Forward Induction, Strong Beliefs, and Unawareness in Dynamic Psychological Games*

Carsten S. Nielsen† Alexander Sebald‡

April 26, 2011

Abstract

We provide an epistemic model of dynamic psychological games with unawareness. To formulate epistemic statements, we extend Battigalli and Siniscalchi (1999)’s hierarchies of conditional beliefs by constructing a space of infinite (coherent) hierarchies for each possible awareness level. The interpretation is that only hierarchies of beliefs at the highest awareness level are full descriptions of all relevant beliefs. Hierarchies at lower awareness levels are subjective portraits of situations in the mind of players who are unaware. With this at hand, we suggest a non-equilibrium solution concept which embodies forward induction reasoning in our class of games. To capture forward induction reasoning we extend Battigalli and Siniscalchi (2002)’s notion of strong beliefs and explore implications of forward induction in our framework.

Keywords: Forward induction; Unawareness; Psychological games; Strong belief; Belief-dependent preferences; Extensive-form games.

JEL-Classifications: C72, C73, D80

*We would like to thank Peter Norman Sørensen for helpful comments and suggestions. All errors are our own.
†Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353, Copenhagen K, Denmark. Phone: (+45) 3532-3051. Fax: (+45) 3532-3064. E-mail: carsten.nielsen@econ.ku.dk. Web: http://www.econ.ku.dk/phdstudent/nielsen.
‡Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353, Copenhagen K, Denmark. Phone: (+45) 3532-4418. Fax: (+45) 3532-3064. E-mail: alexander.sebald@econ.ku.dk. Web: http://www.econ.ku.dk/sebald.
1 Introduction

Players are often unaware of some of the details affecting the strategic situation they find
themselves in, and may therefore make choices which are not in their best interest. For
example, players may have asymmetric awareness levels concerning their own as well as
others’ feasible choices, although they are part of the same strategic environment. Quite
naturally, the more complex a strategic situation, the more difficult it is for players to
comprehend all its relevant details.

In contrast to the existing literature on the impact and importance of unawareness in
strategic interactions [see e.g. Fagin and Halpern (1988), Dekel et al. (1998), Modica and
Rustichini (1999), Halpern (2001), Heifetz et al. (2006), Halpern and Rêgo (2008), Heifetz
et al. (2008), Feinberg (2009), Grant and Quiggin (2009), Li (2009) and Heifetz et al. (2010)],
we focus on unawareness in dynamic psychological games [see Battigalli and Dufwenberg
(2009)] in which complexity not only concerns players’ strategic alternatives, but also players’
belief-dependent emotions like reciprocity [see e.g. Dufwenberg and Kirchsteiger (2004)] and
guilt [see e.g. Battigalli and Dufwenberg (2007)]. Dynamic psychological games provide a
framework in which players’ preferences depend upon updated, higher-order beliefs. The
belief-dependent nature of players’ motivations makes it natural to define these games using
epistemics which involves statements about players’ conditional beliefs, their conditional
beliefs about each other’s conditional beliefs, etc. To represent such statements in strategic
environments without unawareness, Battigalli and Siniscalchi (1999) construct a standard
state-space of infinite (coherent) hierarchies of beliefs of conditional probability systems.

As a first contribution, in order to allow for epistemic statements in environments with
unawareness we generalize conditional probability systems to dynamic settings which allow
for unawareness concerning feasible paths of play. More specifically, we present an extensive
form structure that allows for unawareness concerning feasible paths of play [see also Nielsen
and Sebald (2011)] and extend hierarchies of conditional beliefs to it. Our unawareness
framework constitutes a simplification of the generalized extensive form model given by
Heifetz et al. (2010) which allow for, among other things, imaginary actions. We restrict
our analysis to unawareness of observable paths of play. Hence, instead of generalizing the
extensive form, we focus on extending the beliefs defined on our simplified unawareness
structure to embody infinite hierarchies. Such infinite hierarchies are important for that
analysis of psychological games. We proof the existence of such hierarchies of beliefs for
each possible awareness level (Proposition 1). The interpretation is that only hierarchies of
beliefs at the highest awareness level are full descriptions of all relevant beliefs. Hierarchies
at lower awareness levels are subjective portraits of situations in the mind of players who are unaware.

Having described player’s dispositions to act and player’s dispositions to hold beliefs in our unawareness structure, we are able to define states of partial worlds. States are arrays of strategy/belief hierarchy pairs of each player at a given awareness level, and a partial world is a collection of states defining a particular awareness level. It follows from Proposition 1 that a player’s awareness restricted hierarchy of beliefs defines beliefs over the other players’ states at all possible awareness levels he is aware of. Heifetz et al. (2006, 2008) propose a generalized state-space which satisfies the properties of unawareness proposed by Fagin and Halpern (1988), Dekel et al. (1998), Modica and Rustichini (1999) and Halpern (2001). They define partial worlds which differ in their expressiveness, i.e. the richness of the vocabulary by which their states can describe the situation in question. The expressiveness of our partial worlds is given by the richness of the strategic situation and the relevant hierarchies of beliefs. Our states can therefore be seen as a special instance of theirs.

Our epistemic characterization allows us to apply a broad array of solution concepts—including non-equilibrium concepts. Equilibrium analysis ideally involves interpreting hierarchical beliefs as steady states of transparent reasoning processes. Clearly, such an interpretation is very demanding in a dynamic setting with unawareness in which every increase of awareness by definition is a shock or surprise. Once a player’s view of the game itself is challenged in the course of play, it is hard to justify that a new set of equilibrium beliefs for the continuation of the game is readily available. Thus, less demanding non-equilibrium concepts might be more appropriate in dynamic settings in which players have belief-dependent preferences and are unaware of parts of the strategic environment.

As a second contribution we define such a non-equilibrium concept for our class of dynamic psychological games with unawareness. More specifically, we characterize a non-equilibrium solution concept which embodies the forward induction principle: a player should use all information he acquired about the other players’ past behavior in order to improve his prediction of their future, simultaneous, and past (unobserved) behavior, relying on the assumption that they are rational [Battigalli (1996)]. For example, if a player observes an unexpected move, he should revise his beliefs so as to reflect its likely purpose. Of course, in order to figure out the purpose of unexpected moves, a player must formulate assumptions about the other players’ rationality and strategic reasoning. In psychological games players’ behavior is affected by hierarchical beliefs, so other players’ rationality must be defined on a whole state rather than only strategies.
To capture the forward induction principle, we propose a generalized notion of ‘strong beliefs’ based on Battigalli and Siniscalchi (2002)’s definition, and use it for a formal analysis of forward induction reasoning in our class of dynamic psychological games with unawareness. In particular, we say that a player strongly believes an event if he assigns probability one to it, so long as it is consistent with the history he has reached or copies thereof in subforms he is aware of. With this at hand, we show that the conditions that each player is rational, strongly believes that ‘each of the other players is rational’, strongly believes that ‘each of the other players is rational and strongly believes others are rational,’ etc. formally capture the idea of forward induction reasoning in dynamic psychological games with unawareness.

Our analysis proceeds as follows: In section 2 we define our dynamic framework with unawareness and belief-dependent preferences, sequential rationality and conditional belief-operators. Following this, in section 3 we formally characterize strong beliefs and demonstrate forward induction reasoning using an augmented battle-of-the-sexes game with forward induction. Rationality and common strong belief of rationality are discussed in section 4. Sections 5 and 6 respectively contain a discussion of some extensions and related literature as well as a conclusion.

2 The framework

In this section we introduce a class of extensive forms with unawareness (2.1), model an awareness restricted universal belief space (2.2), and put forward definitions of a psychological multi-stage game with unawareness (2.3), sequential rationality (2.4), and conditional belief operators (2.5).

2.1 Extensive forms with unawareness

In order to keep notation at a minimum, our analysis shall deal with multistage games with observable actions and two players, $i$ and $j$.

A finite extensive-form is a tuple $\{i, j\}, H$ where $H$ is the finite set of histories, including the initial history $h_0$, and a set of terminal histories $Z$. Throughout this paper we shall take the point of view of player $i$, similar arguments can trivially be made for player $j$.

Let $C_i$ be the set of all choices player $i$ can make in $H$. A history of length $l \in L$ is a sequence of choices $h = (c^1, \ldots, c^l)$ where each $c^t = (c^t_i, c^t_j)$ represents the pair of choices made

\footnote{Note that all our arguments generalize immediately to the case of more than two players.}
at stage $t$ ($1 \leq t \leq l$). The history $\tilde{h} = (\tilde{c}^1, \ldots, \tilde{c}^k)$ precedes $h = (c^1, \ldots, c^l)$, written $\tilde{h} < h$, if $\tilde{h}$ is a prefix of $h$ (i.e., $k < l$ and $(\tilde{c}^1, \ldots, \tilde{c}^k) = (c^1, \ldots, c^k)$).

Denote by $H$ the family of subforms of $H$, partially ordered $\preceq$ by the inclusion of paths of play. That is,

$$H = \{ H_T \subseteq H : \exists D \in 2^Z \setminus \{ \emptyset \}, H_T = \{ h : \exists z \in D : h \preceq z \} \},$$

where $h \preceq z$ means that $h$ is $z$ or a prefix of $z$. Each subform $H_T \in H$ represents a set of feasible paths of play. The ‘largest’ of these forms is the set $H$ itself. Such a construction of subforms ensures that any $H_T \in H$ starts at the initial history $h^0$, that it is naturally ordered by proper subhistories, and furthermore implies that each terminal history of each subform $z \in D$ is associated with a well defined terminal history in $Z$. Each $H_T$ consists of copies $h_T$ of the histories $h \in H$. More precisely, whenever two histories $h_T = (c^1, \ldots, c^l) \in H_T$ and $h_T' = (c^1, \ldots, c^l) \in H_T'$ share the same path of choices, then we will say that the two are copies of each other. To ease notation we sometimes use $T \in T$ instead of $H_T \in H$ where no confusion can arise.

To model that players may have different views on the set of feasible paths of play in different histories we define players’ perceptions concerning the strategic environment.

**Definition 1.** For player $i$ there exists a perception function:

$$\varphi_i : \left( \bigcup_{T \in T} H_T \right) \rightarrow \left( \bigcup_{T \in T} H_T \right),$$

which defines for each $h_T$ player $i$’s perception $\varphi_i(h_T)$.

The following properties of this function parallel the properties in Heifetz et al. (2006, p. 83) and Heifetz et al. (2010, p. 47):

(i) **Confined Awareness**: If $h_T \in H_T$, then $\varphi_i(h_T) \in H_T'$, with $H_T' \preceq H_T$.

(ii) **Generalized Reflexivity**: If $H_T' \preceq H_T$, $h_T \in H_T$, $\varphi_i(h_T) \in H_T'$ and $H_T'$ contains a copy $h_{T'}$ of $h_T$, then $h_{T'} = \varphi_i(h_T)$.

(iii) **Subforms Preserve Awareness**: If $h_T \in H_T$, $h_T = \varphi_i(h_T)$, $H_T' \preceq H_T$ and $H_T'$ contains a copy $h_{T'}$ of $h_T$, then $h_{T'} = \varphi_i(h_{T'})$.

---

2. We have chosen the term ‘perception function’ instead of possibility correspondence in order to avoid any confounding with settings of imperfect information.

3. Note that introspection does not play a role in our setting as our setting is restricted to observable actions, i.e. singleton information sets.
(iv) **Subforms Preserve Ignorance**: If \( H_T, H_T' \in H \) with \( H_T \prec H_T', h_T, h_T' \in H_T, \varphi_i(h_T) \in H_T' \) and \( H_T' \) contains the copy \( h_{T'} \) of \( h_T \), then \( \varphi_i(h_{T'}) = \varphi_i(h_T) \).

(v) **Subforms Preserve Knowledge**: If \( H_T'' \preceq H_T \preceq H_T', h_T, h_T' \in H_T, \varphi_i(h_T) \in H_T' \) and \( H_T'' \) contains a copy \( h_{T''} \) of \( h_T \), then \( \varphi_i(h_{T''}) \) consists of the copy that exists in \( H_T'' \) of the node \( \varphi_i(h_T) \).

(vi) **Dynamic Awareness**: For any two histories \( h_T, h_T' \in H_T \) directly preceding each other (i.e. \( h_T = (\tilde{h}_T, c) \) and \( \varphi_i(\tilde{h}_T) \in H_T' \), then (i) \( \varphi_i(h_T) \in H_T' \), if \( H_T' \) contains the copy \( h_{T'} \) of \( h_T \), or (ii) \( \varphi_i(h_T) \in H_T'' \) with \( H_T' \preceq H_T'' \).

The perception function and its properties describe for all possible histories the players’ perceptions and change in perceptions about the strategic environment. More specifically, ‘Confined Awareness’ says that the players’ perceptions in some history \( h_T \) is confined to subforms ‘smaller or equal’ to the subform \( h_T \) belongs to. The property of ‘Generalized Reflexivity’ implies that at some history \( h_T \), players know the (observable) choices that have led to it. Properties (iii) – (v) guarantee the coherence of the knowledge and the awareness of players down the partial order. ‘Subforms Preserve Awareness’ implies that if player \( i \) can perceive a history in a subform, then he must also be able to perceive copies of the history in ‘smaller’ subforms. ‘Subforms Preserve Ignorance’ says that at a history in a ‘smaller’ subform player \( i \) cannot perceive anything, that he cannot perceive at copies of the history in ‘larger’ subforms. And ‘Subforms Preserve Knowledge’ means that if player \( i \) perceives to be in a history, then he also perceives the copies of that history in all ‘smaller’ subforms. Finally, the property of ‘Dynamic Awareness’ regards the dynamic nature of the strategic interaction: At a history, player \( i \) perceives to be in a subform which is at least consistent with choices observed.

For any two subforms \( H_T, H_T' \in H \) we (abuse notation slightly and) write \( T \triangleright T' \) whenever for some history \( h_T \in H_T \) it is the case that \( \varphi_i(h_T) \in H_{T'} \). Denote by \( \triangleright \) the transitive closure of \( \triangleright \). That is, \( T \triangleright T'' \) if there is a sequence of trees \( H_T, H_T', ..., H_T'' \in H \) satisfying \( T \triangleright T' \triangleright ... \triangleright T'' \). If \( h_T \in H_T \) but \( T \not\triangleright T' \), then a player may be interpreted as being unaware of histories in \( H_{T'} \). We denote by \( h_T = \{ h_{T'} \}_{T \triangleright T'} \) the ‘historical event’ that a history and copies thereof that a player is aware of obtains. \( H_T \) is the set of such events and \( Z_T \) denotes the set of terminal historical events.

Let the set of choices of player \( i \) in the subform \( H_T \) be given by \( C_i^T \). As the game progresses, each player is informed of the history that has just occurred. The set of feasible choices for player \( i \) in any history \( h_T \) is denoted \( C_i(h_T) \). Player \( i \) is active at \( h_T \in H_T \) if
and only if $C_i(h_T)$ contains more than one element. Players $i$ and $j$ choose simultaneously in any history $h_T$ if both are active at $h_T$.

Moreover, we shall denote by $S_i^{H_T}$ the set of strategies available to player $i$ (where a strategy is defined as a function $s_i^T : H_T \rightarrow (\cup_{h_T \in H_T} C_i(h_T))$ such that $s_i^T(h_T) \in C_i(h_T)$ for all $h_T$). The set of $i$’s strategies that allow for history $h_T$ is denoted $S_i^{H_T}(h_T)$. Since player $i$ considers the set of $j$’s strategies $\bigcup_{T \rightarrow T'} S_j^{H_{T'}}$ as possible, we have that a typical strategy profile is given by $(s_i^T, s_j^{T'}) \in S_i^{H_T} \times \bigcup_{T \rightarrow T'} S_j^{H_{T'}}$. The path function which defines the terminal history $z_T$ induced by $(s_i^T, s_j^{T'})$ is denoted by $\zeta(s_i^T, s_j^{T'}) \in Z_T$. \(^4\) That is, player $i$ evaluates player $j$’s strategies in the subform he thinks he is in. Finally, for every strategy $s_i^T$, we let $H_T(s_i^T) = \{h_T \in H_T \setminus Z_T : s_i^T \in S_i^{H_T}(h_T)\}$ denote the collection of non-terminal histories $h_T \in H_T$ consistent with $s_i^T$.

### 2.2 Belief hierarchies in the unawareness structure

At the beginning of the game, i.e. at the initial history, player $i$ does not know the true strategies of player $j$. He only learns the true strategy by updating his beliefs as the game unfolds. To account for this, we represent beliefs by means of conditional probability systems.

Consider a player who is uncertain about which element in a set $X$ is true. Assume $X$ is a compact Polish space. Each player assigns probabilities to events $E, F, \ldots$ in the Borel sigma-algebra $\mathcal{B}_X$ of $X$ according to some (countably additive) probability measure. Let $\Delta(X)$ denote the set of all probability measures on $(X, \mathcal{B}_X)$. As events unfold players update their beliefs. Let $\mathcal{C} \subseteq \mathcal{B}_X$ be a nonempty, finite or countable collection, such that each $\emptyset \notin \mathcal{B}_X$. The interpretation is that any given player $i$ is uncertain about the element $x \in X$, and $\mathcal{C}$ represents a collection of ‘relevant hypotheses’.

**Definition 2.** A conditional probability system (cps) on $(X, \mathcal{B}_X, \mathcal{C})$ is a mapping $\mu(\cdot|\cdot) : \mathcal{B}_X \times \mathcal{C} \rightarrow [0, 1]$ such that, for all $E \in \mathcal{B}_X$ and $F', F \in \mathcal{C}$, (i) $\mu(\cdot|\cdot) \in \Delta(X)$, (ii) $\mu(F|F) = 1$, and (iii) $E \subseteq F' \subseteq F$ implies $\mu(E|F) = \mu(E|F') \cdot \mu(F'|F)$.

We regard the set of cps’ on $(X, \mathcal{B}_X, \mathcal{C})$ as a subset of the topological space $[\Delta(X)]^\mathcal{C}$ (the set of mappings from $\mathcal{C}$ to $\Delta(X)$) and it is denoted by $\Delta^\mathcal{C}(X)$. Accordingly, we often write $\mu = (\mu(\cdot|F))_{F \in \mathcal{C}} \in \Delta^\mathcal{C}(X)$. $\Delta(X)$ is endowed with the topology of weak convergence of measures (which makes it Polish), and $[\Delta(X)]^\mathcal{C}$ is endowed with the product topology.

\(^4\)The path function $\zeta : S_i^{H_T} \times \bigcup_{T \rightarrow T'} S_j^{H_{T'}} \rightarrow Z_T$ is defined such that $z_T = (c^1, \ldots, c^L) = \zeta(s_i^T, s_j^{T'})$ if and only if $c^1 = (s_i^T, s_j^{T'})$ and $c^t+1 = (s_i^{T, c^1}, c^t, s_j^{T'})$ for all $t \in \{1, \ldots, L-1\}$. 

---

To present player $i$’s higher-order beliefs, we introduce the notion of a hierarchical cps space. Hierarchies of cps’ are in our unawareness structure defined recursively as follows:

$$X^0_{j,T} = S^H_{j,T};$$

for all $k \geq 1$, 

$$X^k_{j,T} = X^{k-1}_{j,T} \times \Delta^\emptyset \left( \bigcup_{T \rightarrow T'} X^{k-1}_{i,T'} \right).$$

A cps $\mu^k_{i,T} \in \Delta^\emptyset \left( \bigcup_{T \rightarrow T'} X^{k-1}_{j,T'} \right)$ is called a $k$-order cps. A hierarchy of cps’ is a countably infinite sequence of cps’ $\mu_{i,T} = (\mu^1_{i,T}, \mu^2_{i,T}, \ldots) \in \prod_{k \geq 1} \Delta^\emptyset \left( \bigcup_{T \rightarrow T'} X^{k-1}_{i,T'} \right)$. Let $B_{i,T}$ be the set of hierarchies of cps’ that are known with common certainty of coherency at the subform $i$ is confined to. The finite disjoint union of Polish spaces is Polish and each $X^k_{j,T}$ is thus a cross-product of compact Polish spaces, hence $B_{i,T}$ is itself a compact Polish space.\(^5\)

We will be interested in the relevant hypotheses, $F \in \mathcal{C}$, that a certain partial history has occurred. Fix some subform $H_T \in \mathcal{H}$. Player $i$’s first-order cps’ about $j$’s behavior in any subform he is aware of may be represented by taking $X = \left( \bigcup_{T \rightarrow T'} S^H_{j,T'} \right)$ and $\mathcal{C} = \left\{ F \subseteq \left( \bigcup_{T \rightarrow T'} S^H_{j,T'} \right) : F = \left( \bigcup_{T \rightarrow T'} S^H_{j,T'}(h_{T'}) \right) \right\}$ for copies $h_{T'}$ of $h_T \in H_T$. Since each element of $\mathcal{C}$ represents the historical event that a history and copies thereof that a player is aware of obtains, we simplify our notation of cps’ on $\left( \bigcup_{T \rightarrow T'} S^H_{j,T'} \right)$ and replace $\mathcal{C}$ with $H_T$. The collection of cps’ on $\left( \bigcup_{T \rightarrow T'} S^H_{j,T'} \right)$ is thus defined by $\Delta^H_T \left( \bigcup_{T \rightarrow T'} S^H_{j,T'} \right)$. Since $\left( \bigcup_{T \rightarrow T'} S^H_{j,T'} \right)$ and $H_T$ are finite, $\Delta^H_T \left( \bigcup_{T \rightarrow T'} S^H_{j,T'} \right)$ is easily seen to be a closed subset of Euclidean $|H_T| \cdot \left| \bigcup_{T \rightarrow T'} S^H_{j,T'} \right|$-dimensional space.

Player $i$’s $T$-partial world describes his dispositions to act (his strategies) and his dispositions to believe (his hierarchies of cps’). Let $i$’s $T$-partial world be denoted by $\Omega_{i,T} = S^H_{i,T} \times B_{i,T}$. The set of states of the $T$-partial world are thus $\omega = (\omega_i, \omega_j) \in \Omega_T = \Omega_{i,T} \times \Omega_{j,T}$ (for a discussion on how to interpret such states see Section 5.1).

Player $i$ also has higher-order cps’ about $j$’s strategies and beliefs in any subform he is aware of. Therefore the structure $(X, \mathcal{C})$ is specified as follows: $X = \bigcup_{T \rightarrow T'} \Omega_{j,T'}$ and

$$\mathcal{C} = \left\{ F \in \mathcal{B} : \left( s^T_{j,T}, \mu_{j,T'} \right) \subseteq \left( \bigcup_{T \rightarrow T'} \Omega_{j,T'} \right) = s^T_{j,T} \in S^H_{j,T'}(h_{T'}) \right\}$$

for the copies $h_{T'}$ of $h_T \in H_T$.\(^6\)

\(^5\)Coherency is the condition that various orders of cps’ of a player cannot contradict each other, i.e. $\mu^k_{i,T}(h_T) = \text{margin}_{X_{j,T'}} \mu^k_{i,T}(h_T)$ for all $k \geq 1$ and $h_T \in H_T$.

\(^6\)If player $i$ is assigned with the lowest level of awareness ($T \neq T'$ for all $T' \in T$ different from $T$), then the hierarchy will be equal to the that provided by Battigalli and Dufwenberg (2009).
The set of cps’ on \((X,\mathcal{C})\) will be denoted by \(\Delta^{HT}\left(\bigcup_{T\sim T'}^{T,T'} \Omega_{j,T'}\right)\) – a compact Polish space.

The following definition establish that countably infinite hierarchies of cps’ are sufficient for the epistemic analysis; a state \(\omega = (s^T_i, \mu_{i,T}, \omega_j)\) in each \(T\)-partial world describes not just the strategies that players choose in this world, but also each player’s entire (awareness restricted) hierarchy of beliefs about the strategies chosen in each of the partial worlds they are aware of, about the other player’s (awareness restricted) beliefs about this, and so on.

**Proposition 1.** For player \(i:\)

\[
f_{i,T} = (f_{i,h_T})_{h_T \in H_T} : B_{i,T} \to \Delta^{HT}\left(\bigcup_{T\sim T'}^{T,T'} \Omega_{j,T'}\right)
\]

is a 1-to-1 and onto continuous mapping whose inverse is also continuous. (Similar for player \(j\).)

**Proof.** See Appendix (A).

We let \(B^{k}_{i,T}\) denote the set of \(k\)-order cps’ consistent with common certainty of coherency, that is, the projection of \(B_{i,T}\) on \(\Delta^{HT}\left(\bigcup_{T\sim T'}^{T,T'} X^{k-1}_{j,T'}\right)\). For example, the set of player \(i\)’s second-order beliefs \((B^2_{i,T})\) are about \(j\)’s strategies and first-order beliefs \((S^{HT'}_j \times B^1_{j,T'}\)\) in any sub-form he is confined to. One might be concerned as to why the isomorphism \(f_{i,h_T}\) is ‘natural’. The reason is that the marginal probability assigned by each \(f_{i,h_T}(\mu^{1}_{i,T}, \mu^{2}_{i,T}, \ldots)\) to a given event in \([\bigcup_{T\sim T', X^{k-1}_{j,T'}}]\) is equal to the probability that \(\mu^{k}_{i,T}\) assigns to that same event. That is, in deriving probabilities on the product space \([\bigcup_{T\sim T', X^{\infty}_{j,T'}}]\) = \([\bigcup_{T\sim T', S^{HT'}_j \times B^1_{j,T'} \times B^2_{j,T'} \times \ldots}\) from \((\mu^{1}_{i,T}, \mu^{2}_{i,T}, \ldots)\), the function \(f_{i,h_T}\) preserves the probabilities specified by \(\mu^{k}_{i,T}\) on each \([\bigcup_{T\sim T', X^{k-1}_{j,T'}}]\).

**Lemma 1.** Each coordinate function \(f_{i,h_T}\) is such that for all \(\mu_{i,T} = (\mu^{1}_{i,T}, \mu^{2}_{i,T}, \ldots) \in B_{i,T}\), and \(k \geq 1:\)

\[
\mu^{k}_{i,T}(h_T) = \text{marg}_{\left[\bigcup_{T\sim T', S^{HT'}_j \times B^1_{j,T'} \times B^2_{j,T'} \times \ldots}\right]} f_{i,h_T}(\mu_{i,T}).
\]

Absent in our definition of a hierarchical cps space is the description of the beliefs of a player about himself. We omit such beliefs. Thus, beliefs about oneself do not play an explicit role. However, our analysis is consistent with the standard assumption that a player knows his beliefs and assigns probability one to the strategy he intends to carry out.

### 2.3 Psychological multi-stage games with unawareness

We are now ready to state our definition of a dynamic psychological game with unawareness:
Definition 3. A dynamic psychological game with unawareness and belief-dependent preferences is a tuple
\[ \Gamma = \left\{ \{i,j\}, \left( \bigcup_{T \in T} \Omega_T \right), \{\varphi_i, \varphi_j\}, \{u_i, u_j\} \right\}, \]
where \( u_i = (u_{i,T})_{T \in T} \) and \( u_{i,T} : Z_T \times B_{i,T} \to \mathbb{R} \) is a continuous psychological payoff function of player \( i \) who is confined to the subform \( H_T \). Player \( j \)'s utility \( u_j \) is defined analogously.

In a game where some players are unaware of some paths of play, other players will, in general, be aware of this possibility. A game with unawareness is therefore not common knowledge among the players, and should be interpreted as the modelers' point of view.

To capture the players' perspective, the standard assumption of common knowledge of the game is replaced with a structure in which each player assigns to others a level of awareness. For this purpose we define partial games as follows:

Definition 4. For any \( \Omega_T \), a \( T \)-partial game is a tuple
\[ G_T = \left\{ \{i,j\}, \Omega_T, \{u_{i,T}, u_{j,T}\} \right\}. \]

From the modelers point of view there exists a set of \( T \)-partial games \( G = \{G_T\}_{T \in T} \), with the partial order \( \preceq \) on \( G \) defined (with slight abuse of notation) by the transitive closure \( \Rightarrow \) generated by the relational requirement \( \gg \) on subforms. Since \( G \) is a finite set of \( T \)-partial games, any 'awareness chain' in \( G \) must have both a minimal element under \( \preceq \), characterized as a strategic situation in which each player thinks that the other player is aware of the same paths of play as himself, and a maximal element under \( \preceq \), namely the partial game which fully describes the complete strategic situation (the modelers game).

2.4 Sequential Rationality

Fix some \( T \)-partial game \( G_T \in G \). Our basic behavioral assumption is that player \( i \) chooses and carries out a strategy \( s_i^T \in S_i^{H_T} \) that is optimal, given his awareness and his beliefs, conditional upon any history consistent with \( s_i^T \). This does not impose restrictions on the choices specified at histories that cannot obtain if player \( i \) follows strategy \( s_i^T \). Thus, we use a sequential best response property which applies to plans of choices\(^7\) as well as strategies\(^8\).

---

\(^7\)If we were to make such a common knowledge assumption here, then the domain and codomain of the perception function \( \varphi_i \) would become the same for all players. The game would therefore just be a Battigalli and Dufwenberg (2009, Definition 4) game.

\(^8\)Intuitively, a plan of choices for player \( i \) is silent about which actions would be taken by \( i \) if \( i \) did not follow that plan. Formally, a plan of choices is a class of realization-equivalent strategies.
We will say that player \( i \) is sequentially rational at state \( (s^T_T, \mu_{i,T}, \omega_{-i}) \) if and only if \( s^T_T \) is a best response to his conditional expected utility

\[
\mathbb{E}_{s^T_T, \mu_{i,T}}[u_{i,T}|h_T] := \sum_{T \rightarrow T'} \mu^1_{i,T}(h_{T'}|h_T) \times \sum_{s^T_{j'} \in S^H_{i,T'}(h_{T'})} \mu^1_{j,T}(s^T_{j'}|h_{T'}) u_{i,T}(\zeta(s^T_{i}, s^T_{j}), \mu_{i,T}),
\]

for all histories \( h_T \in H_T(s^T_i) \) and copies thereof in \( h_T \).

This expression gives the expected payoff from the strategies of \( j \) he is aware of. However, player \( i \) does—in general—not know the awareness and strategies of player \( j \) and thus evaluates his payoff with respect to his first-order belief. Here we first use the idea that the event \( F' = h_{T'} \) is a subset of the event \( F = h_T \) (for \( F', F \in \mathcal{C} \)) such that \( \mu^1_{i}(\cdot|h_{T'}) = \mu^1_{i}(\cdot|h_{T}) \mu^1(\cdot|h_{T'}|h_T) \), and then the fact that the sets \( S^H_{j,T'} \) are disjoint.

**Definition 5.** Fix a cps \( \mu_{i,T} \in B_{i,T} \). A strategy \( s^T_i \in S^{H_T} \) is a sequential best response to \( \mu_{i,T} \) if and only if for every history \( h_T \in H_T(s^T_i) \) and copies thereof in \( h_T \):

\[
s^T_{i,*} \in \arg \max_{s^T_i \in S^{H_T}(h_T)} \mathbb{E}_{s^T_i, \mu_{i,T}}[u_{i,T}|h_T].
\]

For any cps \( \mu_{i,T} \in B_{i,T} \), let \( BR_{i,T}(\mu_{i,T}) \) denote the set of strategies \( s^T_{i,*} \in S^{H_T}_i \) such that \( s^T_{i,*} \) is a sequential best response to \( \mu_{i,T} \). Clearly, \( BR_{i,T} \) is nonempty valued. The belief function \( f_{i,T} \) is continuous (Proposition 1). Since \( u_{i,T} \) is also continuous, then \( \mathbb{E}_{s^T_i, \mu_{i,T}}[u_{i,T}|h_T] \) is continuous (in \( \mu_{i,T} \)), which implies that \( BR_{i,T} \) is an upper hemicontinuous correspondence.

We say player \( i \) is rational at a state \( \omega \) if and only if

\[
\omega \in R_{i,T} = \left\{(s^T_{i,*}, \mu_{i,T}, \omega_{-i}) : s^T_{i,*} \in BR_{i,T}(\mu_{i,T}) \right\}.
\]

Note that \( R_{i,T} \) is closed because the correspondence \( BR_{i,T} \) is upper hemicontinuous. Hence \( R_{i,T} \) is an event about player \( i \). We shall also refer to the events \( R_T = R_{i,T} \cap R_{j,T} \) ('both players confined to the \( T \)-partial game are rational') and \( R_{j,T} \) ('player \( j \) confined to the \( T \)-partial game is rational').

---

9Hence, our analysis could be carried out in a more parsimonious (but less conventional) formal setup, wherein each player’s behavior at a state is described by a plan of choices.
2.5 Conditional belief operators

We now formalize the idea that a player is conditionally certain (has probability-one belief) at a state. This allows us to make statements such as, ‘player $i$ would be certain that player $j$ is rational, were he to observe the historical event $h_T$.’

Fix a $T$-partial game $G_T \in G$. An event in a $T$-partial world is a Borel subset $E_T \subseteq \Omega_T$. An event about player $i$ is any set of states $E_T = E_{i,T} \times \Omega_{j,T}$, where $E_{i,T}$ is a Borel subset of $\Omega_{i,T}$. Let $\mathcal{E}_i$ denote the family of events about $i$. Events about player $j$ is similarly defined; the collection of such is denoted $\mathcal{E}_j$. At state $(s^T_i, \mu_{i,T}, \omega_j)$, player $i$ would believe the collection of events

$$\{ (s^T_i, \mu_{i,T}, \omega_j) : f_{i,h_T}(\mu_{i,T})(\bigcup_{T \subseteq T'} E_{j,T'}) = 1 \}$$

is the event ‘$i$ would be certain that the collection $\bigcup_{T \subseteq T'} E_{j,T'}$ obtains conditional on $h_T$.’ Each $E_{T'}$ may concern the beliefs of player $j$ in some $T'$-partial world.

The conditional (probability-one) belief operator for player $i$ given $h_T \in H_T$ is a mapping $B_{i,h_T} : (\bigcup_{T \subseteq T'} \mathcal{E}_{j,T'}) \mapsto \mathcal{E}_{i,T}$ defined such that for any $E_{T'} \in \mathcal{E}_{j,T'}$,

$$B_{i,h_T} \left( \bigcup_{T \subseteq T'} E_{T'} \right) = \left\{ (s^T_i, \mu_{i,T}, \omega_j) : f_{i,h_T}(\mu_{i,T}) \left( \bigcup_{T \subseteq T'} E_{j,T'} \right) = 1 \right\}.$$

The conditional belief operator is an event $B_{i,h_T} \in \mathcal{E}_{i,T}$. The function $B_{i,h_T} \cdot$ satisfies some standard properties of falsifiable beliefs, in particular monotonicity ($F' \subseteq F$ implies $B_{i,h_T}(F') \subseteq B_{i,h_T}(F)$) and conjunction ($B_{i,h_T}(F' \cap F) = B_{i,h_T}(F') \cap B_{i,h_T}(F)$).

3 Forward induction reasoning

In its simplest form, forward induction reasoning involves the assumption that, upon observing an unexpected (but rational) move of player $j$, player $i$ maintains the ‘working hypothesis’ that $j$ is rational. To capture such a notion of forward inductive reasoning we adapt Battigalli and Siniscalchi (2002)’s concept of strong beliefs in the rationality of others which precisely captures this type of argument (3.1). We then go on to illustrate, by example, how strong beliefs induce forward induction in dynamic psychological games with unawareness (3.2).
3.1 Strong beliefs

Formally, given a $T$-partial game $G_T \in G$ player $i$ strongly believes that an event $E_{T'} \neq \emptyset$ is true (i.e., adopts $E_{T'}$ as a ‘working hypothesis’) if and only if he is certain of $E_{T'}$ at all $h_T$ consistent with $E_{T'}$. We define a ‘strong belief operator’ $SB_{i,T} : (\cup_{T \rightarrow T'} E_{T'}) \rightarrow \mathcal{E}_{i,T}$ by $SB_{i,T}(\emptyset) = \emptyset$ and for any $E_{T'} \in \mathcal{E}_{j,T'}$,

$$SB_{i,T}\left(\bigcup_{T \rightarrow T'} E_{T'}\right) = B_{i,h_T}\left(\bigcap_{h_T \in H_T - \emptyset} B\left(\bigcup_{T \rightarrow T'} E_{T'}\right)\right),$$

for any collection of events such that $E_{T'} \in \mathcal{E}_{j,T'} - \emptyset$, where $[h_T] = \left(\bigcup_{T \rightarrow T'} S_i^{H_{T'}}(h_{T'}) \times B_i(T')\right)$ is the event ‘$h_T$ occurs’. In words, $SB_{i,T}(\cup_{T \rightarrow T'} E_{T'})$ is the event ‘player $i$ would be certain that the collection $\cup_{T \rightarrow T'} E_{T'}$ obtains conditional on every $h_T$ that does not contradict $\cup_{T \rightarrow T'} E_{T'}$. $SB_{i,T}(\cdot)$ is not a monotone operator and satisfies only a weak form of conjunctiveness ($SB_{i,T}(F') \cap SB_{i,T}(F) \subseteq SB_{i,T}(F' \cap F)$).

We will use brackets to denote specific events.\textsuperscript{11} For example, $[s_i^{T,*}] = \{(s_i^T, \mu_i(T, \omega) : s_i^T = s_i^{T,*})\} \in \mathcal{E}_{i,T}$ is the event ‘$i$ plays $s_i^{T,*}$ at $\omega$ in the $T$-partial world.’ $SB_{i,T}\left([s_i^{T,*}]\right)$ is thus the event ‘player $i$ would believe that player $j$ plays the set of optimal strategy $[s_j^{T,*}]$ at each $h_T$ allowed by $[s_i^{T,*}]$’.

Finally, we shall be interested in formalizing assumptions such as ‘player $i$ and player $j$ in the $T$-partial world strongly believe that the other is rational, were they to observe $h_T$.’ In order to simplify notation, we introduce an additional ‘mutual strong belief’ operator. For any Borel subset $E_{T'} \subseteq \Omega_{T'}$ such that $E_{T'} = E_{i,T'} \times E_{j,T'}$, and for any historical event $h_T$, let

$$SB_T\left(\bigcup_{T \rightarrow T'} E_{T'}\right) = SB_{i,T}\left(\bigcup_{T \rightarrow T'} \Omega_i(T') \times E_{j,T'}\right) \cap SB_{j,T}\left(\bigcup_{T \rightarrow T'} \Omega_i(T') \times E_{i,T'}\right).$$

Let $\mathcal{E}_T$ denote the collection of events of the form $E_T = E_{i,T} \times E_{j,T}$. If $E_{T'} = R_{T'}$, then $SB_T(\cup_{T \rightarrow T'} R_{T'}) = SB_{i,T}(\cup_{T \rightarrow T'} R_{j,T'}) \cap SB_{j,T}(\cup_{T \rightarrow T'} R_{i,T'})$. Since $SB_{i,T} \in \mathcal{E}_{i,T}$ and $SB_{j,T} \in \mathcal{E}_{j,T}$, it follows that $SB_T \in \mathcal{E}_T$

3.2 Augmented battle-of-the-sexes and reciprocity

The features of strong beliefs and forward induction in our setting are best illustrated using an example: a battle-of-the-sexes game, augmented by a dominant choice for one player, with reciprocity and an outside option (see Figure 1).\textsuperscript{11} In particular, for any function (random variable) $x : \Omega_T \rightarrow X$ and value $x^* \in X$, we use the notation $[x = x^*] = \{\omega : x(\omega) = x^*\}$. When $x$ is understood, we simply write $[x^*].$
In this game Ann first decides whether she and Bob should go Out to have dinner at a restaurant, or stay In and make their own. If she decides that they should go Out, they can each choose their own dish and the game ends; if she decides that they should stay In and make their own dinner, then she is engaged in an augmented battle-of-the-sexes game with Bob. In this game Ann and Bob prefer to prepare just one dish, but Bob does not share her preference. More specifically, Bob always prefers an Indian dish–Thali–he just learned about, while Ann prefers Salad over the new Thali over Bob’s old favorite Burger.

Assume that Ann and Bob are concerned about their payoffs, but Ann, in addition, cares about being treated kindly or unkindly by Bob. In particular, if Bob seems to be kind to Ann, she will be kind to him; if Bob seems to be unkind, she will be unkind too. To formally model this type of reciprocal behavior in our example with possible unawareness, we use the model of sequential reciprocity defined in Nielsen and Sebald (2011). In this model it is assumed that Ann evaluates Bob’s kindness on the basis of her own awareness which, in turn, implies that Ann does not hold Bob’s unawareness against him.\footnote{This specification is take from Nielsen and Sebald (2011)’s application.}

This means, Ann’s utility is given by:

\[
u_{AT}(\zeta(s^T_A, s^T_B, \mu_A)) = \pi_A(\cdot) + Y_A \times \kappa_{AB}(\cdot) \times \lambda_{ABA}(\cdot),\]

where \(s^T_A \in S^H_A\) and \(s^T_B \in \bigcup_{T \in T'} S^H_B\). The first term is Ann’s material payoff and the remaining terms are her reciprocity payoff with respect to Bob. More specifically, \(\pi_A(\cdot)\) is Ann’s expected monetary payoff which depends on her first-order belief concerning Bob’s...
strategy $\mu_{A,T}^1(s_B^T)$ and her own strategy $s_A^T$. Thus, at $h_T$ Ann’s expected monetary payoff is given by $\pi_A(\mu_{A,T}^1(s_B^T|h_T), s_A^T)$. The exogenous constant $Y_A > 0$ measures how sensitive Ann is to reciprocity. Ann’s perception of Bob’s kindness had Bob been aware of the same as her is given by $\lambda_{ABA}(\cdot)$. Furthermore, Ann’s belief about her kindness towards Bob is $\kappa_{AB}(\cdot)$.

More formally, Ann’s perception of Bob’s kindness towards her at $h_T \in H_T$ is:

$$\lambda_{ABA}(\cdot) = \pi_A(\mu_{A,T}^1(s_B^T|h_T), \mu_{A,T}^2(\cdot|h_T)) - \pi_A^{e_B}(\mu_{A,T}^1(s_B^T|h_T), \mu_{A,T}^2(\cdot|h_T)),$$

where $\mu_{A,T}^1(s_B^T|h_T)$ and $\mu_{A,T}^2(\cdot|h_T)$ respectively are Ann’s (updated) first- and second-order beliefs conditional on $h_T$ in $h_T$. Given this, $\pi_A(\cdot)$ and $\pi_A^{e_B}(\cdot)$ respectively describe what Ann believes Bob would intend for her and the average that Bob would be able to give had he the same awareness level as Ann. The equitable payoff is formally defined as follows:

$$\pi_A^{e_B}(\cdot) = \frac{1}{2} \left( \max \left\{ \pi_A(\mu_{A,T}^1(s_B^T|h_T), \mu_{A,T}^2(\cdot|h_T)), s_B^T \in S_B^{H_T} \right\} + \min \left\{ \pi_A(\mu_{A,T}^1(s_B^T|h_T), \mu_{A,T}^2(\cdot|h_T)), s_B^T \in S_B^{H_T} \right\} \right).$$

The first term in the brackets, $\max \{ \pi_A(\mu_{A,T}^1(s_B^T|h_T), \mu_{A,T}^2(\cdot|h_T)), s_B^T \in S_B^{H_T} \}$, describes Ann’s belief about Bob’s belief about the maximum that he could have given to Ann. Whereas, $\min \{ \pi_A(\mu_{A,T}^1(s_B^T|h_T), \mu_{A,T}^2(\cdot|h_T)), s_B^T \in S_B^{H_T} \}$ describes the minimum.

Lastly, Ann’s kindness towards Bob at $h_T$ can be described as:

$$\kappa_{AB}(\cdot) = \pi_B(\mu_{A,T}^1(s_B^T|h_T), s_A^T) - \pi_A^{e_A}(\mu_{A,T}^1(s_B^T|h_T), s_A^T)$$

Bob’s expected material payoff $\pi_B(\cdot)$ describes what Ann believes Bob gets, given her beliefs concerning Bob’s strategy $s_B^{'T}$ and his own strategy $s_A^T \in S_A^{H_T}$ where $S_A^{H_T}$ is the set of own strategies that Ann is aware of in the historical event $h_T$. Furthermore, $\pi_A^{e_A}(\cdot)$ is Ann’s belief about the average that she can give to Bob. The equitable payoff $\pi_A^{e_A}(\cdot)$ is defined analogously to $\pi_A^{e_B}(\cdot)$ defined above.

### 3.2.1 Full awareness in the augmented battle-of-the-sexes game

To highlight the role of forward induction in unawareness structures, we will start off by analyzing a benchmark scenario with full awareness of Ann and Bob.

In our forward induction analysis, we need to be able to distinguish the event ‘Bob chooses
Thali’, which in the extensive form implies that ‘Ann decides that they should stay In’, from the event ‘Bob would choose Thali if Ann decides they should stay In’ which is a subjunctive conditional, logically independent of whether Ann chooses In or not. Similar considerations hold for Salad and Burger. We use bold letters to denote subjunctive conditionals (which in this case corresponds to strategies of Bob), as in Thali, Salad and Burger.

It is easiest to start the forward induction analysis by looking at the set of states of the world in which Bob is rational. As Bob is only interested in his own payoff it holds that:

$$R_B = \{(s_B, \mu_B, \omega_A) : s_B = \text{Thali}\}.$$  

Given this, the states of the world in which Ann is rational and strongly believes in Bob’s rationality are:

$$R_A \cap SB_A(R_B) = \{(s_A, \mu_A, \omega_B) :$$

for $\mu_A^2$ such that $\lambda_{ABA} > 0$ and $Y_A > \frac{1}{3\lambda_{ABA}} \Rightarrow s_A = (\text{In, Thali}),$

for $\mu_A^2$ such that $\lambda_{ABA} > 0$ and $Y_A \leq \frac{1}{3\lambda_{ABA}} \Rightarrow s_A = (\text{Out, } \cdot)$,

and for $\mu_A^2$ such that $\lambda_{ABA} \leq 0 \Rightarrow s_B = (\text{Out, } \cdot).\}$$

In words, given that Bob chooses Thali, Ann rationally chooses (In, Thali) only if she believes that Bob is kind to her by choosing Thali and her sensitivity to reciprocity is sufficiently high. It follows that if Ann is not motivated enough by reciprocity, she decides that they go Out to have dinner at a restaurant and the game ends. However, if she is motivated enough by reciprocity and she believes that Bob behaves kindly by choosing Thali, she chooses In.

Upon observing In, Bob is certain that Ann is rational and that her strategy is (In, Thali) (i.e. $\mu_B^1((\text{In, Thali})|\text{In}) = 1$). In other words, the states of the world at which Bob strongly believes in the rationality of Ann are:

$$SB_B(R_A) \notin [\mu_B^1((\text{In, Thali})|\text{In}) = 1].$$

Furthermore, the states of the world at which Bob is rational and strongly believes in the rationality of Ann are:

$$R_B \cap SB_B(R_A) \subseteq R_B \cap [\mu_B^1((\text{In, Thali})|\text{In}) = 1] \subseteq [\text{Thali}].$$

Bob is certain that Ann’s strategy is ((In, Thali)|In) and his rational response is to choose
Thali. If Ann now strongly believes that Bob is rational and that he strongly believes that Ann is rational, she infers that Bob’s would choose Thali and that Bob is certain that her strategy when he gets to choose is $(In, Thali)$, i.e. $\mu_A^1(Thali) = 1$, and $\mu_A^2((In, Thali)|In) = 1$. We obtain:

$$R_A \cap SB_A(R_B \cap SB_B(R_A)) \subseteq R_A \cap SB_A([Thali]) \subseteq [(In, Thali)]$$

Ann’s reasoning implies that her perception concerning Bob’s kindness towards her is $\lambda_{ABA} = 1$. Given this, Ann chooses to go Out if her sensitivity to reciprocity is $Y_A < \frac{1}{3}$ and chooses to stay In and cook Bob’s favorite meal Thali, if her sensitivity to reciprocity is above $\frac{1}{3}$. In other words, we identify strategy profiles $((In, Thali), Thali)$ and $((Out, \cdot), Thali)$ depending on Ann’s sensitivity to reciprocity using forward induction reasoning.

### 3.2.2 Asymmetric awareness in the augmented battle-of-the-sexes game

Now consider the following variant of our example with asymmetric awareness: Ann is unaware of the dish Thali. That is, Ann is unaware of Bob’s new favorite meal and believes they are playing a standard battle-of the sexes game with outside option. In addition, Bob is certain that Ann has never heard of Thali. (See Figure 2.)
Formally, Bob is aware of all the subforms \( H_T \geq H_{T'} \), but Ann is at all non-terminal histories only aware of the subform \( H_{T'} \) and all its subforms. For simplicity assume now that \( Y_A = \frac{1}{4} \). Remember, in the previous setting with full awareness \( Y_A = \frac{1}{4} \) implied that Ann would choose to go Out to have dinner in a restaurant.

We will start the forward-induction reasoning by considering the ‘smallest’ subform \( H_{T'} \). In this subform Ann is rational at the following states of the world:

\[
R_{A,T'} = \{ (s_A^{T'}, \mu_{A,T'}, \omega_B) : \\
\text{for } \mu_{1}^{A,T'} \text{ and } \mu_{2}^{A,T'} \text{ such that } u_{A,T'}(\text{Out}, \cdot) \leq u_{A,T'}(\text{In}, \cdot) \Rightarrow s_A^{T'} = (\text{In}, \text{Salad}), \\
\text{for } \mu_{1}^{A,T'} \text{ and } \mu_{2}^{A,T'} \text{ such that } u_{A,T'}(\text{Out}, \cdot) > u_{A,T'}(\text{In}, \cdot) \Rightarrow s_A^{T'} = (\text{Out}, \cdot) \}.
\]

If Bob observes In, he is certain that rational Ann’s strategy is \((\text{In}, \text{Salad})\) (i.e. \(\mu_{I}^{B}((\text{In}, \text{Salad})|\text{In}) = 1\)). Intuitively, the states of the world at which Bob strongly believes in the rationality of Ann are:

\[
SB_{B,T'}(R_{A,T'}) \subseteq [\mu_{I}^{B,T'}((\text{In}, \text{Salad})|\text{In}) = 1].
\]

Furthermore, the states of the world at which Bob is rational in subform \( H_{T'} \) are

\[
R_{B,T'} = \{ (s_B^{T'}, \mu_{B,T'}, \omega_A) : \\
\text{for } \mu_{1}^{B}((\text{In}, \text{Salad})|\text{In})) > \frac{4}{5} \Rightarrow s_B^{T'} = \text{Salad} \\
\text{and for } \mu_{1}^{B}((\text{In}, \text{Salad})|\text{In})) \leq \frac{4}{5} \Rightarrow s_B^{T'} = \text{Burger} \}.
\]

Forward-induction reasoning implies that in \( H_{T'} \) Bob strongly believes in Ann’s rationality:

\[
R_{B,T'} \cap SB_{B,T'}(R_{A,T'}) \subseteq R_{B,T'} \cap [\mu_{I}^{B,T'}((\text{In}, \text{Salad})|\text{In}) = 1] \subseteq [\text{Salad}].
\]

Bob’s rational response in subform \( H_{T'} \) is to choose Salad. Hence, Ann’s forward-induction reasoning in subform \( H_{T'} \) leads to:

\[
R_{A,T'} \cap SB_{A,T'}(R_{B,T'} \cap SB_{B,T'}(R_{A,T'}) \\
\subseteq R_{A,T'} \cap [\mu_{1}^{A,T'}(\text{Salad}) = 1, \mu_{2}^{A,T'}(\text{In}, \text{Salad}) = 1] \\
\subseteq [(\text{In}, \text{Salad})].
\]

That is, Ann who is aware only of \( H_{T'} \) and all subforms thereof will rationally choose \((\text{In}, \text{Salad})\).
What is Bob’s rational behavior in the ‘larger’ subform $H_T$? As Bob only cares about his own payoff and he is aware of everything, he will (trivially) choose Thali following Ann’s choice $In$.

$$R_{B,T} \cap SB_{B,T}(R_{A,T}) \subseteq [\text{Thali}].$$

Hence, in this situation with asymmetric awareness and $Y_A = \frac{1}{4}$ forward induction reasoning leads to $((In, Salad), \text{Thali})$ rather than $((Out, \cdot), \text{Thali})$ as in the case of full awareness.

Intuitively, in the setting with asymmetric awareness the payoffs that Ann is aware of induce a rather selfish move of Ann, i.e. given her sensitivity to reciprocity she does not want to give up 1 unit of her payoff (3 ($Out, \cdot$) vs 4 ($In, Salad$)) to give Bob additional 2 units of payoff (3 ($Out, \cdot$) vs 1 ($In, Salad$)). Similarly, in the situation with full awareness Ann chooses out given $Y_A = \frac{1}{4}$ because she does not want to give up 1 unit of her payoff (3 ($Out, \cdot$) vs 2 ($In, Thali$)) in order to give Bob additional 3 units of payoff (3 ($Out, \cdot$) vs 6 ($In, Thali$)). Irrespective of the fact that the reasoning is the same, forward induction leads to completely different predictions in case of full awareness compared to our scenario with asymmetric awareness.

### 4 Rationality, strong beliefs and forward induction

In this section we formalize forward induction reasoning (Section 3) using the idea of correct strong beliefs. Recall that a player strongly believes an event if he assigns probability one to it, so long as the event is consistent with the history he has reached, or copies thereof. With this, the conditions that each player is rational, strongly believes others are rational, etc. formally captures the idea of forward induction reasoning.

A player believes that the other player is rational as long as this is consistent with his awareness and observed behavior. More generally, a player bestows on the other player the highest degree of ‘strategic sophistication’ consistent with observed behavior. This ‘best rationalization principle’ captures forward induction [see Battigalli (1996, 1997)].

For player $i$, let $R_{i,T}^1$ be the set of states at which $i$ is rational in the $T$-partial game, and for $m > 0$ let $R_{i,T}^{m+1}$ be defined inductively such that:

1. player $i$ is rational: $R_{i,T}^0 = R_{i,T}$;

2. player $i$ is (0) and strongly believes in $j$’s (0): $R_{i,T}^1 = R_{i,T}^0 \cap SB_{i,T}(\cup_{T \rightarrow T'} R_{j,T'}^0)$;
(2) player \( i \) is (1) and strongly beliefs in \( j \)’s (1): \( R^2_{i,T} = R^1_{i,T} \cap SB_{i,T} \left( \bigcup_{T \to T'} R^1_{j,T'} \right) \); \\
(\vdots)

(m) player \( i \) is \((m-1)\) and strongly beliefs in \( j \)’s \((m-1)\): \( R^m_{i,T} = R^{m-1}_{i,T} \cap SB_{i,T} \left( \bigcup_{T \to T'} R^{m-1}_{j,T'} \right) \);

\( \vdots \)

**Definition 6.** Denote mutual rationality of the \( m \)th-order by \( R^m_T = R^m_{i,T} \cap R^m_{j,T} \). If \( \omega \in R^m_T \), say there is rationality and \( m \)th-order strong belief of rationality in the \( T \)-partial game at this state. If \( \omega \in \bigcap_{m=0}^\infty R^m_T \), say there is rationality and common strong belief of rationality at this state.

The set \( R^m_T \) inherits the non-monotonicity property from the strong belief operator. However, since \( R^m_T \subseteq R_T \), the iterations of \( R^m_T \) gives rise to a weakly decreasing sequence of events. Therefore \( \bigcap_{l=0}^m R^l_T \neq \emptyset \) whenever \( R^l_T \neq \emptyset \) for each \( l = 0, \ldots, m \), and we do not have any problems with inconsistent iterated strong beliefs (see Battigalli and Siniscalchi (2002, Section 4.1)).

For every \( m \geq 0 \), the event \( R^m_T \) is associated with \( m \)-th order strategic sophistication. A minimally sophisticated player is simply rational:

\[ R^0_T = R_{i,T} \cap R_{j,T} = R_T. \]

A first-order strategically sophisticated player is rational, and maintains whenever possible the hypothesis that the other player is rational:

\[
R^1_T = R^0_{i,T} \cap SB_{i,T} \left( \bigcup_{T \to T'} R^0_{j,T'} \right) \cap R^0_{j,T} \cap SB_{j,T} \left( \bigcup_{T \to T'} R^0_{i,T'} \right) \\
= R^0_T \cap SB_T \left( \bigcup_{T \to T'} R^0_{T'} \right).
\]

Continuing inductively, a \((m > 1)\)-order strategically sophisticated player is rational, and maintains whenever possible the hypothesis that the other player is \( m-1 \) strategically sophisticated:

\[
R^m_T = R^{m-1}_{i,T} \cap SB_{i,T} \left( \bigcup_{T \to T'} R^{m-1}_{j,T'} \right) \cap R^{m-1}_{j,T} \cap SB_{j,T} \left( \bigcup_{T \to T'} R^{m-1}_{i,T'} \right) \\
= R^0_T \cap SB_T \left( \bigcup_{T \to T'} R^0_{T'} \right) \cap \cdots \cap SB_T \left( \bigcup_{T \to T'} R^{m-1}_{T'} \right). \]

20
To see that this captures forward induction reasoning, recall the augmented battle-of-the-sexes with asymmetric awareness in Section 3. In particular, Bob’s solution in the ‘larger’ $T$-partial game he is confined to requires rationality and strong beliefs in rationality (the solution obtain at all states $\omega \in R_{T}^{1}$), whereas Ann’s solution in her ‘smaller’ $T'$-partial game requires two layers of mutual common strong belief (the solution obtains at all states $\omega \in R_{T'}^{2}$). Remember, she is certain that Bob is confined to the same partial game as she is, such that $R_{T'}^{2} = R_{T'}^{0} \cap SB_{T'}(R_{T'}^{0}) \cap SB_{T'}(R_{T'}^{0} \cap SB_{T'}(R_{T'}^{0})) = R_{T'} \cap SB_{T'}(R_{T'} \cap SB_{T'}(R_{T'}))$.

5 Discussion

In this section we discuss how states of partial worlds should be interpreted (5.1), how imperfectly observably choices can be added to our framework (5.2), and finally we discuss some related literature on forward induction (5.3).

5.1 States of partial worlds

The general framework we propose builds on recent progress in dynamic interactive epistemology. That is, the primitives of the game are given in terms of what the players know and believe–about the game, and about each other’s rationality, strategies, knowledge, and beliefs.

An epistemic model is not a prescriptive model—it does not suggest strategies to the players. Rather, it is a formal framework for talking about strategies, beliefs, and payoffs. For example, it enables us to say whether player $i$ is rational at a given state, whether this is known to player $j$, and so on. But it does not prescribe or suggest rationality, the players simply do what they do.

We introduce states of the $T$-partial worlds as a convenient way of talking about what players would do and believe given their awareness. A state, $\omega = (s_{i}^{T}, \mu_{i,T}, s_{j}^{T}, \mu_{j,T})$, describes the players strategies and hierarchy of beliefs in the $T$-partial game. This is a more elaborate definition of a state than what is used in a single decision maker problem. In Savage (1954), the decision maker cannot affect the state, he can only react to it. While this is convenient in Savage’s one-person context, it is not appropriate for the psychological games we study here. To describe the state of such a world, it is fitting to consider all the players simultaneously at each stage. Each player must take into account his beliefs about strategies of the other player, his beliefs about the other player’s beliefs about his strategies, and so on.
strategies and hierarchy of beliefs should, therefore, be included in the description of the state.

Since we consider an unawareness structure, only states in the ‘largest’ partial world are full descriptions of all aspects of reality relevant to the interaction among the players. States in ‘smaller’ partial worlds are subjective portraits of situations in the mind of players who are unaware of some of these dimensions. Such portraits include beliefs about the other player’s states in partial worlds that he is aware of. This allows us to formalize mutual beliefs about unawareness. If player $i$ considers possible states in a ‘smaller’ partial world, it means that he is unaware of aspects of reality not expressible in that partial world. If at these states player $i$ believes that player $j$ considers as possible only states in an ‘even smaller’ partial world, it means that $i$ believes that $j$ is unaware of some aspect of reality of which he is aware. Such a description of states is in the spirit of Heifetz et al. (2006, 2008).

As indicated above, states of partial worlds are primarily a convenient constructions which enables the modeler to discuss players’ awareness, beliefs, strategies, belief-dependent payoffs, and so on. As for the players themselves, they need not to concern themselves with the complete structure of the game. Actually, in most situations where our framework is relevant they cannot—they are unaware of some of the structure.

### 5.2 Imperfectly observable choices

We choose to focus on games with observable choices for the sake of simplicity. But our concepts and results carry over to the more general case where past choices need not be perfectly observed. Let $I_{i,T}$ be the partition of the finite set of histories $H_T$ into information sets $I_{i,T}$ of player $i$. Assume that perfect recall holds. For player $i$ the ‘perception’ function (Definition 1) now becomes a possibility correspondence

$$\varphi_i : \left( \bigcup_{T \in T} H_T \right) \rightarrow 2^{(\bigcup_{T \in T} H_T)} \setminus \emptyset,$$

with the additional property:

(vii) **Introspection**: If $I'_{i,T} \in \varphi_i(I_{i,T})$ then $\varphi_i(I'_{i,T}) = \varphi_i(I_{i,T})$.

Introspection guarantees that a player knows what he knows, and knows what he ignores provided he is aware of it. It is straight forward to add information sets to properties (i) – (vi).
The set of strategy profiles consistent with any information set $I_{i,T} \in \mathcal{I}_{i,T}$ of player $i$ must have the form $S_{H^T}(I_{i,T}) = S_i^{H^T}(I_{i,T}) \times \left[ \bigcup_{T \rightarrow T'} S_{j}^{H_{T'}}(I_{i,T'}) \right]$ (where $S_i^{H^T}(I_{i,T}) := \bigcup_{h_T \in I_{i,T}} S_i^{H^T}(h_T)$, $S_{j}^{H_{T'}}(I_{i,T'}) := \bigcup_{h_T' \in I_{i,T'}} S_{j}^{H_{T'}}(h_T')$). We denote by $I_{i,T}$ the ‘historical event’ that an information set and copies thereof that a player is aware of obtains. $\mathcal{I}_{i,T}$ is the set of such events. Consider, for the first-order beliefs of player $i$, the collection of conditioning events

$$\mathcal{C} = \left\{ F \subseteq \left( \bigcup_{T \rightarrow T'} S_{j}^{H_{T'}} \right) : F = \left( \bigcup_{T \rightarrow T'} S_{j}^{H_{T'}}(I_{i,T'}) \right) \text{ for the copies } I_{i,T'} \text{ of } I_{i,T} \in \mathcal{I}_{i,T} \right\}.$$ 

The set of infinite hierarchies of beliefs conditional on information sets $B_{i,T}$ is homeomorphic to $\Delta^{\mathcal{I}_{i,T}}(\bigcup_{T \rightarrow T'} S_{j}^{H_{T'}} \times \tilde{B}_{j,T})$ via the belief mapping $f_{i,T} = (f_{i,I_{i,T}})_{I_{i,T} \in \mathcal{I}_{i,T}}$.

Note that $\mathcal{I}_{i,T}$ specifies $i$’s information at each copy he is aware of, including those where $i$ is inactive and in particular including the terminal nodes. This would be redundant in standard games, but is crucial to model some belief-dependent motivations, such as regret or blame avoidance, whereby players’ conditional beliefs matter even if they are inactive (see e.g. Battigalli and Dufwenberg (2007) and Tadelis (2008)). We refer the reader to the discussion (Section 6) of Nielsen and Sebald (2011) for more on how to model chance moves, asymmetric information and strategic information transmission (which is different from ‘awareness’ messages) in our framework.

### 5.3 Related literature

A formal definition of forward induction has proven a little elusive. Some comments on the relationship between Battigalli and Siniscalchi (2002)’s forward induction (used in this paper) and complementary literature are warranted.

#### 5.3.1 Strong beliefs and forward induction

The non-equilibrium analysis approach to modeling forward induction used in this paper is related to work by Stalnaker. Stalnaker (1996) puts forward a normal-form, finite epistemic model, which can also be used to analyze extensive-form reasoning. Stalnaker (1998) uses this model to propose a notion of ‘absolutely robust belief’ that is equivalent to strong beliefs. He employs this notion to sketch a characterization of the following procedure: form two rounds of elimination of weakly dominated strategies, followed by iterated strict dominance. In some simple game, such as the battle-of-the-sexes with an outside option, this procedure singles out the forward induction outcome. From a technical standpoint, Stalnaker’s model specifies beliefs conditional on every event, including unobservable events concerning the
beliefs of the players. This prevents the construction of universal belief models. Stalnaker is thus forced to state his characterization result with the proviso that the incomplete model at hand is ‘sufficiently rich’ to allow for forward induction reasoning in the game under consideration. Adopting a universal belief space (as we do here) makes it possible to avoid these complications.

However, adopting a universal belief structure is in itself restrictive. As noted by Battigalli and Friedenberg (2010), rationality and common strong belief of rationality (RCSBR) may be distinct from forward induction reasoning outcomes for ‘smaller’ belief structures. In their analysis they study the behavioral implications of RCSBR across all belief structures. Formally, they show that for any given belief structure, RCSBR can be characterized by a solution concept they call extensive-form best response sets (EFBRSS). In this paper we stick to the conventional assumption that players in psychological games are capable of reasoning about the universal belief structure (see Geanakoplos et al. (1989) and Battigalli and Dufwenberg (2009)), and therefore adopt the RCSBR solution concept.

5.3.2 Other solution concepts and forward induction

The forward induction idea that, upon observing an unexpected (but rational) move of any other player, a player maintains the ‘working hypothesis’ that the other player is rational, suggests another aspect of forward induction. If there, in the solution in question, is some strategy of a player that the player would never be willing to choose, then it should not make a difference if such a strategy were removed from the game. This motivates that the iterated deletion of weakly dominated strategies supports forward induction outcomes because players would not in any case be willing to choose these strategies. In other words, a player’s belief should only assign positive probability to a restricted set of strategies of other players—the set of strategies satisfying the minimal criteria for rational play.

Two papers suggest to add the additional requirement of admissibility to the analysis of strategic-form games—that is, the avoidance of weakly dominated strategies. Brandenburger et al. (2008a) provide a characterization of admissibility in games. In their model, player’s hierarchical beliefs (i.e. epistemic types) correspond to lexicographical sequences, a generalization of lexicographic probability systems (lps’) (Blume et al., 1991a,b) that allows for an uncountable state space, over the set of other players’ strategies and beliefs. Admissibility and lps can be viewed as a strategic-form analog to strong belief [see Asheim and Søvik (2005) and Brandenburger et al. (2008b)]. Asheim and Dufwenberg (2003) propose a notion of ‘full admissible consistency’ of a player’s preferences with the game being played and with the preferences of the other players. Their main result shows that common ‘certain belief’
of full admissible consistency characterizes ‘fully permissible’ sets of strategies. The authors provide a thorough discussion of the differences and similarities between standard forward induction arguments and full permissibility, as well as between the latter and iterated weak dominance: see Asheim and Dufwenberg (2003, Section 5.1).

Others capture forward induction by refining the sequential equilibrium solution concept (Kreps and Wilson, 1982). Early studies focused on particular classes of two-player games with generic payoffs revealing aspects of what this idea entails. Outside-option games were addressed by Van Damme (1989). He proposed as a minimal requirement that forward induction should select a sequential equilibrium in which a player rejects the outside option if in the subgame there is only one equilibrium whose outcome he prefers to the outside option. Signaling games were addressed by Cho and Kreps (1987). They proposed an intuitive criteria where a candidate outcome fails if a player’s type can deviate to an off-equilibrium path and expect to obtain a higher payoff than he receives from the current outcome, provided that the other player applies forward induction whenever he observes an unexpected move. Battigalli and Siniscalchi (2002, Section 5) shows how to model the intuitive criteria using strong beliefs. Recently, Govindan and Wilson (2009) has given a general definition of forward induction as a refinement of Reny (1992)’s weakly sequential equilibrium. A weakly sequential equilibrium is the same as a sequential equilibrium except that if a player’s strategy excludes an information set from being reached, then his continuation strategy need not be optimal. They say that a strategy is ‘relevant’ for an outcome if there exists a weakly sequential equilibrium with that outcome for which the strategy is optimal at every information set it does not exclude. It is shown that an outcome satisfies forward induction if it results from a weakly sequential equilibrium in which players’ beliefs assign positive probability only to relevant strategies.

6 Conclusion

In our analysis we have provided an epistemic model of dynamic psychological games with unawareness. To this end we presented an extensive form structure that allows for unawareness concerning paths of play and extended Battigalli and Siniscalchi (1999)’s hierarchies of conditional beliefs to it. By characterizing strategies defined on subforms and player’s beliefs in our unawareness structure, we were able to define states of partial worlds. As said before, states are arrays of strategy/belief hierarchy pairs of each player at a given level of awareness, and a partial world is a collection of states defining a particular awareness level. Furthermore, in order to relax the strong requirements imposed by equilibrium thinking, we
have characterized a non-equilibrium concept which captures forward induction reasoning. More precisely, we extended Battigalli and Siniscalchi (2002)’s notion of strong beliefs to our framework with unawareness and demonstrated its practical implications for the strategic interaction of players with belief-dependent preferences using a model of sequential reciprocity and an augmented battle-of-the-sexes game with outside option. As already suggested in Nielsen and Sebald (2011), also forward induction reasoning reveals that unawareness has an important impact on the strategic interaction of players with belief-dependent preferences.
A Appendix

A.1 Proof of proposition 1

The following Lemma and proof of Proposition 1 relies heavily on Battigalli and Siniscalchi (1999)’s extension of Proposition 1 and 2 in Brandenburger and Dekel (1993) to conditional probability systems.

Let the infinite (not necessarily coherent) hierarchies of cps’ be \( \hat{B}_{i,T} = \prod_{k \geq 0} \Delta^{H_T} (\bigcup_{T \rightarrow T'} X^k_{i,T'}) \), and \( \hat{B}_{i,T} \subset \bar{B}_{i,T} \) be the subset satisfying coherency.

**Lemma 2.** Consider the following set:

\[
D = \left\{ (\delta_1^i, \delta_2^i, \ldots) : k \geq 1, \delta_i^k \in \Delta \left( \bigcup_{T \rightarrow T'} X^{k-1}_{j,T'} \right), \text{marg} \left( \bigcup_{T \rightarrow T'} X^{k-1}_{j,T'} \right) \delta_i^{k+1} = \delta_i^k \right\}.
\]

There is a homeomorphism \( h : D \rightarrow \Delta \left( \bigcup_{T \rightarrow T'} S^H_{j,T'} \times \bar{B}_{j,T} \right) \) such that

\[
\forall k \geq 1, \text{marg} \left( \bigcup_{T \rightarrow T'} X^{k-1}_{j,T'} \right) h(\delta_1^i, \delta_2^i, \ldots) = \delta_i^k.
\]

**Proof.** Let \( Z^0_{i,T} = X^0_{i,T} = S^H_{j,T}, \forall k \geq 1, Z^k_{i,T} = \Delta^{H_T} \left( \bigcup_{T \rightarrow T'} X^k_{i,T'} \right) \). Each \( Z^k \) is a Polish Space (remember; the finite disjoint union of Polish spaces is itself Polish) and

\[
D = \left\{ (\delta_1^i, \delta_2^i, \ldots) : k \geq 1, \delta_i^k \in \Delta \left( \bigcup_{T \rightarrow T'} Z^0_{i,T'} \times \cdots \times Z^{k-1}_{i,T'} \right), \text{marg} \left( \bigcup_{T \rightarrow T'} Z^0_{i,T'} \times \cdots \times Z^{k-1}_{i,T'} \right) \delta_i^{k+1} = \delta_i^k \right\}.
\]

The results then follows from Lemma 1 in Brandenburger and Dekel (1993). □

**Proof.** (Proof of Proposition 1) For each \( h_T \in H_T \), let \( \gamma_{h_T} : \hat{B}_{i,T} \rightarrow D \) be the following function:

\[
\gamma_{h_T}(\mu^1_{i,T}, \mu^2_{i,T}, \ldots) = (\mu^1_{i,T}(h_T), \mu^2_{i,T}(h_T), \ldots)
\]

\( \gamma_{h_T} \) is clearly continuous. By Lemma 2 the mapping

\[
gi_{i,h_T} = h \circ \gamma_{h_T} : \hat{B}_{i,T} \rightarrow \Delta \left( \bigcup_{T \rightarrow T'} S^H_{j,T'} \times \bar{B}_{j,T} \right)
\]

is also continuous. Let \( \mu_{i,T}(h_T) = \gamma_{h_T}(\mu^1_{i,T}, \mu^2_{i,T}, \ldots) \). Clearly, \( \mu_{i,T}(h_T|h_T) = 1 \) and for all
$k \geq 1$ common certainty of coherency at the subform is satisfied. Thus the mapping

$$g_{i,T} = (g_{i,h_T})_{h_T \in H_T} : \hat{B}_{i,T} \rightarrow \left[ \Delta \left( \bigcup_{T \rightarrow T'} S^{H_{T'}}_{j} \times \hat{B}_{j,T'} \right) \right]^{H_T}$$

is continuous and satisfies coherency. The latter implies that $g_{i,T}$ is 1-to-1 and the restriction of $g_{i,T}^{-1}$ to $g_{i,T} (\hat{B}_{i,T})$ is continuous.

The next step is to show that $g_{i,T}(\hat{B}_{i,T}) = \Delta^{H_T} \left( \bigcup_{T \rightarrow T'} S^{H_{T'}}_{j} \times \hat{B}_{j,T'} \right)$. Take $\mu_{i,T} \in \Delta^{H_T} \left( \bigcup_{T \rightarrow T'} S^{H_{T'}}_{j} \times \hat{B}_{j,T'} \right)$ and for all $h_T \in H_T$ and $k \geq 1$, define $\mu^k_{i,T}(\cdot|h_T)$ as being coherent. If $(\mu^1_{i,T}, \mu^2_{i,T}, \ldots) \in \hat{B}_{i,T}$, then $g_{i,T}(\mu^1_{i,T}, \mu^2_{i,T}, \ldots) = \mu_{i,T} \in g_{i,T}(\hat{B}_{i,T})$. It is thus sufficient to show that $(\mu^1_{i,T}, \mu^2_{i,T}, \ldots) \in \hat{B}_{i,T}$; Battigalli and Siniscalchi (1999) establishes this in the proof of their Proposition 1 by showing that each $\mu^k_i$ satisfies Definition 2 point (iii), which implies that $g_{i,T}(\hat{B}_{i,T})$ must be contained in $\Delta^{H_T} \left( \bigcup_{T \rightarrow T'} S^{H_{T'}}_{j} \times \hat{B}_{j,T'} \right)$.

Even if $i$’s hierarchy of cp’s $\mu_{i,T}$ is coherent, some elements of $g_{i,T}(\mu^1_{i,T}, \mu^2_{i,T}, \ldots)$ may assign positive probability to sets of incoherent hierarchies of player $j$. Player $i$ is certain of some collection of events $E_{j,T'} \subset S^{H_{T'}}_{j} \times \hat{B}_{j,T'}$ given $h_T \in H_T$ if $g_{i,h_T}(\mu^1_{i,T}, \mu^2_{i,T}, \ldots) (\bigcup_{T \rightarrow T'} E_{j,T'}) = 1$. Common certainty of coherency given every $h_T \in H_T$ can thus be inductively defined as follows:

$$B^1_{i,T} = \hat{B}_{i,T}$$

for all $n \geq 2$,

$$B^n_{i,T} = \left\{ \left( \mu^1_{i,T}, \mu^2_{i,T}, \ldots \right) \in B^{n-1}_{i,T} : \forall h_T \in H_T, g_{i,h_T}(\mu^1_{i,T}, \mu^2_{i,T}, \ldots) \left( \bigcup_{T \rightarrow T'} S^{H_{T'}}_{j} \times B^{n-1}_{j,T'} \right) = 1 \right\},$$

$$B_{i,T} = \bigcap_{n \geq 1} B^n_{i,T}.$$  

$B_{i,T} \times B_{j,T}$ is the set of pairs of hierarchies satisfying common certainty of coherency conditional in the subform $T$.

The final step in the proof is to show that the restriction of $g_{i,T} = (g_{i,h_T})_{h_T \in H_T}$ to $B_{i,T} \subset \hat{B}_{i,T}$ induces a homeomorphism $f_{i,T} = (f_{i,h_T})_{h_T \in H_T} : B_{i,T} \rightarrow \Delta^{H_T} (\bigcup_{T \rightarrow T'} \Omega_{j,T'})$, where $\Omega_{j,T'} = S^{H_{T'}}_{j} \times \hat{B}_{j,T'}$. It is easy to check that $B_{i,T} = \left\{ \left( \mu^1_{i,T}, \mu^2_{i,T}, \ldots \right) \in \hat{B}_{i,T} : g_{i,T}(\mu^1_{i,T}, \mu^2_{i,T}, \ldots) (\bigcup_{T \rightarrow T'} \Omega_{j,T'}) = 1 \right\}$, so

$$g_{i,T}(B_{i,T}) = \left\{ \left( \mu^1_{i,T}, \mu^2_{i,T}, \ldots \right) \in \Delta^{H_T} \left( \bigcup_{T \rightarrow T'} S^{H_{T'}}_{j} \times \hat{B}_{j,T'} \right) : \left( \mu^1_{i,T}, \mu^2_{i,T}, \ldots \right) \left( \bigcup_{T \rightarrow T'} \Omega_{j,T'} \right) = 1 \right\},$$

$g_{i,T}(B_{i,T})$ is homeomorphic to $B_{i,T}$, and $\left( \mu^1_{i,T}, \mu^2_{i,T}, \ldots \right) \in \hat{B}_{i,T}$:
\[ g_{i,T}(\mu_{i,T}^1, \mu_{i,T}^2, \ldots) (\bigcup_{T \rightarrow T'} \Omega_{j,T'}) = 1 \} \text{ is homeomorphic to } \Delta^{H_T} (\bigcup_{T \rightarrow T'} \Omega_{j,T'}). \text{ So } B_{i,T} \text{ is homeomorphic to } \Delta^{H_T} (\bigcup_{T \rightarrow T'} \Omega_{j,T'}). \]
References


