Dynamic Games with Complete Awareness Information [1]

Jing Li

Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104
E-mail: jing.li@econ.upenn.edu

Incomplete, please do not circulate.

March, 2006

---

[1]

Abstract

I study games where players may be unaware of some fundamentals of the game itself, such as actions or players, and such unawareness may be lifted by other players' actions during the play. Based on Li (2006a, 2006b), I propose a framework to deal with such games using standard tools of game theory, including generalization of familiar solution concepts such as Nash equilibrium and sequential equilibrium. I construct examples to show that unawareness is distinctly different from assigning zero probability in games.

Keywords: unawareness, awareness information, generalized information structures, Nash equilibrium, sequential rationality

> "(W)hat happens is always the thing against which one has not made provision in advance."
>
> John Maynard Keynes, in a letter to Jacob Viner

1 Introduction

Consider a multi-issue bargaining environment. There is a set of payoff-relevant issues that should be on the bargaining table, however, the players may not be fully aware of them, and the issues of which one player is aware may differ from what the other is aware. Players alternately propose bargaining agendas as well as bargain over known issues. Upon seeing an issue in the agenda of which the player wasn't aware before, he becomes aware of it and hence has a different view of what the real game is. He may even be reminded of other issues of which he wasn't aware.

Such a game, with its obvious realistic appeal, is beyond the scope of the traditional incomplete information games. The private information bargainers have in this environment is of a particular kind, namely information about the strategy space of the game. That the game itself and players' information structures are common knowledge is assumption zero in game theoretic analysis, yet it is exactly this assumption that is violated in this environment.

In this paper, I formalize a theory of dynamic games with unawareness, where the players may have different perceptions of the real game, and may update their perceptions upon observing unexpected moves of the opponents. Research on games with unawareness is largely unexplored. In two important pioneering papers, using an epistemic framework, Feinberg studies finitely repeated prisoner's dilemma game where one player may be unaware of the defect action until the opponent defects (Feinberg 2004); and proposes a framework for normal-form games with unawareness (Feinberg 2005). I take a different approach in this paper: both the game and the solution concepts are generalizations of those of the standard theory, and standard techniques apply.

2 Formulating Games with Unawareness

A *game* specifies who moves when knowing what as well as the consequences of the actions. A *game with unawareness* is one where players move and incur consequences as specified in the real game, but may be unaware of some actions or players in it.

Let the real game be denoted by Γ. It is described by the tuple (N, H, L, f, u, P) where the components are interpreted as follows:[1]

N: a finite set of players;

---

[1] The formalism follows Osborne and Rubinstein (1994).

H: a set of sequences satisfying

1. Initial history: $\emptyset \in H$;
2. Tree: $(a_1, \cdots, a_k) \in H \Rightarrow (a_1, \cdots, a_l) \in H$ for all $l < k$;
3. Infinite history: $(a_1, \cdots, a_k) \in H$ for all $k \in \mathbb{N} \Rightarrow (a_1, \cdots) \in H$.

H is interpreted as the set of histories. Let $Z \subset H$ denote the set of terminal histories and $D = H \setminus Z$ denote the set of non-terminal histories. I slightly abuse notation and refer to a non-terminal history and the decision node following it interchangeably. Let $A(h) := \{a : (h, a) \in H\}$ be the set of available actions after a nonterminal history h. I require $\#[A(h)] > 1$, where the symbol $\#$ denotes the cardinality of the set: this is a nontriviality condition that says there are at least two actions available at each decision node;

$L : D \to N \cup \{c\}$: the player function, i.e. $L(h)$ is the player who moves after history h. The player could be chance, denoted by c;

$f = \{f(\cdot|h)\}_{h \in \{h : L(h) = c\}}$: a collection of probability measures, one for each chance node, with the interpretation that $f(\cdot|h)$ specifies a probability measure over $A(h)$ where $L(h) = c$;

$u = \{u_i\}_{i \in N}$: a collection of von-Neumann Mongenstern expected utilities, one for each player;

P: a partition over D such that $h \in P(h')$ implies $A(h) = A(h')$ and $L(h) = L(h')$. The interpretation is, at h, $L(h)$ considers the set of histories $P(h)$ possible. Thus P represents players' *factual information structures* in the real game.

Based on the definition of Γ, I formalize the concept of a game with unawareness.

Definition 1 *I say* Γ^u *is a* **game with unawareness** *with the real game* Γ*, if*

1. *the game is played by players in* N*, uncertainties are resolved as specified in* f *and* P*, the possible play paths are described by* H *and* L*, and players receive payoffs as specified by* u*;*
2. *every player has a subjective game that is an incomplete version of* Γ*, in the sense that it is played by a subset of* N*, uncertainties are resolved in a way consistent with* f *and* P*, the possible play paths are described by a subset of* H *and the restriction of* L *on it, payoffs are as specified by* u*. Moreover, every player knows everybody else (of which the former is aware) may have a subjective game that is an incomplete version of* Γ*, and every player knows every player knows everybody else* $\cdots$*, and so on.*

To represent Γ^u, I endow each player at each decision node in Γ an awareness signal which determines his perception of the game, in addition to the factual signals captured by P. Since the set of histories contains the key structure of a game, it is natural to let the awareness signal take the form of a subset of H. Thus, I consider an awareness function $W : D \to 2^H \setminus \{\emptyset\}$. The interpretation is, $W(h) \subseteq H$ contains the set of histories in $L(h)$'s subjective game.[2] The pair (W, P) consists of the generalized information structure discussed in Li (2006a, 2006b), where it is shown this information structure captures the essence of unawareness.

Let $\trianglerighteq$ denote the partial order "precedence" over H, i.e. $h \trianglerighteq h'$ if $h = h'$ or h is a history preceding h'. Let $\triangleright$ denote the asymmetric part of the relation. Let (h, h') denote the concatenation of two sequences h and h'. Consider the following assumptions:

$A0$ (chance): $L(h) = c \Rightarrow W(h) = H$;

$A1$ (tree): $W(h)$ satisfies initial history, tree, and infinite history for all $h \in D$;

$A2$ (default): for all $h \in D$, there exists $\mathbf{h}^d(h) \in P(h), \mathbf{a}^d(h) \in A(h)$ such that,

1. $(\mathbf{h}^d(h), \mathbf{a}^d(h)) \in W(h)$;
2. $h \in W(h') \Rightarrow (\mathbf{h}^d(h), \mathbf{a}^d(h)) \in W(h')$.

$A3$ (partitional generalized information): $h' \in P(h) \Rightarrow (W(h'), P(h')) = (W(h), P(h))$ for all $h, h' \in D$;

$A4$ (complete awareness information): $L(h) = c, h_1 = (h, a, h'), h_2 = (h, b, h'), L(h_1) = L(h_2), W(h_1) \neq W(h_2) \Rightarrow \nexists\, h_1' \trianglelefteq (h, a), h_2' \trianglelefteq (h, b)$ and $h_1' \in P(h_2')$;

$A5$ (monotonicity): $h \trianglerighteq h', L(h) = L(h') \Rightarrow W(h) \subseteq W(h')$.

$A0 - 3$ are regularity conditions. A0 is obvious: it says chance is fully aware. $A1$ requires $W(h)$ consist of a game tree, and hence can be interpreted as the set of histories of an incomplete version of Γ. Slightly abusing notation, let $Z(h) \subseteq W(h)$ denote the set of terminal histories in $W(h)$, and similarly $D(h) = W(h) \setminus Z(h)$. $A2$ is a nontriviality condition. It requires for each decision node, there exist a "default" history leading to it and a "default" action at that decision node, of which the active player is always aware (the first part), and so is every player whose subjective game include a history reaching this information set (the second part). This condition ensures players are not "too unaware" to play the game or to be strategic: the first part ensures players can always move the game forward; the second part ensures players share common knowledge of some minimal game, namely the "default" sequence of moves. In particular, the second

---

[2] For clarity in exposition, I use $L(h)$ to refer to player $L(h)$ at his decision node h. For example, the phrase "$L(h)$'s subjective game" where $L(h) = i$ is to be understood as "i's subjective game at h."

part makes sure $Z(h) \subseteq Z$, which is necessary for the payoffs in the subjective games to be well-defined.[3]

I say $L(h)$ is unaware of a decision node h' if either $h' \notin W(h)$ or $\#[A(h'|h)] = 1$, where $A(h'|h) := \{a : (h', a) \in W(h)\}$ is the set of feasible actions for $L(h')$ in $L(h)$'s subjective game. Notice that by $A2$, $\#[A(h'|h)] = 1$ implies $A(h'|h) = \left\{\mathbf{a}^d(h')\right\}$, i.e. the default action. If $L(h)$ is unaware of all of j's decision nodes, then $L(h)$ is unaware that j is in the game. Thus the set of players in $L(h)$'s subjective game is:

$$N(h) = \{L(h') : h' \in D(h), \#[A(h'|h)] > 1\}$$

$A3$ extends the partitional structure on factual signals in the standard theory to the signal pairs consisting of both factual signals and awareness signals in this environment. It requires players have identical awareness signals at decision nodes that belong to the same factual information set.[4]

Despite their seeming restrictions on (un)awareness structures, $A2$ and $A3$ are better viewed as regularity conditions on how the real game Γ should be represented. Intuitively, subjective games are represented as restrictions of (N, L, f, u, P) on a subset of H, and hence players' subjective perceptions or evaluations appeared in the subjective games must be accommodated in these components in the real game. Example 1 illustrates this point in the case of $A3$.

Example 1. Imagine that Alice and Bob are two big stock holders in a company. There is a good market so that one of them could sell out the stocks at a profitable price. The one who keeps the stocks also benefits from gaining more control of the company. If both sell, the price is driven down and both lose. However, Bob only becomes aware of the strong market upon observing large volumes of trade on the stocks, without realizing this is because of Alice's selling activity.

It seems natural to depict the real game as in figure 1(a), of which Bob is unaware if Alice holds on to the stocks. But then this game violates $A3$ because Bob has different awareness in his decision nodes s and k, which are in the same factual information set. The problem is 1(a) does not correctly specify the game. Intuitively, Alice's actions are really two-dimensional: sell and inform Bob of the market opportunity ($s\&i$), or keep

---

[3] Let $\succcurlyeq_i$ denote i's preference relation over $\Delta(Z)$, the set of lotteries over Z. Suppose $\succcurlyeq_i$ has an expected utility representation. Then it is easy to check that $\succcurlyeq_i$ restricted on $\Delta(Z')$ for any $Z' \subseteq Z$ is also a preference relation and has an expected utility representation. Thus it is without loss of generality for the subjective game to inherit u_i.

[4] The role of this assumption is analogous to the assumption that P is an information partition in the standard theory, which is extensively discussed in Aumann (1987) and Brandenburger and Dekel(1993), among others. Roughly speaking, the generalized information structure (W, P) is merely a coding system, so that for every player, *within the confines of his awareness*, he "knows" it, and "knows" that *within the confines of everyone's own awareness*, everyone "know" it, and so on. Since information structures themselves are not events in an epistemic model, I use the quotation marks to indicate the informal usage of knowledge. For detailed discussion of the partitional generalized information, see Li (2006a).

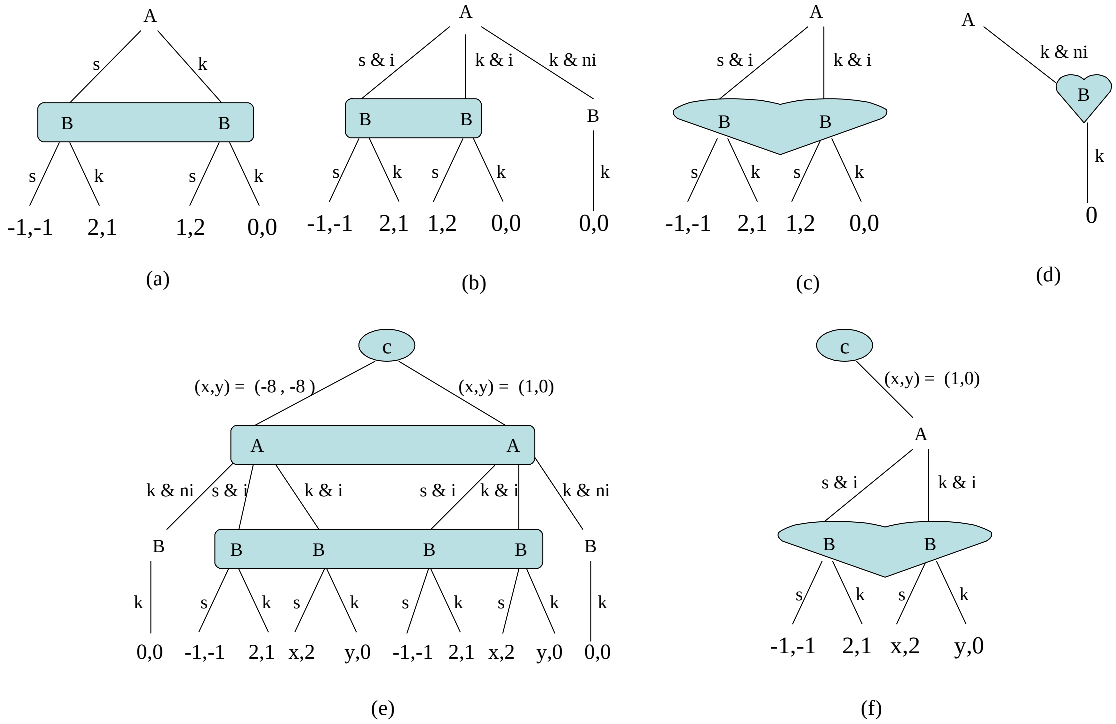

Figure 1: a game with unawareness of the possibility of being unaware

the stocks and hence does not inform Bob of the market opportunity ($k\&ni$). However, Bob is never aware of this awareness information content dimension. In his subjective game at $s\&i$, Alice's actions are simply s and k, the projection of $s\&i, k\&ni$ on the dimension of which he is aware, with the default value of the awareness information content dimension (of which he is unaware) set at i. Thus, to adequately capture the awareness information structures, the action $k\&i$ should be added to the real game. Figure 1(b) depicts the enriched version of the real game. Now it is straightforward to represent players' subjective games: Alice is aware of the real game, while Bob's subjective games are depicted in figures 1(c) and 1(d), where the shaded hearts indicate the corresponding relevant decision nodes.

One may argue such an action $k\&i$ may not be available in the real game. This can always be represented by adding a chance move that specifies a cost of infinity for the undertaker of this move, of which the opponents may be unaware. Figure 1(e) depicts the game where $k\&i$ is not available in the real game. Chance selects $(x, y) = (-\infty, -\infty)$ with probability 1, so $k\&i$ is infeasible for Alice. However, Bob is unaware of the chance move. His subjective game is as depicted in 1(f).

An obviously interesting question is what if Alice is *uncertain about* whether Bob is aware that his awareness of the game depends on her own actions. A natural approach to model such situations is to extend the concept of a "type" to be a hierarchy of beliefs, where different orders of beliefs could reside in different probability spaces.[5] It is not clear whether such an expanded universal type space exists. For simplicity, in this paper, I restrict attention on games with complete awareness information, i.e. games where players do not have exogenous uncertainties about opponents' awareness. This is the content of $A4$: if a player has different awareness at two decision nodes that only differ in chance moves, then no player can have uncertainties about these chance moves.

With this simplification, one can accommodate the standard type of incomplete (factual) information in games with unawareness by including in the primitive specification a "subjective prior" distribution for every possible subjective game. For any $h \in H$ such that $L(h) = c$, let $\mathcal{A}(h)$ be a σ-algebra on $A(h)$ and $\mathcal{A}(h)|_A$ denote the induced sub-σ-algebra on a measurable set $A \subseteq A(h)$.

$f^u = \{f(\cdot|h, A)\}_{A \in \mathcal{A}(h), h \in \{h : L(h) = c\}}$: a collection of probability distributions, one for each measurable subspace of chance moves $A \subseteq A(h)$ where $L(h) = c$, such that $f(\cdot|h, A)$ is a probability measure defined on $\mathcal{A}(h)|_A$, and for all $B \in \mathcal{A}(h)|_A$, $f(B|h, A) \geq f(B|h, A(h))$.

Remark 1 *Roughly, in this model, players can be said to have a common prior f had they been fully aware. Unawareness limits the set of events one can reason about, but does not necessarily create other distortions, such as wrong beliefs. One way to understand it is that players reason in a consistent way. Once the unawareness is lifted, they recognize all relevant factors, and the Harsanyi doctrine would apply. An (un)awareness type specifies both what external uncertainties of which the player is aware and how such unawareness affects his probability assessments, given the "full" prior f. In this sense this model imposes common "prior" in every subjective game. However, it is worth pointing out that players who are aware of the same set of external uncertainties may not assign the same probability distribution over it.*

For any $h' \in D(h)$, recall that $A(h'|h)$ is the set of actions for $L(h')$ of which $L(h)$ is aware. Let $f_h = \{f(\cdot|h', A(h'|h))\}_{h' \in \{h' \in W(h) : L(h') = c\}} \subseteq f^u$ denote the collection of probability measures defined over $A(h'|h)$, one for each chance node h' of which $L(h)$ is aware.

Let $L(\cdot|h) = L|_{D(h)}$ denote the restriction of the player function on the set of decision nodes in $W(h)$. Let $u_h = \{u_i\}_{i \in N(h)}$ denote the set of utility functions for players of which $L(h)$ is aware. The factual signal structures in the subjective game is represented by the restriction of the full information partition P on the set of histories of which $L(h)$ is aware: $P(\cdot|h) = P|_{D(h)}$. Thus, $L(h)$'s perception of the real game, denoted

---

[5] Li (2006a) shows unawareness can be characterized by restrictions on the set of measurable events, or equivalently, probability spaces.

by $\tilde{\Gamma}(h)$, is represented by the following tuple that defines a standard extensive-form game:

$$(N(h), W(h), L(\cdot|h), f_h, u_h, P(\cdot|h)) \tag{2.1}$$

An intriguing aspect in games with unawareness is that players reason about each other's subjective game, and may even take actions to affect the opponent's perception of the game. Intuitively, for $L(h)$, $\tilde{\Gamma}(h)$ plays the role of the real game in his reasoning about other players, in the same way as the real game Γ does in the modeler's analysis of the game with unawareness Γ^u. For ease in exposition, I refer to $\tilde{\Gamma}(h)$ as $L(h)$'s subjective game, to be understood as the $L(h)$'s subjective version of the real game Γ, and use $\tilde{\Gamma}^u(h)$ to denote $L(h)$'s *subjective game with unawareness*, which is $L(h)$'s subjective version of the game with unawareness Γ^u.

Unlike knowledge, there is a natural asymmetry in dealing with higher-order interactive awareness. Given an uncertain natural event E, i can reason about j's awareness of E only if he is aware of both E and j. But then if i can reason about both E and j, it seems natural for him to reason about j's reasoning about E by default. In other words, awareness should be closed under conjunction. Consequently, while it is equally plausible for i to be aware or unaware of a natural event E, it is only plausible for i to be unaware that j might be unaware of E, rather than the other way around. This model reflects this feature: it is never the case that both i and j are aware of an event E, and i is aware of j, yet i is unaware that j is aware of E.

On the other hand, the combination of j's unawareness and i's unawareness of j's unawareness gives rise to extra complications in the game environment. The following example illustrates the problem.

Example 2. Recall example 1, but suppose Alice is unaware that after she plays $s\&i$, Bob could remain unaware that his awareness of the game depends on Alice's action. Bob, being unaware of Alice's action $k\&ni$, is unaware of the awareness type of Alice, too. To cast this game in extensive form, I introduce chance moves to select different types of Bob and Alice. This is depicted in figure 2(a). Chance move $c1$ selects the Alice and Bob as described in the game, while chance move $c2$ selects an Alice Bob has in mind, and a Bob Alice has in mind. Chance chooses $c1$ with probability 1. I refer to the identical subgame following the nature moves $c1$ and $c2$ as the naive form of the real game. Alice at $c2$ as the same perception of the naive form of the real game as Bob at $(c1, s\&i)$, while Bob at $(c2, s\&i)$ has the same perception of the naive form of the real game as Alice at $c1$.

This allows me to model Alice's unawareness of Bob's unawareness as unawareness of chance moves. In real Alice's subjective game (2(b)), Bob's decision node $(c1, s\&i)$ should be viewed as the projection of the set $\{(c1, s\&i), (c2, s\&i)\}$. Thus, Alice's perception of Bob's subjective game at $(c1, s\&i)$ then is naturally everything of which Bob is aware at *either* of these decision nodes and of which Alice is aware herself. Similarly, in real Bob's subjective game (2(c)), Alice's decision node $c1$ should be viewed as the

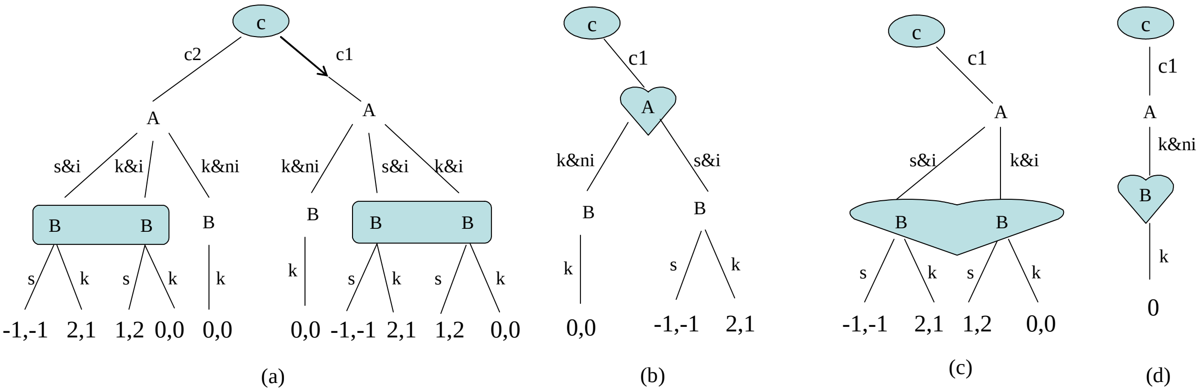

Figure 2: two-player game with higher-order interactive unawareness

projection of $c1$ and $c2$.

Formally, for any $h \in D$, $a, b \in A(h)$, $a \neq b$, I say a and b are *mirror actions* if they are moves purely resolving awareness types, leading to otherwise identical continuation games: $\{h_0 : (h, a, h_0) \in H\} = \{h_0 : (h, b, h_0) \in H\}$, and for any such $h_0, L(h, a, h_0) = L(h, b, h_0)$, $(h, a, h_0) \in Z \Rightarrow u_i(h, a, h_0) = u_i(h, b, h_0)$ for all $i \in N$, $P((h, a, h_0)) = (h, a) \times \{h_\alpha : (h, b, h_\alpha) \in P(h, b, h_0)\}$. Write it as $a \sim b$. I say a history h_1 is a *mirror history* of h_2, denoted by $h_1 \approx h_2$, if they only differ in a pair of mirror actions: $h_1 \approx h_2$ if $h_1 = (h, a, h_0), h_2 = (h, b, h_0)$ and $a \sim b$. In example 2, $c1$ and $c2$ are mirror actions/histories, and $(c1, s\&i) \approx (c2, s\&i)$.

Thus, $L(h)$ is unaware that $L(h')$ is unaware of E while $L(h')$ is unaware of E if and only if $L(h)$ is unaware of the corresponding mirror actions. Let $h \in D$ and $h' \in D(h)$,

$$T(h'|h) = \{h'' \in D \setminus D(h) : h'' \approx h'\} \tag{2.2}$$

This set collects the mirror decision nodes for $L(h')$ in $L(h)$'s subjective game. In example 2, $T((c1, s\&i)|u) = \{(c2, s\&i)\}$. Let $W(\cdot|h) : D(h) \to 2^{W(h)\setminus\{\emptyset\}}$ denote $L(h)$'s *first-order interactive awareness function.* This function describes $L(h)$'s perception of the awareness signal structure in his subjective game $\tilde{\Gamma}(h)$, *subject to the constraint of his own awareness*, $W(h)$. Taking into account mirror actions, $W(\cdot|h)$ is defined as follows:

$$W(h'|h) = W(h) \cap [W(h') \bigcup_{h'' \in T(h'|h)} W(h'')] \tag{2.3}$$

Players can learn new awareness during the play of the game: W may not be constant over $\{h : L(h) = i\}$. A natural restriction on this learning aspect is that players do not forget their own awareness. This is the content of A6: a player's awareness is nondecreasing along a play path. Given the nature of awareness, such monotonicity seems compelling.[6] $A5$ ensures that in his own subjective game, $L(h)$ has the same awareness

---

[6] It may be worth summarizing here that while $A0-3$ are regularity conditions, $A4-5$ do impose

in all future decision nodes $W(h'|h) = W(h)$ if $L(h') = L(h), h \trianglerighteq h'$. In particular, this says players never *anticipate* to perceive a different game at future decision nodes.[7]

Higher-order interactive awareness are recursively defined as follows. Suppose all interactive awareness functions are defined up to n-th order. Let $W(\cdot|(h_1, \cdots, h_n))$ denote the n-th interactive awareness function. For notational ease, let $q^m = (h_1, \cdots, h_m)$ for all $m = 1, \cdots, n$. This function maps histories in $D(h_n|q^{n-1})$, the set of non-terminal histories in $W(h_n|q^{n-1})$, to a non-empty subset of $W(h_n|q^{n-1})$. Let $h_{n+1} \in D(h_n|q^{n-1})$.

$$T(h_{n+1}|q^n) = \left\{h \in D(h_{n-1}|q^{n-2}) \setminus D(h_n|q^{n-1}) : h \approx h_{n+1}\right\} \tag{2.4}$$

$$W(h_{n+1}|q^n) = W(h_n|q^{n-1}) \cap [W(h_{n+1}|q^{n-1}) \bigcup_{h \in T(h_{n+1}|q^n)} W(h|q^{n-1})] \tag{2.5}$$

Roughly speaking, in $n+1$-order reasoning, $L(h_1), \cdots, L(h_{n-1})$ are fully aware in the higher-order subjective game $\tilde{\Gamma}(h_{n-1}|q^{n-2})$, which plays the role of the real game Γ in Γ^u. The set $T(h_{n+1}|q^n)$ contains the relevant mirror decision nodes of h_{n+1} in the "real game" $\tilde{\Gamma}(h_{n-1}|q^{n-2})$, of which $L(h_1), \cdots, L_{h-1}$ is aware that $L(h_n)$ is unaware, and hence takes into account when reasoning about h_n's reasoning about h_{n+1}'s awareness.

That every player can only reason within his own subjective game restricts the domain of higher-order interactive awareness.

Definition 2 *A sequence $q^n = (h_1, \cdots, h_n)$ is* **permissible** *if $h_1 \in D$, and for all $m = 2, \cdots, n$, $h_m \in W(h_{m-1}|q^{m-2})$.*

The main result of this section is that under the construction in 2.2–2.5, $A0-5$ are the conditions one needs to specify a game with unawareness. The result follows from the observation that the tuple 2.1 is a well-defined extensive-form game under $A0-5$, while the interactive awareness function preserves the higher-order counterparts of these assumptions.

Lemma 1 *Let H satisfy initial history, tree and infinite history, and $W : D \to 2^H \setminus \{\emptyset\}$ satisfy $A0-5$. Then for any $h \in D$, the function $W(\cdot|h)$ as defined in equation 2.3 satisfies the adaptation of $A0-5$ to $W(h)$.*

Under the recursive construction of higher-order interactive awareness, the following results are immediate. Let the components in higher-order subjective games be denoted in the obvious fashion.[8]

---

real restrictions on (un)awareness structures. In particular, while $A4$ is a simplification assumption that should, and would eventually be dispensed of, perhaps in future work, $A5$ seems to be an assumption that ought to persist in this environment.

[7] That the real game is a standard game rules out the possibility that players actively acquire new awareness information and anticipate to see a different game of which they are unaware. On the other hand, the current framework is capable of modeling players' subjective perceptions of a continuation game. In this sense, ruling out active acquisition of awareness information is without loss of generality.

[8] I give the precise definition in the appendix.

Theorem 2 *Let* $\Gamma^u = (N, H, L, f^u, u, W, P)$ *where* $\Gamma = (N, H, L, f, u, P)$ *is a standard extensive-form game and* $W : D \to 2^H \setminus \{\emptyset\}$ *satisfies* $A0 - 5$*. Then for any* n*, any permissible* q^n*, the tuple*

$$\tilde{\Gamma}(q^n) = (N(h_n|q^{n-1}), W(h_n|q^{n-1}), L(\cdot|q^n), f_{q^n}, u_{q^n}, P(\cdot|q^n)) \tag{2.6}$$

where $W(h_n|q^{n-1})$ *is as defined in equations 2.2–2.5, defines an extensive-form game which is an incomplete version of* $\tilde{\Gamma}(q^{n-1})$*. Moreover, if* Γ *has perfect recall, so does* $\tilde{\Gamma}(q^n)$*.*

Notice that $A5$ is necessary for $\tilde{\Gamma}(q^n)$ to inherit perfect recall from Γ. By definition 1, Γ^u satisfying $A0-5$ defines a game with unawareness. In fact, each player's subjective perception of Γ^u is a game with unawareness itself. For any $h' \in W(h)$, let $\mathcal{A}(h'|h) \equiv \mathcal{A}(h)|_{A(h'|h)}$ denote the induced sub-σ-algebra of $\mathcal{A}(h)$ on $A(h'|h)$, the set of actions at h' of which $L(h)$ is aware. The subjective game with unawareness is equipped with the collection of "subjective priors" for every possible incomplete perception of $\tilde{\Gamma}(h)$. Let this collection be denoted by f^u_h, and defined by:

$$f^u_h = \{f(\cdot|h', A)\}_{A \in \mathcal{A}(h'|h), h' \in \{h' \in W(h) : L(h') = c\}}$$

The higher-order counterpart of f^u is obtained by replacing $W(h)$ and $A(h'|h)$ with their appropriate higher-order counter-parts.

Theorem 3 *Let* $\Gamma^u = (N, H, L, f^u, u, W, P)$ *and* W *satisfy* $A0-5$*. Then for any* n*, any permissible* q^n*, the tuple*

$$\tilde{\Gamma}^u(q^n) = (N(h_n|q^{n-1}), W(h_n|q^{n-1}), L(\cdot|q^n), f^u_{q^n}, u_{q^n}, W(\cdot|q^n), P(\cdot|q^n)) \tag{2.7}$$

defines a game with unawareness with the real game $\tilde{\Gamma}(q^n)$*.*

The hierarchy of subjective games in this model satisfy intuitive properties in this environments. For example, the game in example 2 is completely characterized by the following hierarchy of subjective games:

1. $\tilde{\Gamma}(q^n) = \tilde{\Gamma}(c1) = 2(b)$ for all q^n such that $h_1 = c1, h_j \in \{c1, (c1, s\&i), (c1, k\&i)\}$ for all $j = 2, \cdots, n$;

2. $\tilde{\Gamma}(q^n) = \tilde{\Gamma}((c1, k\&ni)) = 2(e)$ for all q^n such that $h_1 = c1, h_j \in \{c1, (c1, s\&i), (c1, k\&i)\}$ for all $j = 2, \cdots, m, m < n$ and $h_j = (c1, k\&ni)$ for $j = m+1$, $h_j \in \{c1, (c1, k\&ni)\}$ for $j = m+2, \cdots, n$;

3. $\tilde{\Gamma}(q^n) = \tilde{\Gamma}((c1, s\&i)) = 2(c)$ for all q^n such that $h_1 \in \{(c1, s\&i), (c1, k\&i)\}$, $h_j \in \{c1, (c1, s\&i), (c1, k\&i)\}$ for all $j = 2, \cdots, n$;

4. $\tilde{\Gamma}(q^n) = \tilde{\Gamma}((c1, k\&ni)) = 2(e)$ for all q^n such that $h_1 = (c1, k\&ni)$, $h_j \in \{c1, (c1, k\&ni)\}$ for all $j = 2, \cdots, n$.

In this example, although there is no common knowledge of any game between Alice and Bob, there is a sense in saying Alice "knows" the game 2(b) is "common knowledge" between her and Bob at $(c1, s\&i)$: hierarchy 1 says Alice "knows" the game 2(b), and she "knows" Bob "knows" it, and she "knows" Bob "knows" she "knows" it, and so on (hierarchy 1).[9] Similarly, Bob at $(c1, s\&i)$ "knows" the game 2(c) is "common knowledge" between himself and Alice (hierarchy 2). Intuitively, players cannot reason about things of which they are unaware themselves, and this is "common knowledge" among all players. Consequently, every player "knows" the opponents' interactive awareness, including what the opponents consider to be "common knowledge." The following results formalize the above observations.

Theorem 4 *Let* $\Gamma^u = (N, H, L, f^u, u, W, P)$ *be a game with unawareness. For any* n *and any permissible* q^{n+1}*, if* $h_i \trianglerighteq h_j, L(h_i) = L(h_j)$ *for some* $1 \leq i < j \leq n$*, then*

$$W(h_{n+1}|q^n) = W(h_{n+1}|(h_1, \cdots, h_{j-1}, h_{j+1}, \cdots, h_n)) \tag{2.8}$$

Equation (2.8) says players always "knows" they are fully aware (in their own subjective games), and there is "common knowledge" of this.

Corollary 5 *Let* $\Gamma^u = (N, H, L, f^u, u, W, P)$ *be a game with unawareness. For any* n *and any permissible* q^{n+1}*,*

$$W(h_{n+1}|q^n) = W(h_{n+1}|(h_1, r_1^{m_1}, h_2, r_2^{m_2}, \cdots, h_n, r_n^{m_n})) \tag{2.9}$$

for any $m_1, \cdots, m_n \in N$*, where* $r_i^{m_i} = (r_1^i, \cdots, r_{m_i}^i)$ *is a permissible sequence such that* $r_j^i \in \{h_1, \cdots, h_i\}, r_1^i \in W(h_{i-1}|q^{i-2})$ *for all* $j = 1, \cdots, m_i$*.*

Proof. Repeatedly apply theorem 4 to eliminate all the $r_i^{m_i}$'s from the right hand side of equation (2.9). ☐

Equation (2.9) roughly says, every player "knows" there is "common knowledge" of higher-order subjective games among relevant players, *from each player's subjective perspective.* In the above example, Alice and Bob have different subjective games, and

---

[9] Here I use "know" and "common knowledge" in an informal sense and hence the quotation marks. Strictly speaking, the hierarchy only says Alice is *aware* that Bob is *aware* of 2(b), and that Bob is *aware* that Alice is *aware* of 2(b), and so on. For the purpose of this section, it is enough to note that in an epistemic model based on the generalized information structure (W, P), the state spaces are constructed from sets of histories. Since it is tautological to say agents know their subjective state spaces, awareness of a game is essentially equivalent to knowledge of the game. For a rigorous treatment, see Li (2006a) and Section **??**.

both of them are different from the real game, yet each thinks her/his own subjective game is "common knowledge" between them. In games with more than two players, there is a sequence of nested games *from each player's perspective* that he thinks to be "common knowledge" between different subsets of players. For example, suppose $L(h_1), L(h_2), L(h_3)$ are aware of each other, then both subjective games defined by $W(h_3|(h_1, h_2))$ and $W(h_2|(h_1, h_3))$ are "common knowledge" among them from different subjective perspectives, yet it could be the case that $W(h_3|(h_1, h_2)) \neq W(h_2|(h_1, h_3))$.

3 Solution Concepts

Throughout this section I focus attention on cases where the underlying real game Γ is a finite dynamic game with perfect recall.

The lack of common knowledge, or even mutual knowledge of the real game raises intriguing conceptual issues when it comes to solutions of games with unawareness. Osborne and Rubinstein observes that:

> There are two conflicting interpretations of solutions for ... games. ... The *steady state* interpretation treats a game as a model designed to explain some regularity observed in a family of similar situations. Each participant "knows" the equilibrium and tests the optimality of his behavior given this knowledge, which he has acquired from his long experience. The *deductive* interpretation ... treats a game in isolation, as a "one-shot" event, and attempts to infer the restrictions that rationality imposes on the outcome; it assumes that each player deduces how the other players will behave simply from principles of rationality.[10]

A solution of a game ultimately concerns what the players do, which in turn depends on what they know or believe. The common ground in the above two interpretations is a basic rationality assumption linking behavior and information, i.e. expected utility maximization. On the other hand, the two interpretations differ critically on what the players know in a solution. In a deductive solution, players' information is entirely derived from the basic behavioral assumption of rationality. In contrast, the steady state interpretation typically invokes some unmodeled information dissimilation process that instills knowledge or beliefs about equilibrium actions into the players.

The concept of rationality is fundamental to both interpretations. With unawareness, while it seems natural to define rationality to be optimization within one's awareness constraints, higher-order beliefs of rationality become problematic given that players may be aware of different games. On the other hand, in the steady state interpretation, although one could to some extent avoid modeling players' interactive reasoning processes explicitly, some interpretations are clearly less plausible than others.

---

[10] Osborne and Rubinstein (1994), page 5.

3.1 Strategies.

Intuitively, players can only reason about and choose an action of which he is aware. To analyze the hierarchy of subjective games in a game with unawareness, one needs appropriate notions of strategies to account for each of them.

By $A2$, $L(h)$ is always aware of some history $\mathbf{h}^d(h)$ leading to $P(h)$. Let $\mathring{A}(h) = A(\mathbf{h}^d(h)|h)$ denote the set of actions of which $L(h)$ is aware at h.[11] The set of *objective pure strategy profiles* in Γ^u is $S(\Gamma^u) = \times_{h \in D} \mathring{A}(h)$. The set of objective pure strategies for player $i \in N$ is denoted by $S_i(\Gamma^u) \equiv \times_{h \in D, L(h)=i} \mathring{A}(h)$. The set of mixed strategy profile is $\Sigma(\Gamma^u) = \Delta[\times_{h \in D} \mathring{A}(h)]$ where $\Delta(A)$ denote the set of probability distributions over the set A. Similarly let $\Sigma_i(\Gamma^u)$ denote the set of mixed strategies for i.[12] Given perfect recall, the equivalence between behavioral strategies and mixed strategies holds for every game. Thus in this paper I do not differentiate between mixed strategies and behavioral strategies.

An objective strategy specifies an action for each decision node, not only at decision nodes that may not be reached under this strategy, but also at those of which the player may not even be aware under this strategy. The interpretation is a subtle issue, and I postpone the discussion of it until after the discussion on solution concepts. For now, it can be thought of as a "complete contingent play" from an outsider's perspective, for example, the modeler's perspective.

Similarly, one can define the set of subjective strategies associated with each subjective game with unawareness. Fix $h \in D$, define the effective action set for each decision node $h' \in D(h)$ by $\mathring{A}(h'|h) = A(\mathbf{h}^d(h')|(h, h'))$. The set of subjective pure strategy profiles for $L(h)$ is then $S(\tilde{\Gamma}^u(h)) = \times_{h' \in D(h)} \mathring{A}(h'|h)$. For each $i \in N(h)$, the set $S_i(\tilde{\Gamma}^u(h)) = \times_{h' \in D(h), L(h')=i} \mathring{A}(h'|h)$ represents $L(h)$*'s perception* of i's complete contingent play in $L(h)$'s subjective game. Notice $L(h') = i$ himself could be unaware of some of his own future actions in $L(h)$'s subjective game, and hence need not be aware of $S_i(\tilde{\Gamma}^u(h)$. In general, for any n and any permissible q^n, the set of subjective pure strategy profiles in the subjective game with unawareness $\tilde{\Gamma}^u(q^n)$ is defined by:

$$S(\tilde{\Gamma}^u(q^n)) = \times_{h \in D(h_n|q^{n-1})} \mathring{A}(h|q^n) \tag{3.1}$$

where $\mathring{A}(h|q^n) = A(\mathbf{h}^d(h)|(q^n, h))$. Mixed strategies are defined analogously. As customary, for any strategy profile $\sigma \in \Sigma(\tilde{\Gamma}^u(q^n))$ and any $i \in N(h_n|q^{n-1})$, let $\sigma_{-i} =$

---

[11] $A3$ makes sure this is well-defined.

[12] Since players could be unaware of their own actions, objective strategies for $\Gamma^u = (N, H, L, f^u, u, W, P)$ need not coincide with strategies in the real game $\Gamma = (N, H, L, f, u, P)$. On the other hand, there is a sense in identifying objective strategies in a game with unawareness with strategies in the underlying real game. The set $\mathring{A}(h)$ can be viewed as the set of effective actions at $P(h)$. For all practical purposes, one could redefine a real game to only include the actions in $\mathring{A}(h)$ at every $h \in D$, and view the tuple (N, H, L, f^u, u, P) as an elaborated version of the real game, where spurious moves representing players' subjective perceptions of the real game are included in order to facilitate the definition of a game with unawareness.

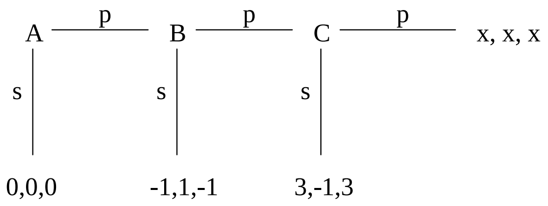

Figure 3: a game with dynamic unawareness

$\times_{j \in N(h_n|q^{n-1}), j \neq i} \sigma_j$ denote the strategy profile for all relevant players except i.

3.2 Nash solution with unawareness.

Given a standard game Γ, a strategy profile $\sigma^* \in \Sigma(\Gamma)$ is a Nash equilibrium if, for all $i \in N, \sigma_i \in \Sigma_i(\Gamma)$,

$$u_i(\sigma^*) \geq u_i(\sigma_i, \sigma^*_{-i})$$

Nash equilibrium is frequently interpreted as providing plausible predictions for the outcomes of a game, viewed as a one-shot interaction. According to this interpretation, players play best responses to their *beliefs* about opponents' equilibrium strategies, which coincide with the truth. In other words, one can rewrite the definition of Nash equilibrium as follows. Let $b^i_j \in \Sigma_j(\Gamma)$ represent i's belief about j's strategy and $b^i = \times_{j \neq i} b^i_j$. Then σ^* is a Nash equilibrium if:

1. (correct beliefs) for all $i \in N$, $b^i = \sigma^*_{-i}$;
2. (rationality) for all $i \in N$, given belief b^i, for any $\sigma_i \in \Sigma_i(\Gamma)$, $u_i(\sigma^*_i, b^i) \geq u_i(\sigma_i, b^i)$.

While rationality condition is natural and appealing, it is much less so when it comes to why players should have correct beliefs about opponents' equilibrium strategies. Common interpretations include:

1. Pre-play communication. Suppose players can engage in unlimited communication with each other, reach an agreement, and then play the game independently in isolation. Nash equilibrium characterizes the set of self-enforcing agreements.

   This interpretation is problematic in games with unawareness. Intuitively, players can use such communication opportunities strategically to convey awareness information about the game. For example, consider the simple three-step centipede game depicted in figure 3 with $x = 2$. Suppose the awareness structure is as follows:

$$W(h) = \begin{cases} \{\emptyset, s, p, (p,s)\} & \text{if } h = \emptyset \\ \{\emptyset, s, p, (p,p), (p,s), (p,p,s), (p,p,p)\} & \text{if } h = p, (p,p) \end{cases}$$

That is, Alice is unaware of Bob's action p and the continuation game following that. It is not hard to see that Bob has incentives to inform Alice of his role in the game leading to the terminal node (p, p, p), to induce Alice to play p, and then stop the game at his own information set, receiving a payoff of 1 instead of 0. But then such pre-play communication should be modeled explicitly in the game.

2. Random matching. There is a population of players, who are randomly matched to play the game repeatedly. The random matching removes the link between different stage games, so the interaction is essentially one-shot. Players observe past plays and form beliefs based on that. Though a plausible story, in games with unawareness, this process calls for an explicit modeling of how players update their awareness upon observing past play, which brings one back to the question of finding a solution concept for one-shot interaction with unawareness.

3. Mediator. Suppose a mediator makes suggestions to players about how to play the game, until no one wants to switch to a different strategy alone. Then players play the game in isolation.

   This process as described also has the problem of imposing awareness information on players. To adapt to the unawareness structure in the game, one can think of the following alternative interpretation: suppose instead of having a mediator announcing strategy profiles to the players, the mediator only reveals the part of the strategy if asked, and is free to say anything if the question does not include the action specified in the strategy. The process stops if after all questions asked and answered, no one wants to switch to a different strategy alone. Then players play the game in isolation.

This interpretation says players play best responses in their own subjective games, given "correct" beliefs which is also constrained by their own awareness. Slightly abusing notation, let $b^h_{h'} \in \Sigma(\tilde{\Gamma}^u(h))$ denote $L(h)$'s belief of $L(h')$'s strategy and $b^h = \times_{h' \in D(h)} b^h_{h'}$.

Definition 3 *Fix a game with unawareness* Γ^u *and a subjective game* $\tilde{\Gamma}$ *with set of histories* $\tilde{H} \subseteq H$. *Let* $\sigma \in \Sigma(\Gamma^u)$. *I say* $\tilde{\sigma} \in \Sigma(\tilde{\Gamma})$ *is a* **projection** *of* σ *on* $\tilde{\Gamma}$ *if:*

1. $\tilde{\sigma}(h) = \sigma(h)$ *if* $(h, a) \in \tilde{H}$ *for all* a *that is in the support of* $\sigma(h)$*;*

2. $\tilde{\sigma}(h)(a) \geq \sigma(h)(a)$ *if* $(h, a) \in \tilde{H}$ *but there exists some* b *that is in the support of* $\sigma(h)$ *but* $(h, b) \notin \tilde{H}$. *Here* $\sigma(h)(a)$ *denotes the probability the strategy* $\sigma(h)$ *assigns to the action* a.

Slightly abusing notation, let $\sigma|_{\tilde{\Gamma}}$ denote the set of projections of σ on $\Sigma(\tilde{\Gamma})$.

Definition 4 *Let* $\Gamma^u = (N, H, L, f^u, u, W, P)$ *be a game with unawareness. An objective strategy profile* $\sigma^* \in \Sigma(\Gamma^u)$ *is a* **Nash solution** *if:*

1. *(subjective correct beliefs) for all* $h \in D$, $b^h \in \sigma^*_{-h}|_{\tilde{\Gamma}(h)}$*;*
2. *(subjective rationality) for all* $h \in D$*, given belief* b^h*, for any* $\sigma_{L(h)} \in \Sigma_{L(h)}(\tilde{\Gamma}(h))$ *such that* $\sigma_{L(h)}(h') = \sigma^*_{L(h)}|_{\tilde{\Gamma}(h)}(h')$ *for all* $h' \trianglerighteq h, L(h') = L(h)$*,*

$$u_{L(h)}(\sigma^*_{L(h)}|_{\tilde{\Gamma}(h)}, b^h) \geq u_{L(h)}(\sigma_{L(h)}, b^h)$$

Subjective correct beliefs require players' beliefs to coincide with the projection of the true strategies on their subjective games. Subjective rationality requires players to play best responses to such awareness constrained consistent beliefs within their own subjective games. Notice since players may update their games during the play, subjective rationality is required for each decision node instead of each player, with the restriction that players can only revise their strategies at current and future decision nodes.

Theorem 6 *Nash solution exists.*

Example 3. Recall the game in example 2. There are three subjective games: $\tilde{\Gamma}^u(c1)$, $\tilde{\Gamma}^u((c1, s\&i))$ which are essentially standard games; and $\tilde{\Gamma}^u((c1, h\&ni))$ which is a trivial game and can be ignored. The collection of relevant strategy profiles are:

$$\begin{aligned} S(\Gamma^u) = S(\tilde{\Gamma}^u(c1)) &= \{s\&i, k\&ni\} \times \{s, k\} \times \{k\} \\ S(\tilde{\Gamma}^u((c1, s\&i)) &= \{s\&i, k\&i\} \times \{s, k\} \end{aligned}$$

There are two pure strategy Nash solutions for this game: $\{(s\&i, (k,k)), (k\&ni, (s,k))\}$. In the second solution profile, Alice plays $k\&ni$ of which Bob is unaware, and hence any belief from the space $\Delta(\{s\&i, k\&i\})$ is subjectively correct. This solution raises intriguing questions regarding the plausibility of sequential rationality in this environment. Since Alice is unaware of Bob's unawareness, from her perspective, this solution involves an "empty threat." Moreover, from her perspective, the game between herself and Bob at $(c1, s\&i)$ is essentially a standard game.

It is tempting to investigate how would an appropriate subjective version of sequential rationality impact situations such as above. Let $\mu = \{\mu^h\}_{h \in D}$ be a collection of probability measures, one for each factual information set in the real game: i.e. μ^h assigns a probability over $P(h)$. Let $u_i(\sigma; \mu|_h)$ denote i's payoffs in the continuation game following h under the strategy profile σ and beliefs μ conditional on reaching $P(h)$.

Definition 5 *Fix a game with unawareness* Γ^u*. I say an assessment* (σ^*, μ) *where* $\sigma^* \in \Sigma(\Gamma^u)$ *is a* **sequential solution** *if:*

1. *(subjective consistent beliefs) there exists a sequence of assessments* (σ_n, μ_n) *where* $\sigma_n \in \Sigma(\Gamma^u)$ *is completely mixed, and the pair satisfies:*

   (a) *beliefs are consistent: for all* $h \in D$, $(\sigma_n|_{\tilde{\Gamma}(h)}, \tilde{\mu}_n(h))$ *converges to* $(b^h, \tilde{\mu}(h))$, *where* $\tilde{\mu}(h) = \left\{\tilde{\mu}^{h'}(h)\right\}_{h' \in D(h)}$ *is a belief system where* $\tilde{\mu}^{h'}$ *assigns a probability measure over* $P(h') \cap W(h)$, *and it satisfies:*

      i. *factual consistency:* μ_n *is derived from Bayes rule;*

      ii. *awareness consistency: for all* $h' \in D(h)$, *all action sets* $A \subseteq [P(h') \cap W(h)]$, $\tilde{\mu}_n^{h'}(h)(A) \geq \mu_n^{h'}(h)(A)$, *and* $\mu_n^{h'}(h)[P(h') \cap W(h)] = 1$;

   (b) *beliefs are subjectively correct: for all* $h \in D$, $b^h \in \sigma^*_{-h}|_{\tilde{\Gamma}(h)}$.

2. *(subjective sequential rationality) for all* $h \in D$, *given* $b^h \in \Sigma_{-h}(\tilde{\Gamma}^u(h))$, *a belief system* $\tilde{\mu}(h)$, *for any* $\sigma_{L(h)} \in \Sigma_{L(h)}(\tilde{\Gamma}(h))$,

$$u_{L(h)}(\sigma^*_{L(h)}|_{\tilde{\Gamma}(h)}, b^h; \tilde{\mu}(h)|_{h'}) \geq u_{L(h)}(\sigma_{L(h)}, b^h; \tilde{\mu}(h)|_{h'})$$

   holds for all $h' \in D(h), h' \trianglelefteq h, L(h') = L(h)$.

In words, in addition to the standard consistency requirement, I require belief assessments to be confined in players' own subjective probability space; while subjective sequential rationality requires these *subjective* versions of the assessment satisfy the usual sequential rationality for future decision nodes in the subjective games.

However, requiring sequentiality in example 3 turns out to be vacuous. Both Nash solutions survive the refinement. Indeed, it is not hard to see in Bob's subjective game depicted in figure 2(c), the subjective strategy $(s,k)|_{\tilde{\Gamma}((c1, s\&i))} = s$ is sequentially rational under belief $s\&i$, which is subjectively correct given his unawareness of $(c1, k\&ni)$. On the other hand, from Alice's perspective, Bob is not sequentially rational. This is because for Alice, the subjective game depicted in figure 2(b) is common knowledge between herself and Bob. In this game, Bob's factual information set is the singleton set $s\&i$, to which Bob's best response is k.

Moreover, consider the following alternative unawareness structure in the same real game. Suppose when Bob becomes aware of the game he also becomes aware that his awareness depends on Alice's action. The game is represented as follows:

$$\begin{array}{rcl} H & = & \{\emptyset, s, k, (s,s), (s,k), (k,k)\} \\ W(h) & = & \left\{ \begin{array}{ll} \{\emptyset, s, k, (s,s), (s,k), (k,k)\} & \text{if } h = \emptyset, s \\ \{\emptyset, k, (k,k)\} & \text{if } h = k \end{array} \right. \end{array}$$

While this game has the same set of Nash solutions as the previous game, applying subjective sequential rationality does eliminate the solution $(k, (s,k))$. This example shows how the strength of sequential rationality relies on the assumption of common knowledge of the game. Since Alice and Bob at s have common knowledge of the real game, under the "correct beliefs" condition, subjective sequential rationality implies higher-order interactive beliefs of subjective sequential rationality, which in turn indicates Alice should not play k.

But Alice's hierarchy of subjective games in the two unawareness structures are identical. Alice considers her subjective game to be (subjective) common knowledge between herself and Bob in both games, yet imposing both sequential rationality under subjectively correct equilibrium beliefs has different implications for her behavior.

In fact, without common knowledge of the game, there is a tension between even mutual belief of rationality and that beliefs be correct in a solution. Consider the game depicted in figure 3 with $x = 4$. Let the unawareness structure in the game be as follows:

$$W(h) = \begin{cases} \{\emptyset, s, p, (p,p), (p,s), (p,p,s)\} & \text{if } h = \emptyset \\ \{\emptyset, s, p, (p,p), (p,s), (p,p,s), (p,p,p)\} & \text{if } h = p, (p,p) \end{cases}$$

In words, Alice is unaware that Cathy is in the game, in particular, she is unaware of the terminal node (p,p,p). One Nash solution for this game is (p,p,p). But in Alice's subjective game, the solution profile has the projection (p,p,s), where Bob chooses his own payoff -1 over 1. It seems hard to justify Bob's rationality from what Alice knows about the game.

Notice that at the initial node, Alice has a standard game that is subjective common knowledge between herself and Bob. This raises intriguing questions regarding the interpretations of Nash equilibrium and Nash solution with unawareness. The common interpretations for Nash equilibrium are typically motivated by the view that it is a steady state: *given the strategic environment described in the game*, Nash equilibrium describes the regularity in observed behavior. With common awareness of the game, the strategic environment is by definition in a steady state itself. This is no longer the case with games with unawareness. While Nash solution describes the regularity in observed behaviors *given the strategic environment*, the environment itself is not in a steady state. In the above example, arguably Alice updates her subjective game to the real game when she receives a payoff of 4 instead of 3. Indeed, given her updated subjective game, Bob is rational.

Intuitively, outcomes inconsistent with mutual beliefs of subjective rationality are those involving "surprises," and hence causing players to update their subjective games. In this sense, a steady state interpretation for games with unawareness requires a solution concept that involves no surprise outcomes.

3.3 Nash steady state with unawareness.

It seems compelling to require players to be *at least* aware of *some* play path leading to the equilibrium payoffs, on on-equilibrium path. Fix $\sigma \in \Sigma(\Gamma^u)$. Let $Z(\sigma)$ denote the set of terminal histories under σ. Slightly abusing terminology, I refer to a terminal history z and the play path leading to z, i.e. $\{h : h \trianglerighteq z\}$, interchangeably when confusions are not likely to arise.

Definition 6 *Two play paths $z, z' \in Z$ are* **indistinguishable** *for $L(h)$, denoted by $z =_h z'$, if:*

1. *$L(h)$ receives the same payoff: $u_{L(h)}(z) = u_{L(h)}(z')$;*

2. *The two play paths go through the same factual information sets: $\{P(h) : h \trianglerighteq z\} = \{P(h) : h \trianglerighteq z'\}$;*

3. *The two play paths coincide at every decision node of which $L(h)$ is aware: $(h', a) \trianglerighteq z'$, $\#[A(P(h')|h)] > 1 \Rightarrow (h', a) \trianglerighteq z$.*

The set of "surprise" outcomes the strategy profile σ induces for $L(h)$ is:

$$Z_h(\sigma) := \{z \in z(\sigma) : \nexists\, z' \in W(h) \text{ such that } z' =_h z\} \tag{3.2}$$

Players along the play path update their awareness upon observing the outcome of σ. Let $R(\sigma) = \{h \trianglelefteq h_0 : h_0 \trianglerighteq z \text{ for some } z \in Z(\sigma), L(h) = L(h_0)\}$ denote the set of the relevant decision nodes. The updated awareness function is:

$$W_\sigma(h) = \begin{cases} W(h) \cup \{h' \in D : h' \trianglerighteq z \text{ for some } z \in Z_h(\sigma)\} & \text{if } h \in R(\sigma) \\ W(h) & \text{otherwise.} \end{cases} \tag{3.3}$$

The function W_σ satisfies $A0 - 5$, thus the game with unawareness is well-defined under W_σ. I explore two steady state solution concepts.

Definition 7 *Let $\Gamma^u = (N, H, L, f^u, u, W, P)$ be a game with unawareness. An objective strategy profile $\sigma^* \in \Sigma(\Gamma^u)$ is a* **weak Nash steady state** *if there exists a sequence of strategy profiles $\sigma_1, \cdots, \sigma_n$ such that:*

1. *$\sigma_{k+1} \in \Sigma(\Gamma^u_k)$ is a Nash solution of Γ^u_k, where: $\Gamma^u_1 = \Gamma^u$, $W^1 = W_{\sigma^1}$; for all $k = 1, \cdots, n-1$, $W^k = W^{k-1}_{\sigma_k}$, $\Gamma^u_k = (N, H, L, f^u, u, W^k, P)$;*

2. *σ^* is a Nash solution of $\Gamma^u_n = (N, H, L, f^u, u, W^n, P)$;*

3. *$W^n = W^n_{\sigma^*}$.*

I call Γ^u_n a stable game with unawareness. If σ^* is a Nash solution in all stable games with unawareness, then it is called a **strong Nash steady state**.

Theorem 7 *Weak Nash steady state exists.*

Proof. This follows easily from the existence result of Nash solution and the observation that $(W^1, \cdots, W^n)$ is an increasing sequence with a point-wise upper bound H, and hence converges point-wise. □

4 Appendix

4.1 Proof for Lemma 1.

Proof.

$A0(h)$: $h' \in W(h), L(h') = c \Rightarrow W(h'|h) = W(h)$.

By $A0$, $W(h') = H$. It follows $W(h'|h) = W(h) \cap H = W(h)$.

$A1(h)$: $W(\cdot|h)$ satisfies initial history, tree and infinite history for all $h' \in D(h)$.

Suffices to note that the union and intersection operation preserve the game tree structure.

$A2(h)$: for any $h' \in D(h)$, $\exists \mathbf{h}^d(h') \in P(h'|h), \mathbf{a}^d(h') \in A(h'|h)$ such that,

1. $(\mathbf{h}^d(h'), \mathbf{a}^d(h')) \in W(h'|h)$;
2. $h' \in W(h''|h) \Rightarrow (\mathbf{h}^d(h'), \mathbf{a}^d(h')) \in W(h''|h)$.

By $A2.1$, the default history and action for h' is in $W(h')$; by $A2.2$, since $h' \in W(h)$, the default history and action is in $W(h)$ too: $(\mathbf{h}^d(h'), \mathbf{a}^d(h')) \in W(h') \cap W(h) \subseteq W(h'|h)$. This proves $A2(h).1$. For $A2(h).2$, observe that $h' \in W(h''|h) \Rightarrow h' \in W(h) \cap W(h_0)$ for $h_0 = h''$ or $h_0 \in T(h''|h)$. But then by $A2.2, (\mathbf{h}^d(h'), \mathbf{a}^d(h')) \in W(h) \cap W(h_0) \subseteq W(h''|h)$.

$A3(h) : h_1 \in P(h_2), h_1, h_2 \in D(h) \Rightarrow (W(h_1|h), P(h_1|h)) = (W(h_2|h), P(h_2|h))$.

By $A3, (W(h_1), P(h_1)) = (W(h_2), P(h_2))$. By definition of $P(\cdot|h)$, $P(h_1|h) = P(h_1) \cap D(h) = P(h_2) \cap D(h) = P(h_2|h)$.

Consider $T(h_1|h)$ and $T(h_2|h)$. If they are empty then the claim follows from $W(h_1) = W(h_2)$. Suppose $T(h_1|h) \neq \emptyset$ and let $h'_1 \in T(h_1|h)$. Write them as $h'_1 = (\bar{h}, a, h^t_1), h_1 = (\bar{h}_1, b, h^t_1)$ where $a \sim b$. It follows $(\bar{h}, a) \notin W(h_1)$. On the other hand, $h_1 \in P(h_2), h_1 \approx h'_1 \Rightarrow$ there exists $h'_2 \trianglelefteq (\bar{h}, a), h'_2 \approx h_2, h'_2 \in P(h'_1)$. It follows $h'_2 \in T(h_2|h), W(h'_2) = W(h'_1)$. Thus $\cup_{h'_1 \in T(h_1|h)} W(h'_1) = \cup_{h'_2 \in T(h_2|h)} W(h'_2)$, and hence $W(h_1|h) = W(h_2|h)$.

$A4(h) : L(\bar{h}) = c, h_1 = (\bar{h}, a, h'), h_2 = (\bar{h}, b, h'), L(h_1) = L(h_2), W(h_1|h) \neq W(h_2|h) \Rightarrow \nexists h'_1 \trianglelefteq (\bar{h}, a), h'_2 \trianglelefteq (\bar{h}, b)$ and $h'_1 \in P(h'_2|h)$.

Suppose not. Let $h'_1 \trianglelefteq (\bar{h}, a), h'_2 \trianglelefteq (\bar{h}, b)$ and $h'_1 \in P(h'_2|h)$. Since $P(h'_2|h) \subseteq P(h'_2)$, we have $h'_1 \in P(h'_2)$. By $A4$, we must have $W(h_1) = W(h_2)$. I derive a contradiction by proving $W(h_1|h) = W(h_2|h)$.

Consider $T(h_1|h)$ and $T(h_2|h)$. The claim is trivially true, if (1) $T(h_1|h) = T(h_2|h) = \emptyset$; or (2) $W(\tilde{h}_i) = W(h_i)$ for all $\tilde{h}_i \in T(h_1|h)$, $i = 1, 2$. Thus suppose $\tilde{h}_1 \in T(h_1|h)$, i.e. $\tilde{h}_1 \approx h_1$ and $\tilde{h}_1 \notin W(h)$; and $W(\tilde{h}_1) \neq W(h_1)$.

Let $\wedge$ denote the operation of taking the maximal common history. That is, for general histories h^1, h^2, $[h^1 \wedge h^2] \trianglerighteq h^1, [h^1 \wedge h^2] \trianglerighteq h^2$ and there exists no h^0 such that $h^0 \trianglerighteq h^1, h^0 \trianglerighteq h^2$ and $[h^1 \wedge h^2] \trianglerighteq h^0$.

Suppose $[\tilde{h}_1 \wedge h_1] \trianglelefteq (\bar{h}, a)$, that is, the mirror action leading to $\tilde{h}_1$ takes place after the chance move. But then since $\tilde{h}_1 \trianglelefteq (\bar{h}, a), \tilde{h}_1 \approx h_1$ implies that a relabeling of the mirror action would violate $A4$: $W(\tilde{h}_1) \neq W(h_1) = W(h_2)$.

Thus $[\tilde{h}_1 \wedge h_1] \trianglerighteq \bar{h} \trianglerighteq h_1'$, that is, the mirror action takes place before the chance move, hence before the information set $P(h_1')$. Then we can trace $\tilde{h}_1$ back to find a mirror history for h_1': $\exists \tilde{h}_1' \approx h_1', \tilde{h}_1' \trianglerighteq \tilde{h}_1$. But then the information set pulls in h_2 as part of the continuation game, and hence expand the set of mirror histories to include those for h_2 and its predecessors, which are linked with the mirror histories for h_1 and its predecessors again through the mirror image of the information set $P(h_1')$. Formally, $h_2' \in P(h_1'), \tilde{h}_1' \approx h_1' \Rightarrow$ there exists $\tilde{h}_2'$ such that $\tilde{h}_2' \in P(\tilde{h}_1'), \tilde{h}_2' \approx h_2'$ and $\tilde{h}_2 \approx h_2$. Apply $A4$ on $\tilde{h}_1, \tilde{h}_2$, we must have $W(\tilde{h}_1) = W(\tilde{h}_2)$. Since $[\tilde{h}_1 \wedge h_1] = [\tilde{h}_2 \wedge h_2]$, $\tilde{h}_2 \notin W(h) \Rightarrow \tilde{h}_2 \in T(h_2|h)$. Therefore $W(h_1) \cup_{\tilde{h}_1 \in T(h_1|h)} W(\tilde{h}_1) = W(h_2) \cup_{\tilde{h}_2 \in T(h_2|h)} W(\tilde{h}_2) \Rightarrow W(h_1|h) = W(h_2|h)$.

$A5(h) : h' \in D(h), \bar{h}' \trianglerighteq h', L(\bar{h}') = L(h') \Rightarrow W(\bar{h}'|h) \subseteq W(h'|h)$.

Consider $T(\bar{h}'|h)$ and $T(h'|h)$. Let $\bar{h}_\alpha \in T(\bar{h}'|h)$. Then $\bar{h}_\alpha \approx \bar{h}'$, $\bar{h}_\alpha \notin W(h)$. By definition of mirror action, there exists $h_\alpha \trianglelefteq \bar{h}_\alpha$ and $h_\alpha \approx h'$. By $A1$, $h_\alpha \notin W(h) \Rightarrow h_\alpha \in T(h'|h)$. By $A5, W(\bar{h}_\alpha) \subseteq W(h_\alpha), W(\bar{h}') \subseteq W(h')$. It follows $W(\bar{h}') \cup_{\bar{h}_\alpha \in T(\bar{h}'|h)} W(\bar{h}_\alpha) \subseteq W(h') \cup_{h_\alpha \in T(h'|h)} W(h_\alpha) \Rightarrow W(\bar{h}'|h) \subseteq W(h'|h)$.

$\square$

4.2 Definition for higher-order subjective games (with unawareness).

For any n and any permissible q^n, $W(h_n|q^{n-1})$ consists of a game tree by applying Lemma 1 recursively. Let $Z(h_n|q^{n-1})$ denote the set of terminal histories for $W(h_n|q^{n-1})$ and let $D(h_n|q^{n-1}) \equiv W(h_n|q^{n-1}) \setminus Z(h_n|q^{n-1})$. The action sets are defined by: for any $h \in W(h_n|q^{n-1}), A(h|q^n) = \{a : (h, a) \in W(h_n|q^{n-1})\}$. Then, $L(h_1)$'s perception of $L(h_2)$'s perception of $\cdots$ of $L(h_n)$'s subjective game is the tuple

$$\tilde{\Gamma}(q^n) = (N(h_n|q^{n-1}), W(h_n|q^{n-1}), L(\cdot|q^n), f_{q^n}, u_{q^n}, P(\cdot|q^n))$$

where:

$$\begin{aligned}
N(h_n|q^{n-1}) &= \{L(h) : h \in D(h_n|q^{n-1}), \#[A(h|q^n)] > 1\} \\
L(\cdot|q^n) &= L_{|D(h_n|q^{n-1})} \\
f_{q^n} &= \{f(\cdot|h, A(h_n|q^{n-1}))\}_{h\in\{h\in W(h_n|q^{n-1}):L(h)=c\}} \\
u_{q^n} &= \{u_i\}_{i\in N(h_n|q^{n-1})} \\
P(\cdot|q^n)) &= P_{|D(h_n|q^{n-1})}
\end{aligned}$$

Let $\mathcal{A}(h|q^n)$ denote the induced sub-σ-algebra of $\mathcal{A}(h)$ on $A(h|q^n) \subseteq A(h)$, i.e. $\mathcal{A}(h|q^n) \equiv \mathcal{A}(h)|_{A(h|q^n)}$. The collection of probability measures for every possible probability space in this game is the restriction of f^u to the sets of chance moves in the game $\tilde{\Gamma}(q^n)$, and is defined by:

$$f^u_{q^n} = \{f(\cdot|h, A)\}_{A\in\mathcal{A}(h|q^n), h\in\{h\in W(h_n|q^{n-1}):L(h)=c\}}$$

Finally, $L(h_1)$'s perception of $L(h_2)$'s perception of $\cdots$ of $L(h_n)$'s subjective game with unawareness is the tuple

$$\tilde{\Gamma}^u(q^n) = (N(h_n|q^{n-1}), W(h_n|q^{n-1}), L(\cdot|q^n), f^u_{q^n}, u_{q^n}, W(\cdot|q^n), P(\cdot|q^n))$$

4.3 Proof for Theorem 4.

Proof.

Given the recursive structure, it suffices to show for all $h \in W(h_j|q^{j-1}), h_i \trianglerighteq h_j, L(h_i) = L(h_j), 1 \le i < j$ implies that,

$$W(h|q^j) = W(h|q^{j-1})$$

The key to this result is that the conditions imply

$$W(h_j|q^{j-1}) = W(h_{j-1}|q^{j-2}) \tag{4.1}$$

For ease of exposition, I first show equation (4.1) implies the result, then prove equation (4.1).

$$\begin{aligned}
W(h|q^j) &= W(h_j|q^{j-1}) \cap [W(h|q^{j-1}) \bigcup_{h'\in T(h|q^j)} W(h'|q^{j-1})] \\
&= W(h_{j-1}|q^{j-2}) \cap [W(h|q^{j-1}) \bigcup_{h'\in T(h|q^j)} W(h'|q^{j-1})]
\end{aligned}$$

But observe equation (4.1) implies

$$\begin{aligned}
T(h|q^j) &= \{h' \in D(h_{j-1}|q^{j-2}) \setminus D(h_j|q^{j-1}) : h' \approx h\} \\
&= \{h' \in D(h_{j-1}|q^{j-2}) \setminus D(h_{j-1}|q^{j-2}) : h' \approx h\} \\
&= \emptyset
\end{aligned}$$

Therefore $W(h|q^j) = W(h_{j-1}|q^{j-2}) \cap W(h|q^{j-1})$. But then by definition (2.5), higher-order interactive awareness are nested with respect to the reasoning sequence, i.e. $W(h|q^{j-1}) \subseteq W(h_{j-1}|q^{j-2})$. It follows $W(h|q^j) = W(h|q^{j-1})$.

I prove equation (4.1) via induction on the distance between i and j.
Step 1. Let $j = i+1$. For notational ease, let $h_{n+1} = h$. Need to show:

$$h_i \trianglerighteq h_{i+1}, L(h_i) = L(h_{i+1}) \Rightarrow W(h_{i+1}|q^i) = W(h_i|q^{i-1})$$

Claim 1.1 For any n, any permissible q^n, $h \trianglerighteq h', L(h) = L(h'), h, h' \in W(h_n|q^{n-1})$ implies $W(h|q^n) \subseteq W(h'|q^n)$.

Let $n = 1$.

$$\begin{aligned} W(h|h_1) &= W(h_1) \cap [W(h) \underset{h_0 \in T(h|h_1)}{\cup} W(h_0)] \\ W(h'|h_1) &= W(h_1) \cap [W(h') \underset{h'_0 \in T(h'|h_1)}{\cup} W(h'_0)] \end{aligned}$$

By $A5$, $W(h) \subseteq W(h')$. Let $h_0 \in T(h|h_1)$. Then $h_0 \approx h, h_0 \notin W(h_1)$. By definition of mirror histories, there exists $h'_0 \trianglelefteq h_0, h'_0 \approx h'$. By $A1$, $h'_0 \notin W(h_1)$. Thus $h'_0 \in T(h'|h_1)$. Again by $A5$, $W(h_0) \subseteq W(h'_0)$. It follows that

$$[W(h) \underset{h_0 \in T(h|h_1)}{\cup} W(h_0)] \subseteq [W(h') \underset{h'_0 \in T(h'|h_1)}{\cup} W(h'_0)]$$

Therefore we have $W(h|h_1) \subseteq W(h'|h_1)$.

Now suppose the claim is true for all permissible sequences of orders up to $n-1$. Expand both sides of the expression for order n:

$$\begin{aligned} W(h|q^n) &= W(h_n|q^{n-1}) \cap [W(h|q^{n-1}) \underset{h_0 \in T(h|q^n)}{\cup} W(h_0|q^{n-1})] \\ W(h'|q^n) &= W(h_n|q^{n-1}) \cap [W(h'|q^{n-1}) \underset{h'_0 \in T(h'|q^n)}{\cup} W(h'_0|q^{n-1})] \end{aligned}$$

Compare $T(h|q^n)$ and $T(h'|q^n)$.

$$\begin{aligned} T(h|q^n) &= \left\{h_0 \in D(h_{n-1}|q^{n-2}) \setminus D(h_n|q^{n-1}) : h_0 \approx h\right\} \\ T(h'|q^n) &= \left\{h'_0 \in D(h_{n-1}|q^{n-2}) \setminus D(h_n|q^{n-1}) : h'_0 \approx h'\right\} \end{aligned}$$

By arguments similar to the case of $n = 1$, for every $h_0 \in T(h|q^n)$, there must exist $h'_0 \in T(h'|q^n)$ such that $h_0 \trianglerighteq h'_0$. Hence by $A5$ and the induction hypothesis,

$$[W(h|q^{n-1}) \underset{h_0 \in T(h|q^n)}{\cup} W(h_0|q^{n-1})] \subseteq [W(h'|q^{n-1}) \underset{h'_0 \in T(h'|q^n)}{\cup} W(h'_0|q^{n-1})]$$

It follows that $W(h|q^n) \subseteq W(h'|q^n)$ and hence the claim.

Therefore, $h_i \trianglerighteq h_{i+1}, L(h_i) = L(h_{i+1})$ implies that

$$\begin{aligned} W(h_{i+1}|q^i) &= W(h_i|q^{i-1}) \cap [W(h_{i+1}|q^{i-1}) \bigcup_{h \in T(h_{i+1}|q^i)} W(h|q^{i-1})] \\ &= W(h_i|q^{i-1}) \end{aligned}$$

Step 2. Now suppose $W(h_j|q^{j-1}) = W(h_{j-1}|q^{j-2})$ holds if $h_i \trianglerighteq h_j, L(h_i) = L(h_j)$ and $j - i \leq m - 1$. Suppose $h_i \trianglerighteq h_j, L(h_i) = L(h_j)$ and $j = i + m$.

$$W(h_j|q^{j-1}) = W(h_{j-1}|q^{j-2}) \cap [W(h_j|q^{j-2}) \bigcup_{h \in T(h_j|q^{j-1})} W(h|q^{j-2})]$$

By induction hypothesis, $W(h_j|q^{j-2}) = W(h_{j-2}|q^{j-3})$. But then since interactive awareness are nested, $W(h_{j-1}|q^{j-2}) \subseteq W(h_{j-2}|q^{j-3})$, and hence $W(h_j|q^{j-1}) = W(h_{j-1}|q^{j-2})$. This concludes the proof. □

4.4 Proof for Theorem 6.

Proof. Without loss of generality, let $L(h) \neq L(h')$ whenever $h \neq h'$. For simplicity here I use h as the index for players. The set of objective strategy profiles is $\Sigma(\Gamma^u) = \times_{h \in D} \mathring{A}(h)$ where $\mathring{A}(h) = A(h|h) \subseteq A(h)$. However, $L(h')$'s perception of $L(h)$'s strategy may involve actions in $A(h) \setminus A(h|h)$. Let $\Sigma(\Gamma) = \times_{h \in D} A(h)$ denote the collection of all strategy profiles.

Let $P_h : \Sigma(\Gamma^u) \rightrightarrows \Sigma(\tilde{\Gamma}^u_h)$ denote h's the projection correspondence, where $\Sigma(\tilde{\Gamma}^u_h) = \times_{h' \in D(h)}(\Delta(A(h')) \cap \Delta[A(h'|h)])$ is the set of strategy profiles in $L(h)$'s subjective game. Intuitively, $P_h(\sigma)$ is the set of possible perceptions $L(h)$ could have for an objective strategy profile σ. By definition 3, P_h is continuous. By $A2$, it is non-empty.

$L(h)$'s *subjective* best response correspondence, denoted by B_h, maps strategy profiles in his own subjective game to the space of his own strategies. By definition of the objective strategies, $\Sigma_h(Gamma^u) = \Sigma_h(\tilde{\Gamma}_h)$. Thus $B_h : \Sigma(\tilde{\Gamma}^u_h) \rightrightarrows \Sigma_h(\Gamma^u)$. Now consider the correspondence $B \circ R = (B_h \circ P_h)_{h \in D} : \Sigma(\Gamma^u) \rightrightarrows \Sigma(\Gamma^u)$. This correspondence satisfies all standard assumptions and hence has a fixed point.

□

References

Aumann, Robert J., "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 1987, *55* (1), 1–18.

Brandenburger, Adam and Eddie Dekel, "Hierarchies of Beliefs and Common Knowledge," *Journal of Economic Theory*, 1993, *59*, 189–198.

Feinberg, Yossi, “Subjective Reasoning – Games with Unawareness,” 2004. Working paper, Stanford University.

——, “Games with Incomplete Unawareness,” 2005. Working paper, Stanford University.

Li, Jing, “Information Structures with Unawareness,” 2006a. Working paper, University of Pennsylvania.

——, “Modeling Unawareness Without Impossible States,” 2006b. Working paper, University of Pennsylvania.

Osborne, Martin J. and Ariel Rubinstein, *A Course in Game Theory*, The MIT Press, 1994.