rather than extensive-form rationalizability.

Our framework for dynamic interaction under unawareness seems to be simpler than the one proposed by Halpern and Rêgo (2006) and Rêgo and Halpern (2007), in which they investigated the notions of Nash and sequential equilibrium, respectively. Feinberg (2004) addressed dynamic unawareness with a syntactic approach, while we propose a set-theoretic approach which generalizes directly the standard framework of extensive-form games. Li (2006) considered dynamic unawareness with perfect information, while our framework allows for both unawareness and imperfect information.

Ozbay (2007) studies sender-receiver games, in which an 'announcer' can make an unaware decision maker aware of more states of nature before the decision maker takes an action. Such games can also be naturally formulated as a particular instance of our framework. For these games Ozbay studies an equilibrium notion incorporating forward-induction reasoning. Filiz-Ozbay (2007) studies a related setting in which the aware announcer is a risk neutral insurer, while the decision maker is a risk averse or ambiguity averse insuree. At equilibrium, the insurer does not always reveal all relevant contingencies to the insuree.[2]

2 Generalized extensive-form games

To define a generalized extensive-form game Γ, consider first, as a building block, a finite perfect information game with a set of players I, a set of decision nodes N_0, active players I_n at node n with finite action sets A_n^i of player $i \in I_n$ (for $n \in N_0$), chance nodes C_0, and terminal nodes Z_0 with a payoff vector $(p_i^z)_{i \in I} \in \mathbb{R}^I$ for the players for every $z \in Z_0$. The nodes $\bar{N}_0 = N_0 \cup C_0 \cup Z_0$ constitute a tree.

2.1 Partially ordered set of trees

Consider now a family $\mathbf{T}$ of subtrees of $\bar{N}_0$, partially ordered ($\preceq$) by inclusion. One of the trees $T_1 \in \mathbf{T}$ is meant to represent the modeler's view of the paths of play that are *objectively* feasible; each other tree represents the feasible paths of play as *subjectively* viewed by some player at some node at one of the trees.

In each tree $T \in \mathbf{T}$ denote by n_T the copy in T of the node $n \in N_0$ whenever the copy

[2]Currently we are unaware of further papers focusing directly and explicitly on dynamic games with unawareness. The literature on unawareness in general is growing fast – see e.g. http://www.econ.ucdavis.edu/faculty/schipper/unaw.htm

of n is part of the tree T. However, in what follows we will typically avoid the subscript T when no confusion may arise.

Denote by N_i^T the set of nodes in which player $i \in I$ is active in the tree $T \in \mathbf{T}$.

We require two properties:

1. All the terminal nodes in each tree $T \in \mathbf{T}$ are copies of nodes in Z_0.

2. If for two decision nodes $n, n' \in N_i^T$ (i.e. $i \in I_n \cap I_{n'}$) it is the case that $A_n^i \cap A_{n'}^i \neq \emptyset$, then $A_n^i = A_{n'}^i$.[3]

Property 1 is needed to ensure that each terminal node of each tree $T \in \mathbf{T}$ is associated with well defined payoffs to the players. Property 2 means that i's active nodes N_i^T are partitioned into equivalence classes, such that the actions available to player i are identical within each equivalence class and disjoint in distinct equivalence classes. It will be needed for the definition of information sets which follows shortly.

Denote by N the union of all decision nodes in all trees $T \in \mathbf{T}$, by C the union of all chance nodes, by Z the union of terminal nodes, and by $\bar{N} = N \cup C \cup Z$. For a node $n \in \bar{N}$ we denote by T_n the tree containing n.

2.2 Information sets

Next, in each decision node $n \in N$, define for each active player $i \in I_n$ an information set $\pi_i(n)$ with the following properties:

- I0 Confinement: $\pi_i(n) \subseteq T$ for some tree T.

- I1 No delusion: If $\pi_i(n) \subseteq T_n$ then $n \in \pi_i(n)$.

- I2 Introspection: If $n' \in \pi_i(n)$ then $\pi_i(n') = \pi_i(n)$.

- I3 No divining of currently unimaginable paths, no expectation to forget currently conceivable paths: If $n' \in \pi_i(n) \subseteq T'$ (where $T' \in \mathbf{T}$ is a tree) and there is a path $n', \ldots, n'' \in T'$ such that $i \in I_{n'} \cap I_{n''}$ then $\pi_i(n'') \subseteq T'$.

[3] Sometimes the modeler may want to impose an additional property: If in the original tree the probabilities of reaching $\bar{n}_1, \ldots \bar{n}_k \in \bar{N}$ from the chance node $c \in C$ are $p_c^{\bar{n}_1} > 0, \ldots, p_c^{\bar{n}_k} > 0$ but some of these nodes do not appear in the subtree, then the probabilities of reaching the remaining nodes emanating from c are renormalized so as to sum to 1 in the subtree. We do not impose this property here since it may be natural in some contexts but unnatural in others.

I4 No imaginary actions: If $n' \in \pi_i(n)$ then $A^i_{n'} \subseteq A^i_n$.

I5 Distinct action names in disjoint information sets: For a subtree T, if $n, n' \in T$ and $A^i_n = A^i_{n'}$ then $\pi_i(n') = \pi_i(n)$.

I6 Perfect recall: Suppose that player i is active in two distinct nodes n_1 and n_k, and there is a path $n_1, n_2, ..., n_k$ such that at n_1 player i takes the action a_i. If $n' \in \pi_i(n_k)$, then there exists a node $n'_1 \neq n'$ and a path $n'_1, n'_2, ..., n'_\ell = n'$ such that $\pi_i(n'_1) = \pi_i(n_1)$ and at n'_1 player i takes the action a_i.

The following figures (Figure 3) illustrate properties I0 to I6.

Properties (I1), (I2), (I4), and (I5) are standard for extensive-form games, and properties (I0) and (I6) generalize other standard properties of extensive-form games to our generalized setting. The essentially new property is (I3). At each information set of a player, property (I3) confines the player's *anticipation* of her future view of the game to the view she currently holds (even if, as a matter of fact, this anticipation is about to be shuttered as the game evolves).

We denote by H_i the set of i's information sets in all trees. For an information set $h_i \in H_i$, we denote by T_{h_i} the tree containing h_i. For two information sets h_i, h'_i in a given tree T, we say that h_i precedes h'_i (or that h'_i succeeds h_i) if for every $n' \in h'_i$ there is a path $n, ..., n'$ such that $n \in h_i$. We denote $h_i \rightsquigarrow h'_i$.

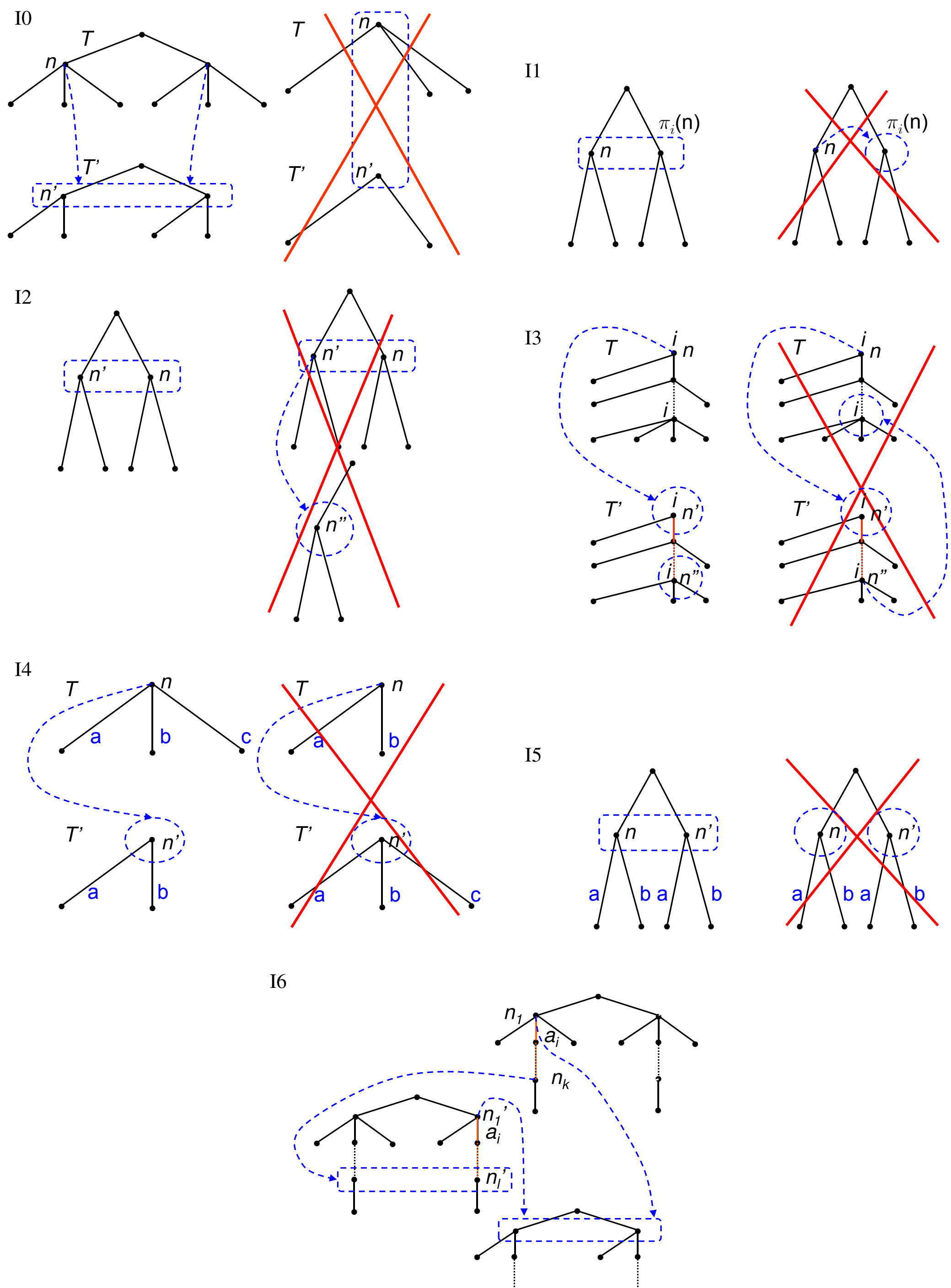
Remark 1 *The following property is implied by I2 and I4: If $n', n'' \in h_i$ where $h_i = \pi_i(n)$ is an information set, then $A^i_{n'} = A^i_{n''}$.*

Proof. If $n', n'' \in h_i$ where $h_i = \pi_i(n)$ is some information set, then by introspection (I3) we must have $\pi_i(n') = \pi_i(n'') = \pi_i(n)$. Hence by (I4) $A^i_{n'} \subseteq A^i_{n''}$ and $A^i_{n''} \subseteq A^i_{n'}$. □

Remark 2 *Properties I0, I1, I2 and I6 imply no absent-mindedness: No information set h_i contains two distinct nodes n, n' on some path in some tree.*

Proof. Suppose by contradiction that there exists an information set h_i with a node $n \in h_i$ such that some other node in h_i precedes n in the tree T_n. Denote by n' the first node on the path from the root to n that is also in h_i. Now apply *I*6 with $n'_l := n'$ to get a path $n'' = n'_1, ..., n'_l = n'$, with $\pi_i(n'') = \pi_i(n_1) = \pi_i(n') = h_i$. By *I*1, we have $n'' \in h_i$ and n'' is a predecessor of n', a contradiction. □

Figure 3: Properties I0 to I6



The perfect recall property I6 and Remark 2 guarantee that with the precedence relation $\rightsquigarrow$ player i's information sets H_i form an *arborescence*: For every information set $h'_i \in H_i$, the information sets preceding it $\{h_i \in H_i : h_i \rightsquigarrow h'_i\}$ are totally ordered by $\rightsquigarrow$.

For trees $T, T' \in \mathbf{T}$ we denote $T \rightarrowtail T'$ whenever for some node $n \in T$ and some player $i \in I_n$ it is the case that $\pi_i(n) \subseteq T'$. Denote by $\hookrightarrow$ the transitive closure of $\rightarrowtail$. That is, $T \hookrightarrow T''$ iff there is a sequence of trees $T, T', \ldots, T'' \in \mathbf{T}$ satisfying $T \rightarrowtail T' \rightarrowtail \cdots \rightarrowtail T''$.

2.3 Generalized games

A generalized extensive-form game Γ consists of a partially ordered set $\mathbf{T}$ of subtrees of a tree $\bar{N}_0$ satisfying properties 1-2 above, along with information sets $\pi_i(n)$ for every $n \in T$, $T \in \mathbf{T}$ and $i \in I_n$, satisfying properties I0-I6 above.

For every tree $T \in \mathbf{T}$, the T*-partial game* is the partially ordered set of trees including T and all trees T' in Γ satisfying $T \hookrightarrow T'$, with information sets as defined in Γ. A T-partial game is a generalized game, i.e. it satisfies all properties 1-2 and I0-I6.

We denote by H_i^T the set of i's information sets in the T-partial game.

2.4 Strategies

A (pure) strategy

$$s_i \in S_i \equiv \prod_{h_i \in H_i} A_{h_i}$$

for player i specifies an action of player i at each of her information sets $h_i \in H_i$. Denote by

$$S = \prod_{j \in I} S_j$$

the set of strategy profiles in the generalized extensive-form game.

If $s_i = (a_{h_i})_{h_i \in H_i} \in S_i$, we denote by

$$s_i(h_i) = a_{h_i}$$

the player's action at the information set h_i. If player i is active at node n, we say that at node n the strategy prescribes to her the action $s_i(\pi_i(n))$.

In generalized extensive-form games, a strategy cannot be conceived as an ex ante plan of action. If $h_i \subseteq T$ but $T \not\hookrightarrow T'$, then at h_i player i may be interpreted as being unaware of her information sets in $H_i^{T'} \setminus H_i^T$.

Thus, a strategy of player i should rather be viewed as a list of answers to the hypothetical questions "what would the player do if h_i were the set of nodes she considered as possible?", for $h_i \in H_i$. However, there is no guarantee that such a question about the information set $h'_i \in H_i^{T'}$ would even be meaningful to the player if it were asked at a different information set $h_i \in H_i^T$ when $T \not\hookrightarrow T'$. The answer should therefore be interpreted as given by the modeler, as part of the description of the situation.

For a strategy $s_i \in S_i$ and a tree $T \in \mathbf{T}$, we denote by s_i^T the strategy in the T-partial game induced by s_i. If $R_i \subseteq S_i$ is a set of strategies of player i, denote by R_i^T the set of strategies induced by R_i in the T-partial game, The set of i's strategies in the T-partial game is thus denoted by S_i^T. Denote by $S^T = \prod_{j \in I} S_j^T$ the set of strategy profiles in the T-partial game.

We say that a strategy profile $s \in S$ *reaches* the information set $h_i \in H_i$ if the players' actions and nature's moves (if there are any) in T_{h_i} lead to h_i with a positive probability. (Notice that unlike in standard games, an information set $\pi_i(n)$ may be contained in tree $T' \neq T_n$. In such a case, by definition $s_i(\pi_i(n))$ induces an action to player i also in n and not only in the nodes of $\pi_i(n)$.)

We say that the strategy $s_i \in S_i$ *reaches* the information set h_i if there is a strategy profile $s_{-i} \in S_{-i}$ of the other players such that the strategy profile (s_i, s_{-i}) reaches h_i. Otherwise, we say that the information set h_i is excluded by the strategy s_i.

Similarly, we say that the strategy profile $s_{-i} \in S_{-i}$ *reaches* the information set h_i if there exists a strategy $s_i \in S_i$ such that the strategy profile (s_i, s_{-i}) reaches h_i.

A strategy profile $(s_j)_{j \in I}$ *reaches a node* $n \in T$ if the players' actions $s_j(\pi_j(n'))_{j \in I}$ and nature's moves in the nodes $n' \in T$ lead to n with a positive probability. Since we consider only finite trees, $(s_j)_{j \in I}$ reaches an information set $h_i \in H_i$ if and if there is a node $n \in h_i$ such that $(s_j)_{j \in I}$ reaches n.

As is the case also in standard games, for every given node, a given strategy profile of the players induces a distribution over terminal nodes in each tree, and hence an expected payoff for each player in the tree.

For an information set h_i, let $s_i \tilde{s}_i^{h_i}$ denote the strategy that is obtained by replacing actions prescribed by s_i at the information set h_i and its successors by actions prescribed by $\tilde{s}_i$. The strategy $s_i/\tilde{s}_i^{h_i}$ is called an h_i-replacement of s_i.

The set of *behavioral* strategies is

$$\prod_{h_i \in H_i} \Delta\left(A_{h_i}\right).$$

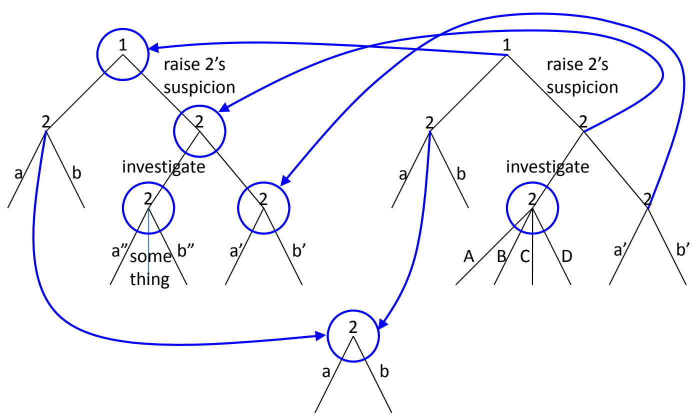
2.5 Awareness of unawareness

In some strategic situations a player may be aware of her unawareness in the sense that she is suspicious that *something* is amiss without being able to conceptualize this 'something'. Such a suspicion may affect her payoff evaluations for actions that she knows are available to her. More importantly, she may take actions to investigate her suspicion if such actions are physically available.

To model awareness of unawareness some of the trees may include *imaginary actions* as placeholders for actions that a player may be unaware of and terminal nodes/evaluations of payoffs that reflect her awareness of unawareness. (The approach of modeling awareness of unawareness by "imaginary moves" was proposed by Halpern and Rêgo, 2006.)

Consider the example in Figure 4.

Figure 4: Game form with awareness of unawareness



In both right and left trees, player 1 can decide whether or not to raise the suspicion of player 2. If he does not, then player 2 can decide between two actions. Since in this case player 2's information set is in the lower tree, she does not even realize that player 1 could have raised her suspicion. If player 1 raises player 2's suspicion, then player 2's information set is in the left tree. She must decide whether to investigate her suspicion or not. If she doesn't, then she can decide between two actions but this time she realizes that player 1 raised her suspicion (and could have refrained from doing so); and that she could have chosen to investigate, in which case she may have had 'something' else to do, that she cannot conceptualize in advance. Once she investigates, she becomes aware of two more actions and her information set is in the right tree. She also realizes that player 1 initially raised her suspicion without being explicitly aware of those actions of hers by himself. Note that before she decides whether or not to investigate, she is not modeled as anticipating to be in the right tree, because she cannot conceptualize the nature of the actions she reveals if and when she investigates.

3 Extensive-form rationalizability

Pearce (1984) defined extensive-form (correlated) rationalizable strategies by a procedure of an iterative elimination of strategies. The idea behind the definition involves a notion of forward induction. In generic perfect-information games, rationalizable strategy profiles yield the backward induction outcome, though they need not be subgame-perfect equilibrium strategies (Reny 1992, Battigalli 1997).

In what follows we extend this definition to generalized extensive-form games.

A *belief system* of player i

$$b_i = \left(b_i\left(h_i\right)\right)_{h_i \in H_i} \in \prod_{h_i \in H_i} \Delta\left(S_{-i}^{T_{h_i}}\right)$$

is a profile of beliefs - a belief $b_i\left(h_i\right) \in \Delta\left(S_{-i}^{T_{h_i}}\right)$ about the other players' strategies in the T_{h_i}-partial game, for each information set $h_i \in H_i$, with the following properties

- $b_i\left(h_i\right)$ reaches h_i, i.e. $b_i\left(h_i\right)$ assigns probability 1 to the set of strategy profiles of the other players that reach h_i.
- If h_i precedes h_i' $\left(h_i \rightsquigarrow h_i'\right)$ then $b_i\left(h_i'\right)$ is derived from $b_i\left(h_i\right)$ by Bayes rule whenever possible.

Denote by B_i the set of player i's belief systems.

For a belief system $b_i \in B_i$, a strategy $s_i \in S_i$ and an information set $h_i \in H_i$, define player i's expected payoff **at** h_i to be the expected payoff for player i in T_{h_i} given $b_i\left(h_i\right)$, the actions prescribed by s_i at h_i and its successors, and conditional on the fact that h_i has been reached.[4]

We say that with the belief system b_i and the strategy s_i player i is *rational* at the information set $h_i \in H_i$ if either s_i doesn't reach h_i in the tree T_{h_i}, or if s_i does reach h_i in the tree T_{h_i} then there exists no h_i-replacement of s_i which yields player i a higher expected payoff in T_{h_i} given the belief $b_i(h_i)$ on the other players' strategies $S_{-i}^{T_{h_i}}$.

We say that with the belief system b_i and the strategy s_i player *i would be rational* at the information set $h_i \in H_i$ if there exists no action $a'_{h_i} \in A_{h_i}$ such that only replacing the action $s_i(h_i)$ by a'_{h_i} results in a new strategy s'_i which yields player i a higher expected payoff at h_i given the belief $b_i(h_i)$ on the other players' strategies $S_{-i}^{T_{h_i}}$.

The difference between these two definitions is as follows. The definition of *rationality* of a strategy s_i at an information set h_i takes a global perspective. It is mute regarding information sets which the strategy s_i itself rules out. Also, at an information set h_i which s_i *does* reach, it considers h_i-replacements, which may alter s_i not only at h_i, but also simultaneously at h_i and/or at some of the succeeding information sets of player i.

In contrast, the second definition takes a local perspective. It takes seriously the reasoning about rationality *assuming* that h_i has been reached, whether this assumption is *realistic* (when h_i can in fact be reached with a positive probability given the actions prescribed by s_i at preceding information sets) or *counterfactual* (when h_i is ruled out by i's own actions with the strategy s_i at preceding information sets). Moreover, it considers alternative actions a'_{h_i} only at h_i itself. This is motivated by the implicit assumption that at h_i, player i is certain that at future information sets she will be acting according to the strategy s_i, but at the same time she also realizes that at each such future information set she will have the opportunity to re-consider her action, and that at h_i she has no way to commit herself to the action she will be taking at such a future information set.

We find the second definition more appealing in the context of unawareness. With unawareness, a player does not necessarily conceive of her entire strategy. Rather, she might be aware only of a subset of her information sets. She may plan what to do if and when such an information set is reached. However, once her level of awareness gets increased along the path of play, she may suspect that a similar revelation can happen again. She may then realize that whatever she plans to do, with her current level of

[4]Even if this condition is counterfactual due to the fact that the strategy s_i does not reach h_i. The conditioning is thus on the event that nature's moves, if there are any, have led to the information set h_i, and assuming that player i's past actions (in the information sets preceding h_i) have led to h_i *even if these actions are distinct than those prescribed by s_i.*

awareness, is in fact subject to reconsideration. That's why with unawareness, what a strategy specifies for future information sets should better be conceptualized as expressing current beliefs about one's future actions rather than as a rigid plan to which the player is bound to conform.

The following lemma describes the close connection between the two definitions when all of the information sets h_i are considered. The lemma follows from the principle of optimality in dynamic programming. The explicit proof appears in the appendix.

Lemma 1 *With a belief system b_i of player i,*

(i) if a strategy s_i of player i would be rational *at* all *information sets $h_i \in H_i$ then it is* rational *at all information sets $h_i \in H_i$; and*

(ii) if a strategy s_i of player i is rational *at* all *information sets $h_i \in H_i$, then there exists a strategy $\hat{s}_i$ which coincides with s_i at all information sets reached by s_i, such that $\hat{s}_i$* would be rational *at all information sets $h_i \in H_i$.*

The connection between the two definitions described in Lemma 1 is related to the notion of a *plan of action* (Rubinstein 1991, Reny 1992). A plan of player i specifies her action when she is called to play, and does not specify what she would do at information sets which are ruled out by that plan. Formally, a plan of action for player i is an equivalence class of strategies $\mathcal{P}_i \subset S_i$ such that two strategies $s_i, \hat{s}_i$ are in $\mathcal{P}_i$ if and only if for every strategy profile s_{-i} of the other players, (s_i, s_{-i}) and $(\hat{s}_i, s_{-i})$ induce the same distribution over terminal nodes in each of the trees of the game Γ. If $s_i \in \mathcal{P}_i$ we say that the strategy s_i *induces* the plan of action $\mathcal{P}_i$.

With this terminology, Lemma 1 implies:

Lemma 2 *For a given belief system b_i of player i, there exists a strategy s_i which* is rational *at all information sets $h_i \in H_i$ and induces the plan of action $\mathcal{P}_i$ if and only if there exists a strategy $\hat{s}_i$ which* would be rational *at all information sets $h_i \in H_i$ and induces the plan of action $\mathcal{P}_i$.*

We now turn to define rationalizability in generalized extensive-form games.

Definition 1 (Would-be rationalizable strategies) *Define, inductively, the following sequence of belief systems and strategies of player i.*

$B_i^1 = B_i$

$S_i^1 = \{s_i \in S_i$: *there exists a belief system* $b_i \in B_i^1$ *with which for every information set* $h_i \in H_i$ *player* i *is rational at* $h_i\}$

$\vdots$

$B_i^k = \{b_i \in B_i^{k-1}$: *for every information set* h_i, *if there exists some profile of the other players' strategies* $s_{-i} \in S_{-i}^{k-1} = \prod_{j \neq i} S_j^{k-1}$ *such that* s_{-i} *reaches* h_i *in the tree* T_{h_i}, *then* $b_i(h_i)$ *assigns probability 1 to* $S_{-i}^{k-1,T_{h_i}}\}$

$S_i^k = \{s_i \in S_i$: *there exists a belief system* $b_i \in B_i^k$ *with which for every information set* $h_i \in H_i$ *player* i *would be rational at* $h_i\}$

The set of player i*'s* ***would-be rationalizable strategies*** *is*

$$S_i^{\infty} = \bigcap_{k=1}^{\infty} S_i^k.$$

Remark 3 $S_i^k \subseteq S_i^{k-1}$ *for every* $k > 1$.

Proof. Consider $s_i \in S_i^k$. By definition, s_i would-be rational at each of player i's information sets given some belief system $b_i \in B_i^k$. Since $B_i^k \subseteq B_i^{k-1}$, s_i would also be rational at each of player i's information sets given a belief system in B_i^{k-1}, namely given b_i. Hence $s_i \in S_i^{k-1}$. $\square$

The generalization of Pearce's (1984) notion of extensive-form correlated rationalizable strategies is introduced next. The inductive definition below generalizes Definition 2 in Battigalli (1997), which he proved to be equivalent to Pearce's original definition.

Definition 2 (Extensive-form correlated rationalizable strategies) *For* $k \geq 1$ *let* $\hat{B}_i^k, \hat{S}_i^k$ *be defined inductively as* B_i^k, S_i^k *above, respectively, the only change being that the phrase "for every information set* $h_i \in H_i$ *player* i ***would be rational*** *at* h_i *" in the definition of* S_i^k *is changed to "for every information set* $h_i \in H_i$ *player* i ***is rational*** *at* h_i*" in the definition of* $\hat{S}_i^k$. *The set of player* i*'s* ***extensive-form correlated rationalizable strategies*** *is*

$$\hat{S}_i^{\infty} = \bigcap_{k=1}^{\infty} \hat{S}_i^k.$$

Remark 4 $\hat{S}_i^k \subseteq \hat{S}_i^{k-1}$ *for every* $k > 1$.

Proof. Analogous to the proof of Remark 3 above. □

Lemma 2 above implies the following proposition.

Proposition 1 *The set of strategies S_i^k is contained in $\hat{S}_i^k$, but S_i^k induces a set of plans of action identical to the set of plans of action induced by $\hat{S}_i^k$. Consequently, the set of would-be rationalizable strategies is contained in the set of extensive-form correlated rationalizable strategies,*

$$S_i^\infty = \bigcap_{k=1}^{\infty} S_i^k \subseteq \hat{S}_i^\infty = \bigcap_{k=1}^{\infty} \hat{S}_i^k$$

but both sets induce the same set of plans of actions.

The inclusion mentioned in the proposition may be strict. For instance, in our first example in the introduction (Figure 1), it is rationalizable for player 1 not to give the car to player 2 and to subsequently go to the Bach concert, but to have gone to the Stravinsky concert (or to the Bach concert, or to the Mozart concert) *had he given the car to 2*. In contrast, the only *would-be* rationalizable strategy of player 1 is not to give the car to player 2 and subsequently attend the Bach concert, but to have gone to the Mozart concert had he given the car to player 2. As the proposition asserts, no difference arises between rationality and would-be rationality along the unique realized path.

Proposition 2 *The set of would-be rationalizable strategies is non-empty. Consequently, the set of extensive-form correlated rationalizable strategies is non-empty.*

The proof is in the appendix.

What are the would-be rationalizable strategies in our battle-of-the-sexes example from the introduction (Figure 2)?

Remark 5 *In the Bach-Stravinsky-Mozart example with unawareness from the introduction (Figure 2), no player has a unique would-be rationalizable strategy.*

Proof. At the first level, any strategy would-be rational for player I except all strategies that prescribe going to the Mozart concert after "don't tell". For player II, both the Bach concert and the Stravinsky concert would-be rational if he is unaware of the Mozart concert. If he is aware of the Mozart concert, then only this concert is rational since it is a dominant action. Thus, $S^1_{II} = \{(B, M), (S, M)\}$. Not telling player II about the Mozart concert and going to the Bach concert would-be rational for player I assuming that she believes with probability at least $\frac{1}{2}$ that player II will go to the Bach concert under such circumstances. Telling player II about the Mozart concert and going to the Mozart concert would-be rational for player I if she believes with probability at least $\frac{1}{2}$ that player II would go to the Stravinsky concert if not told about the Mozart concert. To summarize,

$$S^2_I = \left\{ \begin{array}{c} (\text{“don't tell”}, B, M, B), (\text{“don't tell”}, B, M, S), \\ (\text{“tell”}, B, M, B), (\text{“tell”}, B, M, S), (\text{“tell”}, S, M, B), (\text{“tell”}, S, M, S) \end{array} \right\}$$

where the second (resp. third) component of the strategy vector refers to player I's choice after history “don't tell” (resp. “tell”), and the last component denotes the action in the lower subtree. Finally, note that $S^k_{II} = S^1_{II}$ for $k \geq 1$ and $S^k_I = S^2_I$ for $k \geq 2$. $\square$

When we compare this example to the game in which both players are aware of the Mozart concert but player I has the option of not providing her car for going to this concert (Figure 1), we note that the strategic implications of unawareness of actions are distinct from a situation in which both players are aware of the actions but some action may not always be available. The reason is that if player I keeps player II unaware of the Mozart concert, then player II can not infer the intention of player II to go to the Bach concert. In other words, awareness of an available action (providing the car for going to the Mozart concert) and certainty that it hasn't been taken has stronger strategic implications than unawareness of the very same action.

In the Bach-Stravinsky-Mozart example with unawareness from the introduction (Figure 2), the would-be rationalizable outcome is not unique. This is in contrast to the example with unavailability of actions instead, where there is a unique would-be rationalizable outcome. However, there exist also games where with unavailability of actions there are more would-be rationalizable outcomes than with unawareness of the same actions, as the example in Remark 8 in Section 5 demonstrates.

4 Prudent rationalizability

In normal-form games, iterated admissibility (i.e. iterative elimination of weakly dominated strategies) is a refinement of rationalizability. Van Damme (1989) and more generally Ben-Porath and Dekel (1992) showed that iterated elimination of weakly dominated strategies singles out the forward induction outcome in money-burning games.[5] One interpretation of iterated admissibility is that in every round of elimination, each player is prudent and hence does not exclude completely any strategy profile of the other players which has not been thus far eliminated. In this section, we use the idea of prudence to define an analogous notion of rationalizability for dynamic games:

Definition 3 (Prudent rationalizability) *Let*

$$\bar{S}_i^0 = S_i$$

For $k \geq 1$ define inductively

$$\bar{B}_i^k = \left\{ b_i \in B_i : \begin{array}{l} \textit{for every information set } h_i, \textit{ if there exists some profile} \\ s_{-i} \in \bar{S}_{-i}^{k-1} = \prod_{j \neq i} \bar{S}_j^{k-1} \textit{ of the other players' strategies} \\ \textit{such that } s_{-i} \textit{ reaches } h_i \textit{ in the tree } T_{h_i}, \textit{ then the support} \\ \textit{of } b_i(h_i) \textit{ is the set of strategy profiles } s_{-i} \in \bar{S}_{-i}^{k-1,T_{h_i}} \textit{ that reach } h_i \end{array} \right\}$$

$$\bar{S}_i^k = \left\{ s_i \in \bar{S}_i^{k-1} : \begin{array}{l} \textit{there exists } b_i \in \bar{B}_i^k \textit{ such that for all } h_i \in H_i \textit{ player } i \\ \textit{would be rational at } h_i \end{array} \right\}$$

The set of prudent rationalizable strategies of player i is

$$\bar{S}_i^\infty = \bigcap_{k=1}^{\infty} \bar{S}_i^k$$

Proposition 3 *The set of player i's prudent rationalizable strategies is non-empty.*

Proof. First, observe that $\bar{B}_i^k \neq \emptyset$ for every $k \geq 1$, because if an information set $h_i \in H_i$ is reached by some $s_{-i} \in \bar{S}_{-i}^{k-1}$, then s_{-i} reaches also all of i's information sets that precede h_i in the tree T_{h_i}.

[5] A similar result was shown by Herings and Vannetelbosch (1999) who defined iterated admissibility in terms of full support beliefs and called it trembling-hand perfect rationalizability.

We proceed by induction. $\bar{S}_i^0 = S_i$ and hence non-empty. Notice also that for every $b_i \in \bar{B}_i^1$, a standard backward induction procedure on the arborescence of information sets H_i yields a strategy $s_i \in \bar{S}_i^1$ with which player i would be rational $\forall h_i \in H_i$ given b_i.

Suppose, inductively, we have already shown that $\forall i \in I \quad \bar{S}_i^{k-1} \neq 0$ (and hence that $\bar{S}_{-i}^{k-1} \neq 0$), and also that for every $b_i \in \bar{B}_i^{k-1}$ there exists a strategy $s_i \in \bar{S}_i^{k-1}$ with which player i would be rational $\forall h_i \in H_i$ given b_i.

Let $b_i \in \bar{B}_i^k$. Let $\dot{H}_i \subseteq H_i$ be the set of i's information sets not reached by any profile $s_{-i} \in \bar{S}_{-i}^{k-1}$ but reached by some profile $s_{-i} \in \bar{S}_{-i}^{k-2}$. If $\dot{H}_i \neq \emptyset$, for every $h_i \in \dot{H}_i$ with no predecessor in $\dot{H}_i$, modify (if necessary) $b_i(h_i)$ so as to have full support on the profiles in $\bar{S}_{-i}^{k-2}$ that reach h_i, and in succeeding information sets modify b_i by Bayes rule whenever possible. Denote the modified belief system by $\dot{b}_i$. Then by construction also $\dot{b}_i \in \bar{B}_i^k$.

Consider a sequence of belief systems $b_{i,n} \in \bar{B}_i^{k-1}$ such that

$$\dot{b}_i = \left(\dot{b}_i(h_i')\right)_{h_i' \in H_i} \equiv \left(\lim_{n\to\infty} b_{i,n}(h_i')\right)_{h_i' \in H_i}$$

and given this sequence[6] $b_{i,n} \in \bar{B}_i^{k-1}$ let $s_{i,n} \in \bar{S}_i^{k-1}$ be a corresponding sequence of strategies with the property that given $b_{i,n}$, it is the case that with the strategy $s_{i,n}$ player i would be rational at every $h_i \in H_i$. Since player i has finitely many strategies, some strategy s_i appears infinitely often in the sequence $s_{i,n}$. Since expected utility is linear in beliefs and hence continuous, also given $\dot{b}_i$ it is the case that with the strategy s_i player i would be rational at every $h_i \in H_i$. Hence $s_i \in \bar{S}_i^k$ as well.

Now, since player i's set of strategies S_i is finite and by definition $\bar{S}_i^{k+1} \subseteq \bar{S}_i^k$ for every $k \geq 1$, for some ℓ we eventually get $\bar{S}_i^\ell = \bar{S}_i^{\ell+1}$ $\forall i \in I$ and hence $\bar{B}_i^{\ell+1} = \bar{B}_i^{\ell+2}$ $\forall i \in I$. Inductively,

$$\emptyset \neq \bar{S}_i^\ell = \bar{S}_i^{\ell+1} = \bar{S}_i^{\ell+2} = ...$$

and therefore

$$\bar{S}_i^\infty = \bigcap_{k=1}^{\infty} \bar{S}_i^k = \bar{S}_i^\ell \neq \emptyset$$

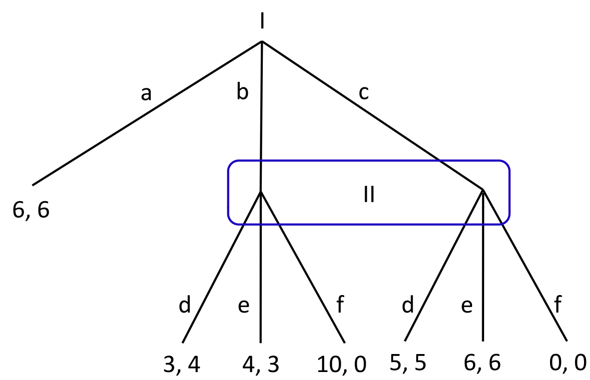
as required. □

[6] To construct such a sequence $b_{i,n} \in \bar{B}_i^{k-1}$, for every information set $h_i' \in H_i$ not reached by any $s_{-i} \in \bar{S}_{-i}^{k-1}$ define $b_{i,n}(h_i') = \dot{b}_i(h_i')$ for every $n \geq 1$; and for every $h_i' \in H_i$ with no predecessors but reached by some profile $s_{-i} \in \bar{S}_{-i}^{k-1}$ define $b_{i,n}(h_i') \in \Delta\left(\bar{S}_{-i}^{k-1}\right)$ to be any converging sequence of beliefs such that for every $n \geq 1$ the support of $b_{i,n}(h_i')$ is the subset of profiles in $\bar{S}_{-i}^{k-2}$ that reach h_i', while $\lim_{n\to\infty} b_{i,n}(h_i') = \dot{b}_i(h_i')$. In succeeding information sets reached by some $s_i \in \bar{S}_{-i}^{k-1}$ define $b_{i,n}(h_i')$ by Bayes rule whenever possible.

4.1 A first tension between rationalization and prudence: Divining the opponent's past behavior

In normal-form games, iterated admissibility is a refinement of rationalizability. Somewhat surprisingly, in extensive-form games prudent rationalizability is *not* a refinement of would-be rationalizability, as the following example (Figure 5) demonstrates.

Figure 5:



In this example, player 1 can guarantee herself the payoff 6 by choosing a and ending the game. If player 2 is called to play, should he believe that player 1 chose b or c? If player 1 is certain that player 2 is rational, she is certain that player 2 will not choose f. Hence, if player 2 is certain that player 1 is certain that he (player 2) is rational, then at his information set player 2 is certain that player 1 chose c. The reason is that among player 1's actions leading to 2's information set, c is the only action which, assuming 2 believes c was chosen and that 2 is rational and will hence choose e, yields player 1 the payoff 6, which is just as high as the payoff she could guarantee herself with the outside option a. Hence (a, e) and (c, e) are the profiles of extensive-form (correlated) rationalizable strategies (as well as would-be rationalizable strategies) in this game.

The notion of prudence, in contrast, embodies the idea that being prudently rational, player 1 shouldn't rule out completely any of 2's possible choices, and hence that c is strictly inferior for player 1 relative to her outside option a. Hence, if 2's information set is ever reached, the only way for 2 to rationalize this is to believe that 1 chose b, based on a belief ascribing a high probability to the event that 2 will foolishly choose f. Player 2's best reply to b is d; and player 1's best reply to d is a. Thus, the only profile of prudent rationalizable strategies in this game is (a, d).

This example demonstrates that in dynamic interactions the notions of rationalization and prudence might involve a *tension*. Extensive-form rationalizability embodies a best-rationalization principle (Battigalli 1997, Battigalli and Siniscalchi 2002); it is driven by the assumption that in each of his information sets, a player assesses the other players' future behavior by attributing to them the 'highest' level of rationality and mutual certainty of rationality consistent with the fact that the information set has indeed been reached. But, with the additional criterion of 'prudence', what should a player believe about the behavior of his opponent if, as in the example, the opponent's only action which is compatible with common certainty of rationality is imprudent on the part of the opponent?

The definition of prudent rationalizability resolves this tension unequivocally in favor of the prudence consideration. It remains open whether and how a more balanced and elaborate definition could resolve the tension in less an extreme fashion. We plan to address this challenge in future work. However, any definition would have to cut the Gordian knot in the above example in one particular way, choosing either d or e, and indeed both potential resolutions are backed by sensible intuitions. This suggests that for dynamic interactions we need not necessarily expect one ultimate definition of rationalizability taking into account both rationalization and prudence.

Remark 6 *The definition of prudent rationalizability employs would-be rationality. For standard extensive-form games, Brandenburger and Friedenberg (2007) defined extensive-form admissible strategies on the basis of rationality (rather than would-be rationality), and studied the connection with extensive-form rationalizability. They showed that under a "no relevant convexities" condition, extensive-form rationalizability and extensive-form admissibility coincide. However, the example in Figure 5 does* not *satisfy this condition, and hence demonstrates that in general extensive-form admissibility is not a refinement of extensive-form rationalizability.*

Nevertheless, as far as paths of play are concerned, in the above example the set of paths induced by prudent rationalizability (the path a) *is* a subset of the paths induced by (would-be) rationalizability (the paths a and (c, e)). This is an instance of a general phenomenon:

Proposition 4 (Prudent rationalizability refines would-be rationalizable paths) *The set of paths induced by profiles of prudent rationalizable strategies is a subset of the* *paths induced by profiles of would-be rationalizable strategies (or, equivalently, the paths induced by profiles of extensive-form correlated rationalizable strategies).*

The proof is in appendix A.

4.2 The refining power of prudent rationalizability

To demonstrate the extra power of prudent rationalizability, consider the following example of dynamic interaction with unawareness, which is a variant of example 3 in Ozbay (2007). There are 3 states of nature, $\omega_1, \omega_2, \omega_3$. A chance move chooses one out of four potential distributions over the states of nature:

$$\begin{aligned}\delta_1 &= (1,0,0)\\ \delta_2 &= (0,1,0)\\ \delta_3 &= (0,0,1)\\ \delta_4 &= \left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)\end{aligned}$$

An Announcer gets to know the distribution (but not the realization of the state of nature). A Decision Maker (DM) is initially aware only of the state ω_1 (and hence the DM is certain that ω_1 will be realized with certainty). However, before the DM chooses what to do, the Announcer can choose to make the DM aware of either ω_2, ω_3, none of them or both of them. Increased awareness makes the DM aware of the relevant marginals of the distributions. For instance, if the Announcer makes the DM aware of ω_2, the DM becomes aware of the set of distributions

$$\begin{aligned}\delta_{1|_{\{\omega_1,\omega_2\}}} &= (1,0)\\ \delta_{2|_{\{\omega_1,\omega_2\}}} &= (0,1)\\ \delta_{4|_{\{\omega_1,\omega_2\}}} &= \left(\frac{1}{2},\frac{1}{2}\right)\end{aligned}$$

and also becomes certain that the Announcer knows which of these is the true distribution.[7]

[7] In the spirit of Footnote 3 above, in Ozbay's example and in what follows the DM's beliefs about these marginal distributions will not be necessarily related to the prior probabilities with which the distributions were chosen by the chance move. That's why we do not even bother to specify the probabilities with which the chance move chooses the different distributions.

Subsequently, the DM should choose one out of three possible actions – left, middle or right. The payoffs to the players as a function of the chosen action and the state of nature appear in the following table:

	left	middle	right
ω_1	$3, 3$	$0, 0$	$2, 2$
ω_2	$0, 0$	$5, 5$	$2, 2$
ω_3	$2, 2$	$0, 0$	$2, 2$

The game is thus described in Figure 6 in the following page.

It is obvious that if the Announcer announces nothing, and hence the DM is certain that ω_1 prevails, the DM will choose 'left'.

What happens if the Announcer makes the DM aware of ω_2? The information set of the DM becomes

$$\left\{\delta_{1|_{\{\omega_1,\omega_2\}}}, \delta_{2|_{\{\omega_1,\omega_2\}}}, \delta_{4|_{\{\omega_1,\omega_2\}}}\right\}$$

The DM may then assign a high probability to $\delta_{1|_{\{\omega_1,\omega_2\}}}$,[8] and this will lead the DM to choose 'left'. Hence, assuming such a belief by the DM, it is rationalizable for the Announcer to make the DM aware of ω_2 when the Announcer knows that the true distribution is δ_1 (i.e. when the Announcer knows that ω_1 will be realized with probability 1).

This is not very sensible, though. After all, the Announcer can ensure that the DM chooses 'left' by not announcing any new state. When the Announcer likes the DM to choose 'left', it makes no sense on the Announcer's part to announce ω_2 and thus face the risk that the DM assigns a low probability to $\delta_{1|_{\{\omega_1,\omega_2\}}}$ and consequently choose 'middle'. This idea is captured by Ozbay's reasoning refinement to his awareness equilibrium notion[9], as well as by prudent rationalizability:

Put differently, instead of describing this game by a partially ordered set of trees, one for each level of awareness as in Figure 6, we could have replaced each tree with an arborescence in which the initial chance move is erased. Allowing for arborescences instead of trees in the framework for dynamic unawareness of Section 2 is straightforward, but for the sake of clarity of the exposition we avoid this explicit generalization in the body of the paper.

[8] That is, the DM may assign a high probability to strategies of the Announcer by which the Announcer announces ω_2 (and cause the DM's information set to become $\left\{\delta_{1|_{\{\omega_1,\omega_2\}}}, \delta_{2|_{\{\omega_1,\omega_2\}}}, \delta_{4|_{\{\omega_1,\omega_2\}}}\right\}$) when the Announcer has learned that the true distribution is δ_1.

[9] As explained in the introduction, we believe that equilibrium notions are somewhat questionable in the context of unawareness, and hence our focus on rationalizability.

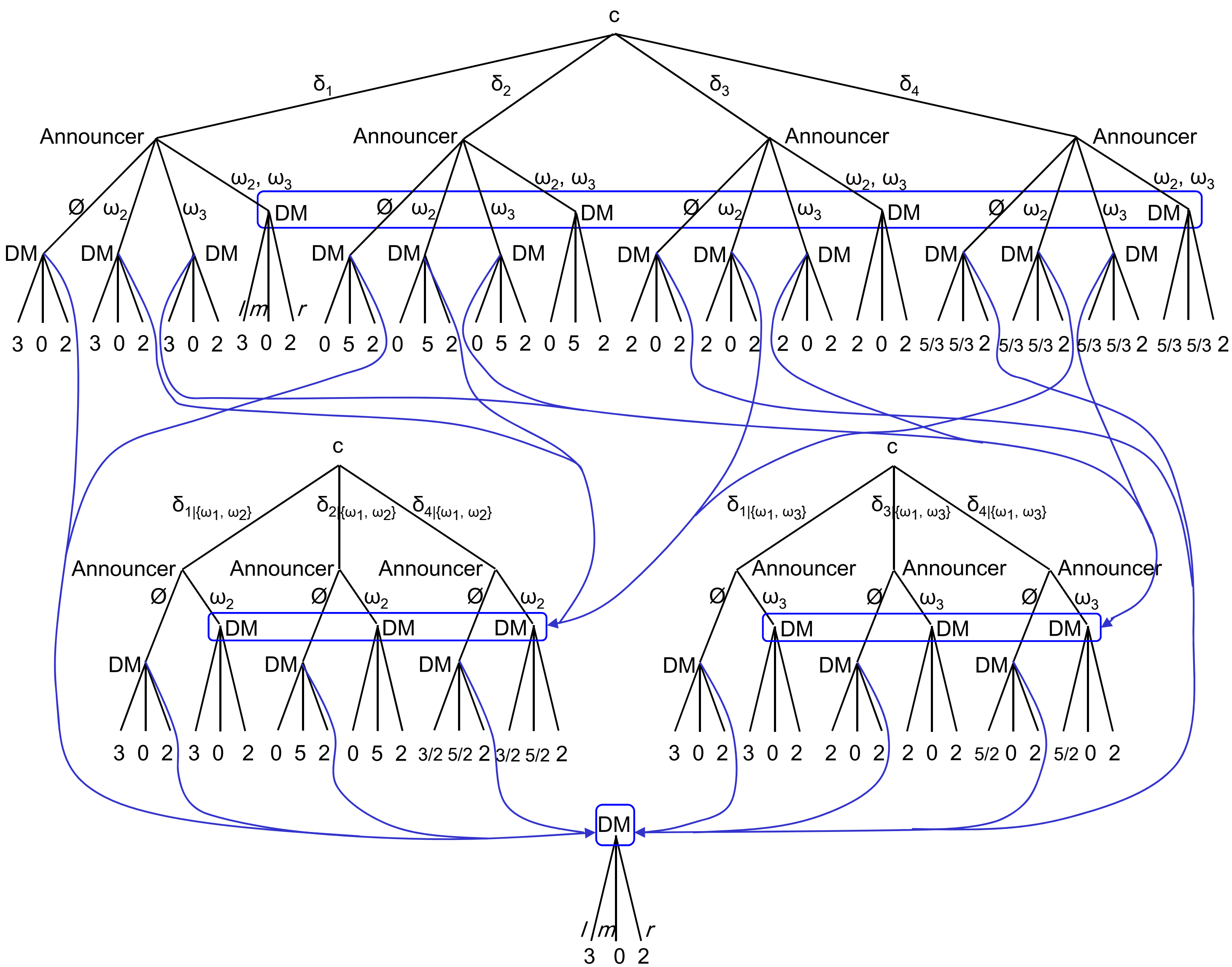

c
δ1
δ2
δ3
δ4
Announcer
Announcer
Announcer
Announcer
Ø ω2 ω3 ω2, ω3
DM
l m r
3 0 2 3 0 2 3 0 2 3 0 2 0 5 2 0 5 2 0 5 2 0 5 2 2 0 2 2 0 2 2 0 2 2 0 2 5/3 5/3 2 5/3 5/3 2 5/3 5/3 2 5/3 5/3 2
c
δ1|{ω1, ω2}
δ2|{ω1, ω2}
δ4|{ω1, ω2}
3 0 2 3 0 2 0 5 2 0 5 2 3/2 5/2 2 3/2 5/2 2
c
δ1|{ω1, ω3}
δ3|{ω1, ω3}
δ4|{ω1, ω3}
3 0 2 3 0 2 2 0 2 2 0 2 5/2 0 2 5/2 0 2
DM
l m r
3 0 2

Figure 6:

Proposition 5 *The DM has a unique prudent rationalizable strategy. With this strategy the DM chooses 'left' when no new state is announced, 'middle' when only ω_2 is announced, 'left' when only ω_3 is announced, and 'right' when both ω_2, ω_3 are announced.*

Proof. $\bar{B}^1_{DM}$ contains belief systems in which in the information set $\left\{\delta_{1|_{\{\omega_1,\omega_2\}}}, \delta_{2|_{\{\omega_1,\omega_2\}}}, \delta_{4|_{\{\omega_1,\omega_2\}}}\right\}$ (which follows the announcement of only ω_2 by the Announcer) the DM's belief assigns high probabilities to $\delta_{2|_{\{\omega_1,\omega_2\}}}, \delta_{4|_{\{\omega_1,\omega_2\}}}$. The strategies in $\bar{S}^1_{DM}$ corresponding to these belief systems prescribe 'middle' to the DM in the information set $\left\{\delta_{1|_{\{\omega_1,\omega_2\}}}, \delta_{2|_{\{\omega_1,\omega_2\}}}, \delta_{4|_{\{\omega_1,\omega_2\}}}\right\}$. The crucial point is that $\bar{B}^2_{Announcer}$ contains *only* belief systems that assign strictly positive probabilities to these strategies of the DM. Thus, with *any* belief system in $\bar{B}^2_{Announcer}$, it is sub-optimal for the Announcer to announce ω_2 in the announcer's information set $\{\delta_1\}$, in which the Announcer is certain of ω_1.[10] Hence, $\bar{S}^2_{Announcer}$ does not contain strategies in which the Announcer announces just ω_2 when the announcer's information set is $\{\delta_1\}$. We conclude that $\bar{B}^3_{DM}$ contains only belief systems in which the belief at the information set $\left\{\delta_{1|_{\{\omega_1,\omega_2\}}}, \delta_{2|_{\{\omega_1,\omega_2\}}}, \delta_{4|_{\{\omega_1,\omega_2\}}}\right\}$ assigns probability zero to $\delta_{1|_{\{\omega_1,\omega_2\}}}$. Hence, $\bar{S}^3_{DM}$ contains only strategies with which the DM chooses 'middle' at the information set $\left\{\delta_{1|_{\{\omega_1,\omega_2\}}}, \delta_{2|_{\{\omega_1,\omega_2\}}}, \delta_{4|_{\{\omega_1,\omega_2\}}}\right\}$.

Furthermore, already $\bar{S}^1_{DM}$ contains only strategies with which the DM chooses 'left' at the information set $\left\{\delta_{1|_{\{\omega_1,\omega_3\}}}, \delta_{3|_{\{\omega_1,\omega_3\}}}, \delta_{4|_{\{\omega_1,\omega_3\}}}\right\}$ (i.e. when the Announcer announces just the new state ω_3). This is because prudent rationalizability implies that all the belief systems in $\bar{B}^1_{DM}$ assign a positive probability to strategies of the Announcer with which the Announcer announces the new state ω_3 even when the Announcer's information set (from the point of view of the DM!) is $\left\{\delta_{1|_{\{\omega_1,\omega_3\}}}\right\}$ or $\left\{\delta_{4|_{\{\omega_1,\omega_3\}}}\right\}$.

Also, $\bar{B}^1_{DM}$ contains belief systems in which the DM's belief in the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$ (when the Announcer announces both new states ω_2, ω_3) assigns high probability to δ_2. The strategies in $\bar{S}^1_{DM}$ corresponding to these belief systems prescribe 'middle' to the DM in the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$. Hence, $\bar{B}^2_{Announcer}$ contains *only* belief systems that assign strictly positive probabilities to these strategies of the DM. Thus, with *any* belief system in $\bar{B}^2_{Announcer}$, it is sub-optimal for the Announcer to announce both ω_2 and ω_3 in the announcer's information sets $\{\delta_1\}$ and $\{\delta_3\}$. Similarly, $\bar{B}^1_{DM}$ contains belief systems in which the DM's belief in the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$ assigns high probability to δ_1. The strategies in $\bar{S}^1_{DM}$ corresponding to these belief systems prescribe 'left' to the DM in the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$. Hence, $\bar{B}^2_{Announcer}$ contains

[10] Because according to every belief system in $\bar{B}^2_{Announcer}$, announcing just ω_2 will lead the DM with a positive probability to choose 'middle'.

only belief systems that assign strictly positive probabilities to these strategies of the DM. Thus, with *any* belief system in $\bar{B}^2_{Announcer}$, it is sub-optimal for the Announcer to announce both ω_2 and ω_3 in the Announcer's information sets $\{\delta_1\}, \{\delta_2\}$ or $\{\delta_3\}$. We conclude that $\bar{B}^3_{DM}$ contains only belief systems in which the belief at the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$ assigns probability zero to $\delta_1, \delta_2, \delta_3$. That is, $\bar{B}^3_{DM}$ contains only belief systems that assign probability 1 to δ_4 at the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$. Hence, $\bar{S}^3_{DM}$ contains only strategies with which the DM chooses 'right' at the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$.

We thus conclude that $\bar{S}^3_{DM}$ contains a unique strategy s^*_{DM}. This strategy prescribes the DM to choose 'left' in the information set $\left\{\delta_{1|_{\{\omega_1\}}}\right\}$ (i.e. when the Announcer does not announce any new state), to choose 'middle' in the information set $\left\{\delta_{1|_{\{\omega_1,\omega_2\}}}, \delta_{2|_{\{\omega_1,\omega_2\}}}, \delta_{4|_{\{\omega_1,\omega_2\}}}\right\}$ (i.e. when the Announcer announces just the new state ω_2), to choose 'left' in the information set $\left\{\delta_{1|_{\{\omega_1,\omega_3\}}}, \delta_{3|_{\{\omega_1,\omega_3\}}}, \delta_{4|_{\{\omega_1,\omega_3\}}}\right\}$ (i.e. when the Announcer announces just the new state ω_3) and to choose 'right' in the information set $\{\delta_1, \delta_2, \delta_3, \delta_4\}$ (i.e. when the Announcer announces both new states ω_2, ω_3).[11] □

4.3 A second tension between rationalizability and prudence: Divining the opponent's future behavior

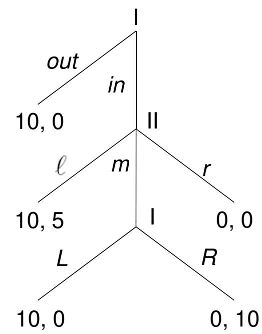
In Figure 5 we demonstrated the tension between the considerations of rationalization and prudence when a player tries to divine his opponent's past actions. A related but distinct tension arises when a player tries to deduce the opponent's future behavior from past actions of that opponent. Consider the following example in Figure 7.

In this example, *in* is imprudent for player 1 (since by going *out* she can guarantee a payoff of 10, while by moving *in* she risks getting 0 if player 2 would rather foolishly choose r). This means that if player 1 does move *in* and player 2 gets to play, no prudent strategy in $\bar{S}^1_1$ reaches 2's information set. Hence, the beliefs $\bar{B}^2_2$ of player 2 about player 1's future actions are not restricted. In particular, it contains beliefs by which if player 2 chooses m, player 1 will foolishly choose R (with a high probability). That's why both m and ℓ are prudent rationalizable for player 2.

However, it is not very sensible on the part of player 2 to believe that following m player 1 may choose R. After all, when player 2 has to move, player 1 has already

[11] This is also the unique strategy of the DM which is part of an awareness equilibrium satisfying reasoning refinement in Ozbay (2007).

Figure 7:



proved to be imprudent, but not irrational. Indeed, player 1's rationalizable (though imprudent) strategy (in, L) yields her the payoff 10 in conjunction with 2's only (would-be) rationalizable strategy ℓ, as well as in conjunction with 2's prudent rationalizable strategy m; and this payoff is the same as the payoff player 1 gets from her only prudent rationalizable strategy (out, L).

Thus, as long as player 1 has been rational (even if imprudent) thus far, it makes more sense for player 2 to believe that player 1 will continue to be rational (though possibly imprudent) in the future. Restricting player 2's beliefs according to this logic would cross out the non-sensical choice m.

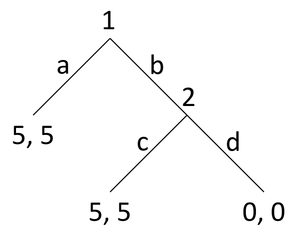
Already Pearce (1984) was well aware of this tension, which motivated his definition of **cautious extensive-form rationalizability**. That definition involves refining the set of rationalizable strategies by another round of strategy elimination with full support beliefs about the other players' surviving strategies; and then repeating this entire procedure – the standard iterative elimination process as in the definition of rationalizability, followed by one round assuming full-support beliefs – ad infinitum. In the above example, cautious extensive-form rationalizability does indeed rule out the strategy m for player 2.

However, as Pearce (1984) himself admits, the definition of cautious extensive-form rationalizability is not really satisfactory, as the following simple example of his shows.

In this example, the strategy d is irrational for player 2. Once d is crossed out, both a and b are extensive-form rationalizable for player 1, and are actually also cautious extensive-form rationalizable. Notice that in contrast, b *does* get crossed out by *prudent rationalizability,* and the only prudent rationalizable strategy for player 1 is a.

To sum up, we believe it is worth exploring further a more fine-tuned *refinement*

Figure 8:



of rationalizability which would take prudence considerations into account, one which would be more subtle than Pearce's cautious extensive-form rationalizability. As the above examples suggest, such a definition would be involved, and would take us beyond the scope of the current paper. We plan to address this issue in future work.

4.4 Strategy elimination vs. belief systems reduction

Definition 1 of would-be rationalizable strategies involves, as in Battigalli (1997), an iterative reduction procedure of belief systems (that is, by definition $B_i^k \subseteq B_i^{k-1}$), and this definition implies (Remark 3) that strategies get iteratively eliminated ($S_i^k \subseteq S_i^{k-1}$); and the same is true also for extensive-form correlated rationalizable strategies – by definition $\hat{B}_i^k \subseteq \hat{B}_i^{k-1}$ and hence $\hat{S}_i^k \subseteq \hat{S}_i^{k-1}$. In contrast, the inductive definition of **prudent** rationalizable strategies involves an iterative elimination of strategies (that is, by definition $\bar{S}_i^k \subseteq \bar{S}_i^{k-1}$, in analogy with the original formulation of Pearce (1984) for extensive-form rationalizability by an iterative elimination procedure), but in the case of prudence it is **not** generally the case that $\bar{B}_i^k \subseteq \bar{B}_i^{k-1}$. Indeed, when $\bar{S}_{-i}^k \subsetneq \bar{S}_{-i}^{k-1}$:

- If the set of strategy profiles in $\bar{S}_{-i}^k$ reaching some information set $h_i \in H_i$ is a proper, non-empty subset of the strategy profiles in $\bar{S}_{-i}^{k-1}$ that reach h_i, then the support of each belief $\bar{b}_i^{k-1}(h_i)$ in each belief system $\bar{b}_i^{k-1} \in \bar{B}_i^{k-1}$ is strictly larger than the support of any belief $\bar{b}_i^k(h_i)$ for $\bar{b}_i^k \in \bar{B}_i^k$.
- For information sets h_i *not* reached by $\bar{S}_{-i}^k$, there is no restriction (beyond Bayes rule) on $\bar{b}_i^k(h_i)$ for $\bar{b}_i^k \in \bar{B}_i^k$. No such restriction is needed, because if we define

$$m_{h_i}^k = \max\left\{m < k : \exists s_{-i} \in \bar{S}_{-i}^m \text{ that reaches } h_i\right\}$$

then for $s_i^k \in \bar{S}_i^k$ the restrictions on i's actions $s_i^k(h_i)$ at h_i were already determined at stage $m^k_{h_i}$, since by definition $s^k_i \in \bar{S}^k_i \subseteq \bar{S}^{m^k_{h_i}}_i$.

Is it nevertheless feasible to define prudent rationalizability via a reduction process of belief systems? Asheim and Perea (2005) proposed to look at **systems of conditional lexicographic probabilities** – belief systems in which each belief at an information set is itself a **lexicographic probability system** (Blume, Brandenburger and Dekel 1991) about the other players' strategy profiles. Using belief systems which are conditional lexicographic probabilities we could, in the spirit of Stahl (1995), put forward an equivalent definition of prudent rationalizable strategies involving an iterative reduction procedure of belief systems rather than an iterative elimination procedure of strategies. In each round of the procedure, the surviving belief systems would be those in which at each information set, ruled-out strategy profiles of the other players (i.e. strategy profiles outside $\bar{S}^{m^k_{h_i}}_{-i}$) would be deemed infinitely less likely than the surviving strategy profiles, but infinitely more likely than strategy profiles which had already been eliminated in previous rounds. We leave the precise formulation of such an equivalent definition to future work.

In their paper, Asheim and Perea (2005) proposed the notion of **quasi-perfect rationalizability**, which also involves the idea of cautious beliefs. Quasi-perfect rationalizability is distinct from our notion of prudent rationalizability. The difference is that with prudent rationalizability (as with would-be rationalizability), a player need not believe that another player's future behavior must be rationalizable to a higher order than that exhibited by that other player in the past; in contrast, with the quasi-perfect rationalizable strategies of Asheim and Perea (2005), a player should ascribe to her opponent the highest possible level of rationality in the future even if this opponent has already proved to be less rational in the past. That's why quasi-perfect rationalizability implies backward induction in generic perfect information games, while our prudent rationalizable strategies need not coincide with the backward induction strategies in such games (though they do generically lead to the backward induction path – the argument is the same as in Reny 1992 and Battigalli 1997, since in generic perfect information games prudent rationalizability coincides with extensive-form rationalizability in terms of realized paths).

5 Characterization by conditional dominance

5.1 Associated normal-form games

Consider a generalized extensive-form game Γ with a partially ordered set of trees $\mathbf{T}$. The *associated normal-form game* G is defined by $\langle I, \langle (S_i^T)_{i\in I}, (u_i^T)_{i\in I}\rangle_{T\in\mathbf{T}}\rangle$, where I is the set of players in Γ and S_i^T is player i's set of T-partial strategies. If player i is not active in trees $T' \in \mathbf{T}$ with $T \hookrightarrow T'$, then $S_i^T = \emptyset$. Recall that if player i is active at node $n \in T$, then at node n the strategy $s_i \in S_i^T$ prescribes to her the action $s_i(\pi_i(n))$. Hence, each profile of strategies in S^T induces a distribution over terminal nodes in T (even if there is a player active in T with no information set in T). $u_i^T(s)$ is the expected value of the payoffs associated with the terminal nodes in T reached by $s \in S^T$ weighted by the probabilities associated to the moves of nature. (Note that while strategy profiles in S^T reach terminal nodes also in trees $T' \in \mathbf{T}$, $T \hookrightarrow T'$, u_i^T concerns payoffs in the tree T only.)

Recall that H_i^T denotes player i's set of extensive form information sets in the T-partial game. For each $h_i \in H_i^T$, let $S^T(h_i) \subseteq S^T$ be the subset of the T-partial strategy space containing T-partial strategy profiles that reach the information set h_i. Define also $S_i^T(h_i) \subseteq S_i^T$ and $S_{-i}^T(h_i) \subseteq S_{-i}^T$ to be the set of player i's T-partial strategies reaching h_i and the set of profiles of the other players' T-partial strategies reaching h_i respectively. For the entire game denote by $S(h_i) \subseteq S$ the set of strategy profiles that reach h_i. Similarly, $S_i(h_i) \subseteq S_i$ and $S_{-i}(h_i) \subseteq S_{-i}$ are the set of player i's strategies reaching h_i and the set of profiles of the other players' strategies reaching h_i respectively.

Given Γ and its associated normal-form game G, define player i's set of *normal-form information sets*[12] by

$$\mathcal{X}_i = \{S^{T_{h_i}}(h_i) : h_i \in H_i\}.$$

These are the "normal form versions" of information sets in the generalized extensive-form game.

For $T \in \mathbf{T}$, any set $Y \subseteq S^T$ is called a *restriction* for player i (or an i-product set) of T-partial strategies if $Y = Y_i \times Y_{-i}$ for some $Y_i \subseteq S_i^T$ and $Y_{-i} \subseteq S_{-i}^T$. Clearly, a player's

[12] We abuse here slightly existing terminology. In the literature on standard games, normal-form information sets refer more generally to subsets of the strategy space of a pure strategy reduced normal-form game for which there exists an extensive-form game with corresponding information sets (see Mailath, Samuelson and Swinkels, 1993). For our characterization, we are just interested in the normal form versions of information sets of a given generalized extensive-form game.

normal-form information set is a restriction. I.e., if $S^{T_{h_i}}(h_i)$ is a normal form information set of player i, then it is a restriction for player i of T_{h_i}-partial strategy profiles.

5.2 Iterated conditional strict dominance and extensive-form rationalizability

We say that $s_i \in S_i^T$ is *strictly dominated* in a restriction $Y \subseteq S^T$ if $s_i \in Y_i$, $Y_{-i} \neq \emptyset$, and there exists a mixed strategy $\sigma_i \in \Delta(Y_i)$ such that $u_i^T(\sigma_i, s_{-i}) > u_i^T(s_i, s_{-i})$ for all $s_{-i} \in Y_{-i}$.

Denote by $\mathbf{S} = \bigcup_{T \in \mathbf{T}} S^T$ and $\mathbf{S}_i = \bigcup_{T \in \mathbf{T}} S_i^T$.

For $T \hookrightarrow T'$ and a T-partial strategy $s_i \in S_i^T$, we denote the T'-partial strategy $s_i^{T'} \in S_i^{T'}$ induced by s_i. For $\tilde{s}_i \in S_i^{T'}$, define

$$[\tilde{s}_i] := \bigcup_{T \hookrightarrow T'} \{s_i \in S_i^T : s_i^{T'} = \tilde{s}_i\}.$$

That is, $[\tilde{s}_i]$ is the set of strategies in $\mathbf{S}_i$ which at information sets $h_i \in H_i^{T'}$ prescribe the same actions as strategy $\tilde{s}_i$.

Let $(Y^T)_{T \in \mathbf{T}}$ be a collection of i-product sets, one for each $T \in \mathbf{T}$. Define $\mathbf{Y} = \bigcup_{T \in \mathbf{T}} Y^T$. Given such a $\mathbf{Y}$, we say that $s_i \in S_i^T$ is *conditionally strictly dominated on* $(\mathcal{X}_i, \mathbf{Y})$ if (1) there exists a normal-form information set $X \in \mathcal{X}_i$, $X \subseteq S^T$ such that s_i is strictly dominated in $X \cap Y^T$ or (2) for some $\tilde{s}_i \in S_i^{T'}$, $T \hookrightarrow T'$, $s_i \in [\tilde{s}_i]$, we have that $\tilde{s}_i$ is strictly dominated in $X \cap Y^{T'}$ for some normal-form information set $X \in \mathcal{X}_i$, $X \subseteq S^{T'}$.

(Note that (2) implies (1), but the explicit distinction between (1) and (2) makes the presentation more transparent.)

For $\mathbf{Y}$ define

$$U_i(\mathbf{Y}) = \{s_i \in \mathbf{S}_i : s_i \text{ is not conditionally strictly dominated on } (\mathcal{X}_i, \mathbf{Y})\},$$

$$U(\mathbf{Y}) = \bigcup_{T \in \mathbf{T}} \prod_{i \in I} \left(U_i(\mathbf{Y}) \cap S_i^T\right),$$

and

$$U_{-i}(\mathbf{Y}) = \bigcup_{T \in \mathbf{T}} \prod_{j \in I \setminus \{i\}} \left(U_j(\mathbf{Y}) \cap S_j^T\right).$$

Define inductively

$U^0(\mathbf{S}) = \mathbf{S}$,

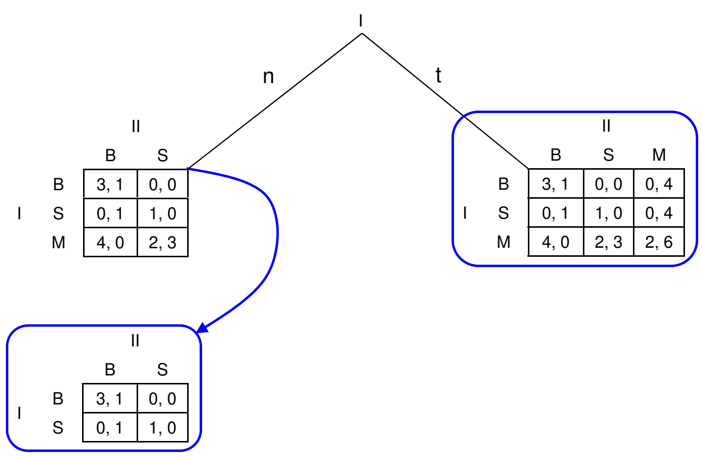
$U^{k+1}(\mathbf{S}) = U(U^k(\mathbf{S}))$ for $k \geq 0$,

$U^\infty(\mathbf{S}) = \bigcap_{k=0}^\infty U^k(\mathbf{S})$,

and similarly for $U_i^k(\mathbf{S})$ and $U_{-i}^k(\mathbf{S})$.

Example. Consider the game below whose extensive form is identical to the Battle-of-the-Sexes game with unawareness from the introduction but whose payoffs are quite different (Figure 9). In this strategic situation, player I may deceive player II by hiding player II's dominant action M. As we will see, this example allows us to demonstrate some features of iterated conditional dominance that we couldn't have demonstrated with the introductory example.

Figure 9:



The associated normal form is given in Figure 10. The lower strategic form game is the normal form associated with the T'-partial extensive-form game and the normal form associated with the T-partial extensive-form game is the upper strategic form game. Player I is the row player, while player II is the column player. For the row player in the upper strategic form, the first component of her strategy refers to actions at the root of the upper tree, the second to her action in the upper left subgame, the third to the upper right subgame, and the last component to the action in the lower game. For the column player, the first component of his strategy refers to the action taken in the upper information set while the second is the action taken in the lower information set.

Each boxed cell is a normal-form information set. The entire upper strategic form is

Figure 10: The Associated Normal Form Game

T	BB	BS	SB	SS	MB	MS
$nBBB$	3, 1	0, 0	3, 1	0, 0	3, 1	0, 0
$nSBB$	0, 1	1, 0	0, 1	1, 0	0, 1	1, 0
$nMBB$	4, 0	2, 3	4, 0	2, 3	4, 0	2, 3
$nBSB$	3, 1	0, 0	3, 1	0, 0	3, 1	0, 0
$nSSB$	0, 1	1, 0	0, 1	1, 0	1, 0	1, 0
$nMSB$	4, 0	2, 3	4, 0	2, 3	4, 0	2, 3
$nBMB$	3, 1	0, 0	3, 1	0, 0	3, 1	0, 0
$nSMB$	0, 1	1, 0	0, 1	1, 0	1, 0	1, 0
$nMMB$	4, 0	2, 3	4, 0	2, 3	4, 0	2, 3
$nBBS$	3, 1	0, 0	3, 1	0, 0	3, 1	0, 0
$nSBS$	0, 1	1, 0	0, 1	1, 0	1, 0	1, 0
$nMBS$	4, 0	2, 3	4, 0	2, 3	4, 0	2, 3
$nBSS$	3, 1	0, 0	3, 1	0, 0	3, 1	0, 0
$nSSS$	0, 1	1, 0	0, 1	1, 0	1, 0	1, 0
$nMSS$	4, 0	2, 3	4, 0	2, 3	4, 0	2, 3
$nBMS$	3, 1	0, 0	3, 1	0, 0	3, 1	0, 0
$nSMS$	0, 1	1, 0	0, 1	1, 0	1, 0	1, 0
$nMMS$	4, 0	2, 3	4, 0	2, 3	4, 0	2, 3
$tBBB$	3, 1	3, 1	0, 0	0, 0	0, 4	0, 4
$tSBB$	3, 1	3, 1	0, 0	0, 0	0, 4	0, 4
$tMBB$	3, 1	3, 1	0, 0	0, 0	0, 4	0, 4
$tBSB$	0, 1	0, 1	1, 0	1, 0	0, 4	0, 4
$tSSB$	0, 1	0, 1	1, 0	1, 0	0, 4	0, 4
$tMSB$	0, 1	0, 1	1, 0	1, 0	0, 4	0, 4
$tBMB$	4, 0	4, 0	2, 3	2, 3	2, 6	2, 6
$tSMB$	4, 0	4, 0	2, 3	2, 3	2, 6	2, 6
$tMMB$	4, 0	4, 0	2, 3	2, 3	2, 6	2, 6
$tBBS$	3, 1	3, 1	0, 0	0, 0	0, 4	0, 4
$tSBS$	3, 1	3, 1	0, 0	0, 0	0, 4	0, 4
$tMBS$	3, 1	3, 1	0, 0	0, 0	0, 4	0, 4
$tBSS$	0, 1	0, 1	1, 0	1, 0	0, 4	0, 4
$tSSS$	0, 1	0, 1	1, 0	1, 0	0, 4	0, 4
$tMSS$	0, 1	0, 1	1, 0	1, 0	0, 4	0, 4
$tBMS$	4, 0	4, 0	2, 3	2, 3	2, 6	2, 6
$tSMS$	4, 0	4, 0	2, 3	2, 3	2, 6	2, 6
$tMMS$	4, 0	4, 0	2, 3	2, 3	2, 6	2, 6

T'	B	S
B	3, 1	0, 0
S	0, 1	1, 0

the normal-form information set of player 1 (but not player 2) associated with player 1's information set at the beginning of the T-partial game (but not in the T'-partial game). We denote this information set by $X_1(\emptyset^T)$. The upper boxed cell in the upper strategic form is the normal-form information set of player 1 (but not of player 2) corresponding to her extensive form information set after the history n in the T-partial game (but not in the T'-partial game). We denote it by $X_1(n)$. The lower boxed cell in the upper strategic form game is the normal-form information set for both player 1 and 2 corresponding to the information sets after history t in the T-partial game (but not in the T'-partial game). We denote it by $X_i(t)$.

Finally, the lower strategic form game is a normal form information set for both player 1 and 2 both for corresponding information sets in the T'-partial normal form and in the T-partial normal form game. It is also the normal-form information set for player 2 corresponding to his information set $\pi_2(n)$ in the T-partial game. We denote it by $X_i(\emptyset^{T'}) = X_2(n)$.

The definition of $\mathbf{S}_i$ is illustrated by the example $\mathbf{S}_2 = \{BB, BS, SB, SS, MB, MS, B, S\}$, while the definition $[\tilde{s}_i]$ can be illustrated by $[''S'']=\{BS,SS,MS,S\}$. These are all the strategies of player 2 that prescribe action "S" ("Stravinsky") at the information set $\pi_2(n)$.

The iterated elimination of conditionally strictly dominated strategies proceeds as follows:

$$\begin{aligned}
U_i^0(\mathbf{S}) &= \mathbf{S}_i, i=1,2 \\
U_1^1(\mathbf{S}) &= \{nMBB, nMSB, nMMB, nMBS, nMSS, nMMS, \\
&\quad\; tBMB, tSMB, tMMB, tBMS, tSMS, tMMS, \\
&\quad\; B, S\} \\
U_2^1(\mathbf{S}) &= \{MB, B\}
\end{aligned}$$

For instance, strategy $nSBB$ is conditionally strictly dominated by $nMBB$ in the normal-form information set $X_1(\emptyset^T)$ or $X_1(n)$. More interestingly, MS is conditionally strictly dominated on $(\mathcal{X}_2,\mathbf{S})$ because $MS\in[''S'']$ and S is strictly dominated by B in $X_2(n)$. So this example demonstrates that an action in the upper normal form may be deleted because of strict dominance in the lower normal form. This is one reason why we chose this game to demonstrate iterated conditional strict dominance rather than the introductory example.

Applying the definitions iteratively yields

$$\begin{aligned}
U_1^2(\mathbf{S}) &= \{nMBB, nMSB, nMMB, \\
&\quad\; tBMB, tSMB, tMMB, tBMS, tSMS, tMMS, B\} \\
U_2^2(\mathbf{S}) &= U_2^1(\mathbf{S}) = \{MB, B\} \\
U_1^3(\mathbf{S}) &= \{nMBB, nMSB, nMMB, B\} \\
&= U_1^k(\mathbf{S}) \text{ for } k\geq 3 \\
U_2^3(\mathbf{S}) &= U_2^2(\mathbf{S}) = \{MB, B\} \\
&= U_2^k(\mathbf{S}) \text{ for } k\geq 1
\end{aligned}$$

Note that $U_i^\infty(\mathbf{S})\cap S_i=\hat{S}_i^\infty$. That is, the set of iterated elimination of conditionally strictly dominated strategies coincides with the set of extensive-form correlated rationalizable strategies, and both predict that player I will not give the car to player II and attend the Mozart concert, while player II will attend the Bach concert.

The following proposition generalizes the observation made in the example.

Proposition 6 *For every finite generalized extensive form game,* $U_i^k(\mathbf{S}) \cap S_i = \hat{S}_i^k$*,* $k \geq 1$*. Consequently,* $U_i^\infty(\mathbf{S}) \cap S_i = \hat{S}_i^\infty$*.*

The proof is in appendix A.

Remark 7 *If in the definition of prudent rationalizability, would-be rationality is replaced by rationality, then prudent rationalizability can be characterized by iterated elimination of conditional* weakly *dominated strategies. The proof is analogous. Instead of using Lemma 3 in Pearce (1984), we would now use Lemma 4 in Pearce (1984). Moreover, iterated conditional weak dominance is equivalent to iterated admissibility in the normal-form. This is so because if a strategy weakly dominates a replacement in an information set, then the payoffs from the strategy and its replacement outside the information set must coincide (since otherwise it wouldn't be a replacement).*

Remark 8 *Consider a game with unavailability of actions analogous to Figure 1 but with the payoffs as in the example of this section. Then the set of would-be rationalizable paths include the one in which player* I *gives the car to player* II *and they both go to the Mozart concert, as well as the paths in which player* I *doesn't give the car to player* II *and then player* I *goes either to the Bach or to the Mozart concert and player* II *goes either to the Bach or to the Stravisnky concert. In contrast with the example in the introduction, this example therefore shows that would-be rationalizability does not necessarily yield a sharper prediction under unavailability of actions than under unawareness of the same actions.*

6 Unawareness

Generalized games can describe many types of games with subjective reasoning. In a generalized game, a player cannot imagine that she can take an action which is physically unavailable to her (property I4), but at a given information set $\pi_i(n)$ she can nevertheless imagine that in a succeeding information set she will have an action which is actually nowhere available in the tree T_n as in the example of Figure 4. Furthermore, she can imagine that along the path of play another player will forget the history of play, i.e.

that at a later information set this other player will imagine he is playing in a game tree which is completely unrelated to the game tree he imagined at an earlier stage along the path.

Since our main motivation is to analyze games with unawareness rather than games with arbitrary kinds of subjective reasoning, it is worthwhile spelling out additional properties of generalized games in which the only reason for players' misconception of the game is unawareness (and mutual unawareness) of available actions. In *extensive-form games with unawareness* the set of trees $\mathbf{T}$ forms a join semi-lattice under the inclusion partial order relation $\preceq$. The maximal tree in this join semi-lattice is the modeler's objective description of feasible paths of play.

The following additional properties parallel properties of static unawareness structures in Heifetz, Meier and Schipper (2006). [13]

U0 Confined awareness: If $n \in T$ and $i \in I_n$ then $\pi_i(n) \subseteq T'$ with $T' \preceq T$.

U1 Generalized reflexivity: If $T' \preceq T$, $n \in T$, $\pi_i(n) \subseteq T'$ and T' contains a copy $n_{T'}$ of n, then $n_{T'} \in \pi_i(n)$.

U2 Introspection: If $n' \in \pi_i(n)$ then $\pi_i(n') = \pi_i(n)$. (I.e. property I2.)

U3 Subtrees preserve awareness: If $n \in T'$, $n \in \pi_i(n)$, $T \preceq T'$, and T contains a copy n_T of n, then $n_T \in \pi_i(n_T)$.

U4 Subtrees preserve ignorance: If $T \preceq T' \preceq T''$, $n \in T''$, $\pi_i(n) \subseteq T$ and T' contains the copy $n_{T'}$ of n, then $\pi_i(n_{T'}) = \pi_i(n)$.

U5 Subtrees preserve knowledge: If $T \preceq T' \preceq T''$, $n \in T''$, $\pi_i(n) \subseteq T'$ and T contains the copy n_T of n, then $\pi_i(n_T)$ consists of the copies that exist in T of the nodes of $\pi_i(n)$.

The following remark is analogous to Remark 3 in Heifetz, Meier and Schipper (2006).

Remark 9 *U5 implies U3.*

Proof. If $n \in T'$, $n \in \pi_i(n)$, $T \preceq T'$, and T contains a copy n_T of n, then by U5 $\pi_i(n_T)$ must consist of the copies that exist in T of the nodes of $\pi_i(n)$. Since by assumption $n \in \pi_i(n)$ and the copy n_T exists in T, we must have $n_T \in \pi_i(n_T)$. □

[13]The number of each property corresponds to the respective property in Heifetz, Meier and Schipper (2006).

Remark 10 *U0 implies I0. U1 implies I1.*

Remark 11 *U0 is equivalent to I0 and $T \rightarrowtail T'$ implies $T' \preceq T$.*

Proof. I0 and $T \rightarrowtail T'$ implies $T' \preceq T$ are equivalent to if there exists $n \in T$ and $i \in I_n$ such that $\pi_i(n) \subseteq T'$ then $T' \preceq T$. □

All these properties are static properties in the sense that they relate nodes on one tree with copies of those nodes in another tree. One may wonder about dynamic properties of unawareness. The following property states that a player can not become unaware during the play.

DA Awareness may only increase along a path: If there is a path $n, \ldots, n'$ in some subtree T such that player i is active in n and n', and $\pi_i(n) \subseteq T$ while $\pi_i(n') \subseteq T'$ then $T' \succeq T$.

Recall that I3 is the only completely new property imposed on information sets in generalized games.

Remark 12 *Suppose that U0 to U2 hold. Then DA if and only if I3.*

Proof. More precisely, we will show first that if I1 holds, then I3 implies DA. Second, if U0 and I2 holds, then DA implies I3. This implies the result by Remark 10.

If $n, ..., n'$ is path in T such that $i \in I_n \cap I_{n'}$, $\pi_i(n) \subseteq T$ while $\pi_i(n') \subseteq T'$ then by I1 we have $n \in \pi_i(n) \subseteq T$. Then by I3, $\pi_i(n') \subseteq T$, which implies DA.

If $n' \in \pi_i(n) \subseteq T'$ and $n', ..., n''$ is path in T' such that $i \in I_{n'} \cap I_{n''}$ then by I2, $\pi_i(n') = \pi_i(n)$ and thus by DA if $pi_i(n'') \subseteq T''$ then $T'' \succeq T'$. By U0, if $n'' \in T'$ then $\pi_i(n'') \subseteq T''$ with $T'' \preceq T'$. Hence $T'' = T'$, which implies I3. □

A Proofs

A.1 Proof of Lemma 1

(i) By (I3), all information sets of player i along a path starting in h_i and ending at a terminal node are contained in T_{h_i}. Therefore, it is enough to show the claim for every subtree T in the generalized extensive-form game. Since player i's belief system b_i satisfies updating consistency as defined by Perea (2002), the proof of Theorem 3.1 of Perea (2002) implies the claim.[14]

(ii) If a strategy $s_i$of player i is rational at all information sets $h_i \in H_i$, then in particular s_i would be rational in all information sets $h_i \in H_i$ reached by s_i. Denote by $H_i^{-s_i}$ the set of information sets not reached by s_i. By (I3), the expected payoff for player i (given the belief system b_i) from choosing an action in $h_i \in H_i^{-s_i}$ does not depend on her choices at information sets outside T_{h_i}.

Furthermore, $H_i^{-s_i}$ is an arborescence with respect to the precedence relation $\rightsquigarrow$. Hence, a standard backward-induction procedure on $H_i^{-s_i}$yields an optimal action $a_{h_i}^* \in A_{h_i}$ for player i (given b_i) at h_i for every information set $h_i \in H_i^{-s_i}$. Replacing by $a_{h_i}^*$ the action prescribed by $s_i$at h_i for every $h_i \in H_i^{-s_i}$ yields a new strategy $\hat{s}_i$ which would be rational at all information sets $h_i \in H_i$. $\square$

A.2 Proof of Proposition 2

We proceed by induction.

B_i^1 is non-empty. Indeed, to construct a belief system b_i, for each information set h_i with no predecessors (according to the precedence relation $\rightsquigarrow$) in the arborescence of information sets H_i, assign to player i a full-support belief $b_i(h_i)$ on the other players' strategies $S_{-i}^{T_{h_i}}$ that reach h_i. The full-support guarantees that Bayes rule is applicable for deriving the beliefs of player i in all her remaining information sets.

Suppose, by induction, we have already shown that B_i^k is non-empty. We have to show that S_i^k is non-empty. For a typical belief system $b_i \in B_i^k$ we have to construct a strategy with which player i would be rational at each of her information sets H_i. Since H_i is an arborescence, it is standard to construct such a strategy s_i by backward induction.

To complete the induction step, observe that B_i^{k+1} is non-empty, because by definition it singles out a non-empty subset of B_i^k.

Now, since player i's set of strategies S_i is finite and by Remark 3$S_i^{k+1} \subseteq S_i^k$ for every $k \geq 1$, for some ℓ we eventually get $S_i^\ell = S_i^{\ell+1}$ for all $i \in I$ and hence $B_i^{\ell+1} = B_i^{\ell+2}$ for all

[14]Formally, theorem 3.1 in Perea (2002) refers to two-player games, but as he remarks at the top of p. 325, the argument can be extended in a straightforward manner to games with more than two players and correlated beliefs about other players' strategies.

$i \in I$. Inductively,

$$\emptyset \neq S_i^\ell = S_i^{\ell+1} = S_i^{\ell+2} = ...$$

and therefore

$$S_i^\infty = \bigcap_{k=1}^{\infty} S_i^k = S_i^\ell \neq \emptyset$$

as required. □

A.3 Proof of Proposition 4

Denote by (a_i, h_i) the copy of the action a_i of player $i \in I$ whenever it appears in the information set h_i. For the purpose of this proof the word "action" will refer to a copy (a_i, h_i) of an action *at a given information set*.

Define a *menu* of a player to be a (possibly empty) subset of (the union of) her actions in her information sets.

Define a *menu profile* to be a profile of menus, one for each player, with the following property: For each information set h_i of player i, her menu in the menu profile contains at least one action in h_i if and only if that information set is reached by a sequence of actions of the players in the menu profile.

For a menu profile M, denote by M_i the menu of player i in M.

For a menu profile M, denote by $P^T(M)$ all the paths from the roots to leaves in the trees of the T-partial game that one can compose from actions in M and moves of nature (if there are any). Denote also by $P(M)$ the set of paths from roots to leaves in all the trees of the generalized games that one can compose from actions in M and moves of nature.

Now, every product of sets of strategies $R = \prod_{i\in I} R_i$ (where R_i is a subset of i's strategies) induces a menu profile, in which player i's menu is defined as follows. For each information set of the player:

1) If the information set is reached by some strategy profile in the set R, the player's menu contains all the actions ascribed in that information set by i's strategies in R_i that reach the information set.

2) If the information set is not reached by any strategy profile in R, then player i's menu contains no action of hers in that information set.

Intuitively, player i's menu is mute about an information set if and only if that information set is excluded by the set of strategy profiles R (case 2); otherwise (case 1) the menu contains all the actions in that information set that appear in some strategy of hers in R_i that reaches that information set.

If M is the menu profile induced by R, then every strategy in R_i together with a belief about R_{-i} induce a belief β^T about the paths of actions in $P^T(M)$ for every tree T of the generalized game.

Next, denote by M^k the menu profile induced by $S^k = \prod_{i\in I} S_i^k$, the set of level k would-be rationalizable strategy profiles; and denote by $\bar{M}^k$ the menu profile induced by $\bar{S}^k = \prod_{i\in I} \bar{S}_i^k$, the set of level k prudent rationalizable strategy profiles.

Proposition 4 is implied by the following lemma:

Lemma 3 *For all $\ell \geq 0$, $\bar{M}^\ell \subseteq M^\ell$. In particular $\bar{M}^\infty \subseteq M^\infty$.*

Proof. The proof is by induction.

For $\ell = 0$ we have $M^0 = \bar{M}^0$, the menu profile which includes all actions at all the information sets of all the players.

Suppose the claim holds for $\ell \leq k$.

By the induction hypothesis $P(\bar{M}^\ell) \subseteq P(M^\ell)$ for every $\ell \leq k$.

We will now prove the claim for $\ell = k+1$, i.e. that $\bar{M}_i^{k+1} \subseteq M_i^{k+1}$ for every player $i \in I$.

To this end we have to show that for every player $i \in I$, every $\bar{s}_i^{k+1} \in \bar{S}_i^{k+1}$, every information set $h_i \in H_i$ which is reached both by $\bar{s}_i^{k+1}$ and by some strategy profile in $\bar{S}_{-i}^{k+1}$ (meaning that $\bar{s}_i^{k+1}(h_i) \in \bar{M}_i^{k+1}$), it is the case that

a) h_i is also reached by S^{k+1}, and

b) $\bar{s}_i^{k+1}(h_i) \in M_i^{k+1}$ as well.

In fact, it is enough to show that b) holds. To see this, proceed inductively along each feasible path of the generalized game (in each of its trees). If player i is the first to play in this path (apart from nature, if there are nature moves in the path), and if h_i is the information set in which she makes this initial move, then condition a) automatically obtains for h_i, and we only need to prove b). Inductively, if we reach a node in the path which is not in $P\left(\bar{M}^{k+1}\right)$, we have nothing to prove for this node's information set when considering this path.[15] If all the nodes $n_1 \ldots n_m$ in an initial segment of the path

[15] We may have to consider this information set again when we analyze another path passing through

are on a path in $P\left(\bar{M}^{k+1}\right)$ and we have already proved conditions a) and b) for all the information sets of these nodes, then it already follows that a) holds for the information set of the next node n_{m+1} in the path [because b) holds for the previous node n_m for the player (or players) active in n_m]. It thus remains to show b) for such an information set.

So we now proceed to prove b).

Suppose h_i is reached by $\bar{S}_{-i}^{k+1}$ and by $\bar{s}_i^{k+1} \in \bar{S}_i^{k+1}$. Since by definition $\bar{S}_{-i}^{k+1} \subseteq \bar{S}_{-i}^{k}$, we have $\bar{s}_i^{k+1} \in \bar{S}_i^k$ and hence $m_i^{k+1}(h_i) = k$. Consider a belief system $b_i \in \bar{B}_i^{k+1}$ with a full-support belief $b_i(h_i)$ on the strategy profiles $\bar{S}_{-i}^k$ that reach h_i, and with which $\bar{s}_i^{k+1}$ would be rational at h_i (i.e. player i cannot improve her expected payoff by changing $\bar{s}_i^{k+1}$ only at h_i, from $\bar{s}_i^{k+1}(h_i)$ to some other action a'_{h_i} available there).

The strategy $\bar{s}_i^{k+1}$ together with the belief $b_i(h_i)$ on the other players' strategies induce a full support belief β on the paths of actions in $P(\bar{M}^k)$ reaching h_i and along which player i uses the strategy $\bar{s}_i^{k+1}$. Since by the induction hypothesis $P(\bar{M}^k) \subseteq P(M^k)$, it follows that β is a belief on the paths of actions in $P(M^k)$ reaching h_i and along which player i uses the strategy $\bar{s}_i^{k+1}$.

Denote by $\bar{s}_i^{k+1}|a'_{h_i}$ the strategy one gets from $\bar{s}_i^{k+1}$ by altering the action at the information set h_i from $\bar{s}_i^{k+1}(h_i)$ to a'_{h_i}. The altered strategy $\bar{s}_i^{k+1}|a'_{h_i}$ together with the belief $b_i(h_i)$ on the other players' strategies induce a full support belief β' on the paths of actions in $P(\bar{M}^k)$ reaching h_i and along which player i uses the strategy $\bar{s}_i^{k+1}|a'_{h_i}$.

The fact that $\bar{s}_i^{k+1}$ would-be rational given the belief system b_i means that in particular at the information set h_i, with the belief $b_i(h_i)$ on the other players' strategies, the expected payoff to player i given β is not smaller than the expected payoff to player i given β'.

This yields the conclusion b) that we wanted, namely that $\bar{s}_i^{k+1}(h_i) \in M_i^{k+1}$. $\square$

A.4 Proof of Proposition 6

A *general belief system* of player i

$$\tilde{b}_i = (\tilde{b}_i(h_i))_{h_i \in H_i} \in \prod_{h_i \in H_i} \Delta(S_{-i}^{T_{h_i}})$$

---

it.

is a profile of beliefs – a belief $\tilde{b}_i(h_i) \in \Delta(S_{-i}^{T_{h_i}})$ about the other players' strategies in the T_{h_i}-partial extensive-form game, for each information set $h_i \in H_i$, such that $\tilde{b}_i(h_i)$ reaches h_i, i.e., $\tilde{b}_i(h_i)$ assigns probability 1 to the set of strategy profiles of the other players that reach h_i. The difference between a belief system and a general belief system is that in the latter *we do not impose Bayes rule.*

For $k \geq 1$ let $\tilde{B}_i^k$ and $\tilde{S}_i^k$ be defined inductively like $\hat{B}_i^k$, $\hat{S}_i^k$ in Definition 2, respectively, the only change being that belief systems are replaced by generalized belief systems.

Lemma 4 $U_i^k(\mathbf{S}) \cap S_i = \tilde{S}_i^k$ *for* $k \geq 1$. *Consequently,* $U_i^\infty(\mathbf{S}) \cap S_i = \tilde{S}_i^\infty$.

Proof of the Lemma. We proceed by induction. The case $k = 0$ is straight-forward since $U_i^0(\mathbf{S}) \cap S_i = S_i = \tilde{S}_i^0$ for all $i \in I$.

Suppose now that we have shown $U_i^k(\mathbf{S}) \cap S_i = \tilde{S}_i^k$ for all $i \in I$. We want to show that $U_i^{k+1}(\mathbf{S}) \cap S_i = \tilde{S}_i^{k+1}$ for all $i \in I$.

"$\subseteq$": First we show, if $s_i \in U_i^{k+1}(\mathbf{S}) \cap S_i$ then $s_i \in \tilde{S}_i^{k+1}$.

$s_i \in U_i^{k+1}(\mathbf{S}) \cap S_i$ if $s_i \in S_i$ is not conditionally strictly dominated on $(\mathcal{X}_i, U^k(\mathbf{S}))$.

$s_i \in S_i$ is not conditionally strictly dominated on $(\mathcal{X}_i, U^k(\mathbf{S}))$ if for all $T' \in \mathbf{T}$ with $T_1 \hookrightarrow T'$ and all $\tilde{s}_i \in S_i^{T'}$ such that $s_i \in [\tilde{s}_i]$, we have that there does not exist a normal-form information set $X \in \mathcal{X}_i$ with $X \subseteq S^{T'}$ such that $\tilde{s}_i$ is strictly dominated in $X \cap U^k(\mathbf{S})$.

For any information set $h_i \in H_i$, if $\tilde{s}_i \in S_i^{T_{h_i}}$ is not strictly dominated in $S^{T_{h_i}}(h_i) \cap U^k(\mathbf{S})$, then

(i) either $\tilde{s}_i$ does not reach h_i, in which case $\tilde{s}_i$ is trivially rational at h_i; or

(ii) by Lemma 3 in Pearce (1984) there exists a belief $\tilde{b}_i(h_i) \in \Delta(S_{-i}^{T_{h_i}}(h_i) \cap U_{-i}^k(\mathbf{S}))$ for which $\tilde{s}_i$ is rational at h_i. Since by the induction hypothesis $U^k(\mathbf{S}) \cap S = \tilde{S}^k$, we have in this case that there exists a belief at h_i with $\tilde{b}_i(h_i)(\tilde{S}_{-i}^{k,T_{h_i}}) = 1$ for which $\tilde{s}_i$ is rational at h_i.

By definitions of $[\tilde{s}_i]$ and "reach", if $\tilde{s}_i$ reaches h_i in the tree T_{h_i} and $s_i \in [\tilde{s}_i]$, then s_i reaches h_i in the tree T_{h_i}. Hence, if $\tilde{s}_i \in S_i^{T_{h_i}}$ is rational at h_i given $\tilde{b}_i(h_i)$, then $s_i \in [\tilde{s}_i]$ is rational at h_i given $\tilde{b}_i(h_i)$.

We need to show that beliefs in (ii) define a generalized belief system in $\tilde{B}_i^{k+1}$. Consider any $\tilde{b}'_i = (\tilde{b}'_i(h_i))_{h_i \in H_i} \in \tilde{B}_i^{k+1}$. For all $h_i \in H_i$ for which there exists a profile of player i's opponents' strategies $s_{-i} \in \tilde{S}^k_{-i}$ that reach h_i, replace $\tilde{b}'_i(h_i)$ by $\tilde{b}_i(h_i)$ as defined in (ii). Call the new belief system $\tilde{b}_i$. Then this is a generalized belief system. Moreover, $\tilde{b}_i \in \tilde{B}^{k+1}_i$.

Hence, if s_i is not conditionally strictly dominated on $(\mathcal{X}_i, U^k(\mathbf{S}))$ then there exists a generalized belief system $\tilde{b}_i \in \tilde{B}^{k+1}_i$ for which s_i is rational at every $h_i \in H_i$. Thus $s_i \in \tilde{S}^{k+1}_i$.

"$\supseteq$": We show next, if $s_i \in \tilde{S}^{k+1}_i$ then $s_i \in U^{k+1}_i(\mathbf{S}) \cap S_i$.

If $s_i \in \tilde{S}^{k+1}_i$ then there exists a generalized belief system $\tilde{b}_i \in \tilde{B}^{k+1}_i$ such that for all $h_i \in H_i$ the strategy s_i is rational given $\tilde{b}_i(h_i)$. That is, either

(I) s_i does not reach h_i, or

(II) s_i reaches h_i and there does not exist an h_i-replacement of s_i which yields a higher expected payoff in T_{h_i} given $\tilde{b}_i(h_i)$ that assigns probability 1 to T_{h_i}-partial strategies of player i's opponents in $\tilde{S}^{k,T_{h_i}}_{-i}$ that reach h_i in T_{h_i}. By the induction hypothesis, $\tilde{S}^k_{-i} = U^k_{-i}(\mathbf{S}) \cap S^{T_{h_i}}_{-i}$. Hence $\tilde{b}_i(h_i) \in \Delta(U^k_{-i}(\mathbf{S}) \cap S^{T_{h_i}}_{-i}(h_i))$.

If $s_i \in [\tilde{s}_i]$ with $\tilde{s}_i \in S^{T_{h_i}}_i$ and s_i reaches h_i in the tree T_{h_i}, then $\tilde{s}_i$ reaches h_i in the tree T_{h_i}. Hence, if $s_i \in [\tilde{s}_i]$ with $\tilde{s}_i \in S^{T_{h_i}}_i$ is rational at h_i given $\tilde{b}_i(h_i)$, then $\tilde{s}_i$ is rational at h_i given $\tilde{b}_i(h_i)$.

Thus, if s_i is rational at h_i given $\tilde{b}_i(h_i)$, then $\tilde{s}_i \in S^{T_{h_i}}_i$ with $s_i \in [\tilde{s}_i]$ is not strictly dominated in $U^k_{-i}(\mathbf{S}) \cap S^{T_{h_i}}_{-i}(h_i)$ either because s_i does not reach h_i (case (I)), or because of Lemma 3 in Pearce (1984) (in case (II)).

It then follows that if the strategy s_i is rational at all $h_i \in H_i$ given $\tilde{b}_i$ then s_i is not conditionally strictly dominated on $(\mathcal{X}_i, U^k(\mathbf{S}))$. Hence $s_i \in U^{k+1}_i(\mathbf{S}) \cap S_i$. $\square$

Lemma 5 $\tilde{S}^k_i = \hat{S}^k_i$ *for* $k \geq 1$. *Consequently,* $\tilde{S}^\infty_i = \hat{S}^\infty_i$.

Proof of the Lemma. $\hat{S}^k_i \subseteq \tilde{S}^k_i$ for $k \geq 1$ since if s_i is rational at each information set $h_i \in H_i$ given the belief system $b_i \in B_i$ then there exists a generalized belief system $\tilde{b}_i \in \tilde{B}^k_i$, namely $\tilde{b}_i = b_i$, such that s_i is rational at each information set $h_i \in H_i$ given $\tilde{b}_i$.

We need to show the reverse inclusion, $\tilde{S}^k_i \subseteq \hat{S}^k_i$ for $k \geq 1$. The first step is to show how to construct a (consistent) belief system from a generalized belief system. Let s_i be rational given $\tilde{b}_i \in \tilde{B}_i^1$, i.e. $s_i \in \tilde{S}_i^1$. Consider an information set $h_i^0 \in H_i$ such that in T_{h_i} there does not exist an information set h_i that precedes h_i^0. Define $b_i(h_i^0) \equiv \tilde{b}_i(h_i^0)$.

Assume, inductively, that we have already defined b_i for a subset of information sets $H_i' \subseteq H_i$ such that for each $h_i' \in H_i'$ all the predecessors of h_i' are also in H_i'. For each successor information set h_i'' of each information set $h_i' \in H_i'$ such that $h_i'' \notin H_i'$ define $b_i\left(h_i''\right)$ as follows:

- If $b_i\left(h_i'\right)$ reaches h_i'' define $b_i\left(h_i''\right)$ by using Bayes rule, i.e. if $s_{-i}^{T_{h_i'}} \in S_{-i}(h_i'')$

$$b_i\left(h_i''\right)\left(s_{-i}^{T_{h_i'}}\right) = \frac{b_i\left(h_i'\right)\left(s_{-i}^{T_{h_i'}}\right)}{\sum_{\tilde{s}_{-i}^{T_{h_i'}} \in S_{-i}(h_i'')} b_i(h_i')(\tilde{s}_{-i}^{T_{h_i'}})}$$

  and $b_i\left(h_i''\right)\left(s_{-i}^{T_{h_i'}}\right) = 0$ else.

- If $b_i\left(h_i'\right)$ does not reach h_i'' let $b_i(h_i'') \equiv \tilde{b}_i(h_i'')$.

Since there are finitely many information sets in H_i, this inductive definition will be concluded in a finite number of steps.

Next, assuming that s_i is rational at each information set $h_i \in H_i$ with the generalized belief system $\tilde{b}_i$, we will show that s_i is also rational at each information set $h_i \in H_i$ according to the belief system b_i.

Consider again $h_i^0 \in H_i$ with no predecessors in $T_{h_i^0}$. Since $b_i(h_i^0) = \tilde{b}_i(h_i^0)$ and s_i is rational at h_i^0 given $\tilde{b}_i\left(h_i^0\right)$, s_i is also rational at h_i^0 given $b_i\left(h_i^0\right)$.

Assume, inductively, that we have already shown the claim for a subset of information sets $H_i' \subseteq H_i$ such that for each $h_i' \in H_i'$ all the predecessors of h_i' are also in H_i'. Consider a successor information set h_i'' of an information set $h_i' \in H_i'$ such that $h_i'' \notin H_i'$. Notice that each h_i''-replacement is also an h_i'-replacement. Therefore,

- If $b_i\left(h_i'\right)$ reaches h_i'', $b_i\left(h_i''\right)$ is derived from $b_i(h_i')$ by Bayes rule, and hence any h_i''-replacement improving player i's expected payoff according to $b_i\left(h_i''\right)$ would improve player i's payoff also according to $b_i(h_i')$, contradicting the induction hypothesis. Hence s_i is rational at h_i'' given $b_i\left(h_i''\right)$.

- If $b_i\left(h_i'\right)$ does not reach h_i'', then $b_i\left(h_i''\right) = \tilde{b}_i(h_i'')$. Hence, s_i is rational at h_i'' also given $b_i\left(h_i''\right)$.

Applying the same argument inductively yields $\tilde{S}_i^k = \hat{S}_i^k \; \forall k \geq 1$. This concludes the proof of the lemma. $\square$

Lemmata 4 and 5 together yield $U_i^k(\mathbf{S}) \cap S_i = \hat{S}_i^k$ for $k \geq 1$. Since it applies for all $k \geq 1$ and $i \in I$, this completes the proof of the proposition. $\square$

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. and M. Siniscalchi (2002). Strong belief and forward induction reasoning, *Journal of Economic Theory* **106**, 356-391.

[4] Ben Porath, E. and E. Dekel (1992), Signaling future actions and the potential for sacrifice, *Journal of Economic Theory* **57**, 36-51.

[5] Blume, L., Brandenburger, A., and E. Dekel (1991). Lexicographic probabilities and choice under uncertainty, *Econometrica* **69**, 61-79.

[6] Brandenburger, A. and A. Friedenberg (2007). The relationship between rationality on the matrix and the tree, mimeo.

[7] Feinberg, Y. (2004). Subjective reasoning - games with unawareness, mimeo.

[8] Filiz-Ozbay, E. (2007). Incorporating unawareness into contract theory, in: Samet, D. (ed.), Proceedings of the 11th conference on Theoretical Aspects of Rationality and Knowledge, Presses Universitaires de Louvain, 135-144.

[9] Halpern, J. and L. Rêgo (2006). Extensive games with possibly unaware players, in: *Proc. Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, 744-751.

[10] Heifetz, A., Meier, M. and B. C. Schipper (2006). Interactive unawareness, *Journal of Economic Theory* **130**, 78-94.

[11] Herings, P.J.J and V.J. Vannetelbosch (1999). Refinements of rationalizability for normal-form games, *International Journal of Game Theory* **28**, 53-68.

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

[13] Li, J. (2006). Dynamic games with perfect awareness information, mimeo.

[14] Mailath, G., Samuelson, L. and J. Swinkels (1993). Extensive form reasoning in normal form games, *Econometrica* **61**, 273-302.

[15] Ozbay, E. (2007). Unawareness and strategic announcements in games with uncertainty, in: Samet, D. (ed.), Proceedings of the 11th conference on Theoretical Aspects of Rationality and Knowledge, Presses Universitaires de Louvain, pp. 231-238.

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

[17] Perea, A. (2002). A note on the one-deviation property in extensive form games, *Games and Economic Behavior* **40**, 322-338.

[18] Rêgo, L. and J. Halpern (2007). Generalized Solution Concepts in Games with Possibly Unaware Players, in: Samet, D. (ed.), Proceedings of the 11th conference on Theoretical Aspects of Rationality and Knowledge, Presses Universitaires de Louvain, pp. 253-262.

[19] Reny, P. (1992). Backward induction, normal form perfection and explicable equilibria, *Econometrica* **60**, 627-649.

[20] Rubinstein, A. (1991), Comments on the Interpretation of Game Theory, *Econometrica* **59**, 909-924.

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

[22] Stahl, D. (1995). Lexicographic rationalizability and iterated admissibility, *Economics Letters* **47**, 155-159.

[23] Van Damme, E. (1989). Stable equilibria and forward induction, *Journal of Economic Theory* **48**, 476-496.