Hierarchies of conditional beliefs derived from commonly known priors

Maastricht University

Elias Tsakas

Hierarchies of conditional beliefs derived from commonly known priors

RM/11/021

METEOR

Maastricht University School of Business and Economics
Maastricht Research School of Economics
of Technology and Organization

P.O. Box 616
NL - 6200 MD Maastricht
The Netherlands

Hierarchies of conditional beliefs derived from commonly known priors*

Elias Tsakas†

Department of Economics, Maastricht University

March 16, 2011

Abstract

In this paper, we consider a finite set of agents with commonly known full-support priors on the fundamental space of uncertainty. Then, we show that if the hierarchies of conditional beliefs á la Battigalli and Siniscalchi (1999) are derived from these priors, then each agent's hierarchy is commonly known, and consequently all types of the same agent yield the same hierarchy. We also show that the previous result does not necessarily hold when the priors are not full-support. Moreover, if the collections of conditioning events does not cover the underlying space of uncertainty, there are always commonly known (non-full-support) priors such that every agent's conditional belief hierarchies are derived from these priors.

Keywords: Epistemic game theory, hierarchies of conditional beliefs, prior beliefs, common knowledge.

JEL Classification: C70, D80, D81, D82.

1. Introduction

Hierarchies of conditional beliefs are an integral tool of modern economic theory. They were first introduced by Battigalli and Siniscalchi (1999), and were proven extremely useful for the epistemic analysis of solution concepts in extensive form games, e.g., Battigalli and Siniscalchi (2002) provide an epistemic characterization of extensive form rationalizable outcomes, whereas Battigalli and Friedenberg (2010) epistemically characterize extensive form best response sets.

*Financial support from the Marie Curie Fellowship (PIEF-GA-2009-237614) is gratefully acknowledged.

†P.O. Box 616, 6200 MD, Maastricht, The Netherlands; Tel: +31-43-38 83649; Fax: +31-43-38 84878; E-mail: e.tsakas@maastrichtuniversity.nl

Conditional beliefs generalize usual probabilistic beliefs, in that every agent is endowed with a collection of conditioning events[1], and forms conditional beliefs given each hypothesis, in a way that Bayesian updating is satisfied whenever possible. Such a collection of measures is called conditional probability system[2] (Rênyi, 1955). An agent's first order conditional beliefs are described by a conditional probability system over the fundamental space of uncertainty[3]; the second order conditional beliefs are described by a conditional probability system over the space of the opponents' first order conditional beliefs; and so on. Obviously, conditional belief hierarchies are very complex objects, thus making it quite hard to work with them. In an attempt to make them more tractable, Battigalli and Siniscalchi (1999) represent them with a compact (type-based) model which mimics Harsanyi's representation of usual belief hierarchies. Namely, for each agent there is a set of types. Each type is associated with a conditional probability system over the product of the fundamental space of uncertainty and the set of the opponents' types. This construction induces a hierarchy of conditional beliefs for each type.

We say that an agent's conditional beliefs are derived from some prior, q, over the fundamental space of uncertainty, whenever the first order conditional beliefs given any non-q-null conditioning event are derived by applying Bayes law on q. The prior captures the beliefs that the agent has about the fundamental space of uncertainty before having received any additional information in the form of some conditioning event (Aumann, 1998). That is, a probabilistic assessment about the natural world, which is represented by the fundamental space of uncertainty, is the product of two sources of information, the one embodied in the prior, and the one contained in the conditioning event. Quite often, it is suggested that differences in probabilistic beliefs should be necessarily attributed to different conditioning events. That is, the prior beliefs of an agent are commonly known, i.e., everybody knows that the agent derives her beliefs from this prior, everybody knows that everybody knows this, and so on. In principle, we are neither in favor, nor against this view. Our aim is to examine how restrictive this condition really is.

The main result of the paper (Theorem 1) shows that if it is common knowledge that every agent's beliefs are derived from some full-support prior – not necessarily common for all agents – then every type of this agent has the same hierarchy of conditional beliefs. This result, though rather simple to prove, is a bit surprising, as it completely rules out information asymmetries. The intuition is as follows: Recall that, whenever the priors are commonly known, differences in beliefs

[1]In an extensive form game, a player's collection of conditioning events, else called conditioning hypotheses, would typically correspond to the information sets controlled by the player.

[2]For the formal definition, see Section 3.

[3]In epistemic game theory, we usually take the strategy space as the fundamental space of uncertainty, whereas in incomplete information games the latter coincides with the set of possible payoff functions.

should be attributed to different conditioning events. However, Battigalli and Siniscalchi (1999) construct the type space in a way such that all types of the same agent share the same conditioning hypotheses. Therefore, two types can in principle differ only in the conditional beliefs given zero-probability conditioning events. However, the latter is not possible if the (commonly known prior) is full-support, as in this case no conditioning hypothesis is a null event. In fact, we illustrate (in Example 1) that the previous result is tight with respect to the full-support assumption. Moreover, we show (Theorem 2) that for an arbitrary type space, if the collection of conditioning hypotheses does not cover the fundamental space of uncertainty, there are always non-full-support priors such that every it is commonly known that the beliefs are derived from these priors. The intuition here is that these priors are concentrated on the states that are not contained in any conditioning hypothesis, implying that all conditioning events are null-events, and therefore are not restricted by the priors.

The paper is organized as follows: Section 2 contains some mathematical preliminaries; Section 3 presents the type-based representation of conditional belief hierarchies; Section 4 contains our results.

2. Preliminaries

We present some preliminaries on Polish spaces. For further reference see Kechris (1995). A topological space $(Z, \mathcal{T})$ is called Polish if it is separable and completely metrizable. The countable product of Polish spaces, endowed with the product topology, is Polish. A closed subspace of a Polish space, endowed with the relative topology, is also Polish.

For any topological space Z, let $\Delta(Z)$ denote the set of all Borel probability measures, endowed with the topology of weak convergence. If Z is Polish then so is $\Delta(Z)$. For some $p \in \Delta(Z)$, let $\Gamma(p)$ denote its support, i.e., the set of all points $z \in Z$ such that every $T \in \mathcal{T}$ with $z \in T$ has positive probability: $\Gamma(p) = \{z \in Z : z \in T \in \mathcal{T} \Rightarrow p(T) > 0\}$. The support is the smallest closed subset of Z with measure equal to 1. If Z is separable and metrizable, the support is unique (Parthasarathy, 1967, pp. 27–28).

3. Hierarchies of conditional beliefs

Let $(\Sigma, \mathcal{A})$ be a measurable space, where Σ is Polish, $\mathcal{A}$ is the Borel σ-algebra, and let $\mathcal{B} \subseteq \mathcal{A} \setminus \{\emptyset\}$ be a collection of non-empty clopen[4] conditioning events (not necessarily an algebra). Throughout the paper we assume that $\mathcal{B}$ is a countable collection of subsets.

[4]The assumption about the elements of $\mathcal{B}$ being both closed and open is rather standard. For further discussion, see Battigalli and Siniscalchi (1999, p. 191).

Definition 1. A conditional probability system (CPS) on $(\Sigma, \mathcal{A}, \mathcal{B})$ is a function $\pi : \mathcal{A} \times \mathcal{B} \to [0,1]$ that satisfies the following properties:

(C_1) $\pi(B|B) = 1$, if $B \in \mathcal{B}$,

(C_2) $\pi(\cdot|B)$ is a probability measure over $(\Sigma, \mathcal{A})$ for every $B \in \mathcal{B}$,

(C_3) $\pi(A|C) = \pi(A|B) \times \pi(B|C)$, if $A \subseteq B \subseteq C$, and $A \in \mathcal{A}$ and $B, C \in \mathcal{B}$.

The underlying idea behind this construction is as follows: There is a fundamental space of uncertainty, Σ, which can be thought as a collection of all possible values that some objective parameters may take, e.g., the set of possible payoff functions in an incomplete information game, or the set of action profiles in a game. The collection $\mathcal{B}$ contains all the conditioning events[5] that an agent could possibly observe. Every conditioning event, $B \in \mathcal{B}$, yields a probability distribution, $\pi(\cdot|B)$, over the measurable space $(\Sigma, \mathcal{A})$, which corresponds to the conditional beliefs given B. Throughout the paper, for notation simplicity, we often skip $\mathcal{A}$ and we simply write $(\Sigma, \mathcal{B})$. Conditional probability systems are due to Rênyi (1955), and were first introduced in a game-theoretic framework by Myerson (1986).

Let $\Delta^{\mathcal{B}}(\Sigma)$ denote the space of CPS's over $(\Sigma, \mathcal{B})$. Observe that $\Delta^{\mathcal{B}}(\Sigma)$ is a subspace of $[\Delta(\Sigma)]^{\mathcal{B}}$, which is a Polish space, endowed with the product topology[6] (of weak convergence of measures). Then, it follows from $\Delta^{\mathcal{B}}(\Sigma)$ being a closed subset of $[\Delta(\Sigma)]^{\mathcal{B}}$, that it is also a Polish space endowed with the relative topology (Battigalli and Siniscalchi, 1999, Lemma 1).

Let $I = \{a, b\}$ be the set of individuals[7], with typical elements i and j. Agent i is endowed with a collection of conditioning hypotheses, $\mathcal{B}_i$. Agent i's first order beliefs consist of a CPS over Σ, thus inducing a collection of probability measures – one for every conditioning hypothesis. The second order beliefs consist of a CPS over $\Sigma \times \Delta^{\mathcal{B}_j}(\Sigma)$, thus inducing a collection of measures – one for each conditioning hypothesis – over the opponent's first order beliefs.

[5] In an extensive form game, $\mathcal{B}$ often corresponds to the collection of information sets.

[6] As usual, the space of probability measures, $\Delta(\Sigma)$, is endowed with the topology of weak convergence. Since Σ is Polish, so is $\Delta(\Sigma)$ (Aliprantis and Border, 1994, p. 515). Then, the product space $[\Delta(\Sigma)]^{\mathcal{B}}$ is also a Polish space, endowed with the product topology (Kechris, 1995, p.).

[7] Our analysis can be directly generalized to any finite set of individuals. Throughout the paper, we stick to two individuals for notation simplicity.

Formally, consider the following sequence:

$$\begin{aligned}
\Theta_i^0 &:= \Sigma & \mathcal{B}_i^0 &:= \mathcal{B}_i \\
\Theta_i^1 &:= \Theta_i^0 \times \Delta^{\mathcal{B}_j^0}(\Theta_j^0) & \mathcal{B}_i^1 &:= \{B_i^0 \times \Delta^{\mathcal{B}_j^0}(\Theta_j^0) \subseteq \Theta_i^1 \mid B_i^0 \in \mathcal{B}_i^0\} \\
&\vdots & &\vdots \\
\Theta_i^{n+1} &:= \Theta_i^n \times \Delta^{\mathcal{B}_j^n}(\Theta_j^n) & \mathcal{B}_i^{n+1} &:= \{B_i^n \times \Delta^{\mathcal{B}_j^n}(\Theta_j^n) \subseteq \Theta_i^{n+1} \mid B_i^n \in \mathcal{B}_i^n\} \\
&\vdots & &\vdots
\end{aligned}$$

For all $n \geq 0$, the CPS $\pi_i^{n+1} \in \Delta^{\mathcal{B}_i^n}(\Theta_i^n)$ denotes agent i's $(n+1)$-th order conditional beliefs, and $\pi_i := (\pi_i^1, \pi_i^2, ...) \in \times_{n=0}^{\infty} \Delta^{\mathcal{B}_i^n}(\Theta_i^n)$ is i's hierarchy of (conditional) beliefs, with $\pi_i^{n+1}(\cdot|B_i^n)$ denoting the conditional measure given the hypothesis $B_i^n \in \mathcal{B}_i^n$.

Observe that, by construction, every element in $\mathcal{B}_i^n$ is a clopen cylinder generated by some element of $\mathcal{B}_i^0$ (Battigalli and Siniscalchi, 1999, p. 193), and therefore i's $(n+1)$-th order conditional beliefs given $B_i^n = B_i^0 \times \Delta^{\mathcal{B}_j^0}(\Theta_j^0) \times \cdots \times \Delta^{\mathcal{B}_j^{n-1}}(\Theta_j^{n-1})$ are essentially determined by the event $B_i^0 \in \mathcal{B}_i^0$. Throughout the paper, with slight abuse of notation, we write $\Delta^{\mathcal{B}_i}(\Theta_i^n)$ instead of $\Delta^{\mathcal{B}_i^n}(\Theta_i^n)$, and $\pi_i^n(\cdot|B_i)$ instead of $\pi_i^n(\cdot|B_i^n)$ (Battigalli and Siniscalchi, 1999, p. 194).

As usual, we impose the standard coherency restriction, which, roughly speaking, says that i's n-th order beliefs do not contradict i's $(n-1)$-th order beliefs. Formally, some $\pi_i \in \times_{n=0}^{\infty} \Delta^{\mathcal{B}_i}(\Theta_i^n)$ is coherent whenever

$$\text{marg}_{\Theta_{n-2}} \pi_i^n = \pi_i^{n-1},$$

with $\text{marg}_{\Theta_{k-2}} \pi_i^n := \big(\text{marg}_{\Theta_{k-2}} \pi_i^n(\cdot|B_i) \ ; \ B_i \in \mathcal{B}_i\big)$. We focus on the hierarchies that satisfy, not only coherency, but also common certainty in coherency. That is, we restrict attention to types that (1) are coherent, (2) assign probability 1 (given every conditioning hypothesis) to the event that the opponent's beliefs are coherent, and so on. We denote the set of these hierarchies by T_i^*, else called i's universal type space. Battigalli and Siniscalchi (1999) show that there is a homeomorphism $T_i^* \mapsto \Delta^{\mathcal{B}_i^*}(\Sigma \times T_j^*)$, implying that every belief hierarchy in T_i^* is associated with a unique CPS over $(\Sigma \times T_j^*, \mathcal{B}_i^*)$, where $\mathcal{B}_i^* := \{B_i \times T_j^* \mid B_i \in \mathcal{B}_i\}$ denotes i's conditioning hypotheses[8] in $\Sigma \times T_j^*$.

Hierarchies of conditional beliefs are very complex objects, and therefore it is quite hard working with them. Following the standard practice first introduced by Harsanyi (1967-68) in a slightly different framework, Battigalli and Siniscalchi (1999) introduce a compact way of representing them with a type space model.

Definition 2. A Σ-based type space model (T-space) is a tuple $(\Sigma, \mathcal{B}_a, \mathcal{B}_b, T_a, T_b, g_a, g_b)$, where Σ is Polish, $\mathcal{B}_i$ is a collection of non-empty clopen subsets of Σ, T_i is a Polish (type) space and $g_i : T_i \to$

[8]For an arbitrary Polish space Y, let $\Delta^{\mathcal{B}_i}(\Sigma \times Y)$ denote the collection of CPS's over $(\Sigma \times Y, \{B_i \times Y \mid B_i \in \mathcal{B}_i\})$.

$\Delta^{\mathcal{B}_i}(\Sigma \times T_j)$ a continuous function, with $g_i^{B_i}[t_i] \in \Delta(\Sigma \times T_j)$ denoting the conditional probability measure associated with $t_i \in T_i$ given the hypothesis $B_i \in \mathcal{B}_i$.

A type, $t_i \in T_i$, is associated with a complete description of the agent's state of mind, thus inducing a hierarchy of conditional beliefs. The first order conditional beliefs are given by the CPS $\pi_i^1(\cdot|t_i) := \big(\pi_i^1(\cdot|t_i, B_i); B_i \in \mathcal{B}_i\big) \in \Delta^{\mathcal{B}_i}(\Theta_i^0)$, where

$$\pi_i^1(\sigma|t_i, B_i) := \int_{t_j \in T_j} g_i^{B_i}[t_i](\sigma, t_j)\ dt_j$$

denotes the probability density that t_i assigns to an arbitrary $\sigma \in \Sigma$ given the hypothesis B_i.

For every $n > 0$, let the correspondence $\beta_j^n : \Delta^{\mathcal{B}_j}(\Theta_j^{n-1}) \twoheadrightarrow T_j \cup \{\emptyset\}$ map every j's n-th order belief hierarchy to a (possibly empty) set of types in T_j. That is, for an arbitrary $\pi_j^n \in \Delta^{\mathcal{B}_j}(\Theta_j^{n-1})$, let

$$\beta_j^n(\pi_j^n) := \{t_j \in T_j : \pi_j^n(\cdot|t_j) = \pi_j^n\}.$$

Then, i's $(n+1)$-th order beliefs are given by the CPS $\pi_i^{n+1}(\cdot|t_i) := \big(\pi_i^{n+1}(\cdot|t_i, B_i); B_i \in \mathcal{B}_i\big) \in \Delta^{\mathcal{B}_i}(\Theta_i^n)$, where

$$\pi_i^{n+1}(\sigma, \pi_j^1, ..., \pi_j^n|t_i, B_i) := \int_{t_j \in \beta_j^1(\pi_j^1) \cap \cdots \cap \beta_j^n(\pi_j^n)} g_i^{B_i}[t_i](\sigma, t_j)\ dt_j$$

denotes the probability density that t_i assigns to an arbitrary $(\sigma, \pi_j^1, ..., \pi_j^n) \in \Sigma \times \Delta^{\mathcal{B}_j}(\Theta_j^0) \times \cdots \times \Delta^{\mathcal{B}_j}(\Theta_j^{n-1})$ given the hypothesis $B_i \times \Delta^{\mathcal{B}_j}(\Theta_j^0) \times \cdots \times \Delta^{\mathcal{B}_j}(\Theta_j^{n-1})$.

4. Conditional belief hierarchies and prior beliefs

Definition 3. We say that t_i's beliefs about Σ are derived from the prior $q_i \in \Delta(\Sigma)$, whenever for an arbitrary $\sigma \in \Sigma$,

$$\pi_i^1(\sigma|t_i, B_i) = q_i(\sigma|B_i)$$

for all $B_i \in \mathcal{B}_i$ with $q_i(B_i) > 0$.

In order to say that it is commonly known that i's beliefs are derived from a prior q_i, we first need to formally define the notion of common knowledge, which corresponds to the intuitive idea that everybody knows that i's prior over Σ is q_i, everybody knows that everybody knows that i's prior over Σ is q_i, and so on. In order to do so, we construct the associated state space model of $(\Sigma, \mathcal{B}_a, \mathcal{B}_b, T_a, T_b, g_a, g_b)$. Let

$$\Omega := \Sigma \times T_a \times T_b$$

be the state space associated with the T-space above, with typical element $\omega = (\sigma, t_a, t_b)$. Let $t_i(\omega) := \text{marg}_{T_i}\, \omega$ and $\sigma(\omega) := \text{marg}_{\Sigma}\, \omega$. Moreover, let $[t_i] := \{\omega \in \Omega : t_i(\omega) = t_i\}$ and $[\sigma] :=$

$\{\omega \in \Omega : \sigma(\omega) = \sigma\}$. Endow every agent with an information partition $\mathcal{P}_i$, with $P_i(\omega)$ containing the states $\omega' \in \Omega$ that i cannot distinguish from ω. We define the information partition as usual (Brandenburger and Dekel, 1993), i.e., for each $\omega \in \Omega$ let

$$P_i(\omega) := \{\omega' \in \Omega : t_i(\omega') = t_i(\omega)\} = [t_i(\omega)].$$

That is, i cannot distinguish between states that correspond to the same T_i-type. As usual, we say that i knows an event $E \subseteq \Omega$ at ω, whenever $P_i(\omega) \subseteq E$. Let the partition $\mathcal{M} := \mathcal{P}_a \wedge \mathcal{P}_b$ denote the finest common coarsening of the information partitions, with $M(\omega)$ denoting the element of $\mathcal{M}$ that contains ω. Then, E is commonly known at ω, if and only if $M(\omega) \subseteq E$ (Aumann, 1976).

Let $[q_i] \subseteq \Omega$ be the set of states where i's beliefs about the underlying space of uncertainty are derived from q_i. Formally, we have

$$[q_i] := \{\omega \in \Omega \ : \ \pi_i^1(\cdot|t_i(\omega), B_i) = q_i(\cdot|B_i), \text{ for all } B_i \in \mathcal{B}_i \text{ with } q_i(B_i) > 0\}.$$

Then it follows from the previous discussion that it is common knowledge at ω that i's beliefs are derived from q_i, whenever $M(\omega) \subseteq [q_i]$.

Theorem 1. *Let $(\Sigma, \mathcal{B}_a, \mathcal{B}_b, T_a, T_b, g_a, g_b)$ be a type space model. If it is commonly known that the beliefs of every $i \in \{a, b\}$ are derived from a full-support prior $q_i \in \Delta(\Sigma)$, then every $t_i \in T_i$ yields the same hierarchy of conditional beliefs, i.e., if for every i there is some $q_i \in \Delta(\Sigma)$ such that $\Gamma(q_i) = \Sigma$ and $M(\omega) \subseteq [q_i]$, then there is some $\tilde{\pi}_i \in \times_{n=0}^{\infty} \Delta^{\mathcal{B}_i}(\Theta_i^n)$ such that $\pi_i(\cdot|t_i) = \tilde{\pi}_i$ for all $t_i \in T_i$.*

Proof. It follows by construction that $M(\omega) = \Omega$, implying that $[q_i] = \Omega$. Since q_i is full-support, it follows that $q_i(B_i) > 0$ for all $B_i \in \mathcal{B}_i$, and therefore $\pi_i^1(\cdot|t_i, B_i) = q_i(\cdot|B_i)$ for all $\sigma \in \Sigma$ and all $t_i \in T_i$. Hence, there is some $\tilde{\pi}_i^1 \in \Delta^{\mathcal{B}_i}(\Theta_i^0)$ such that $\pi_i^1(\cdot|t_i) = \tilde{\pi}_i^1$ for all $t_i \in T_i$, implying that $\beta_i^1(\tilde{\pi}_i^1) = T_i$ for each $i \in \{a, b\}$. The latter implies that for all $t_i \in T_i$, the second order beliefs are such that (a) if $\pi_j^1 \neq \tilde{\pi}_j^1$, then $\pi_i^2(\sigma, \pi_j^1|t_i, B_i) = 0$ for every $\sigma \in \Sigma$, and (b) if $\pi_j^1 = \tilde{\pi}_j^1$, then for an arbitrary $\sigma \in \Sigma$,

$$\begin{aligned}
\pi_i^2(\sigma, \tilde{\pi}_j^1|t_i, B_i) &= \int_{t_j \in \beta_j^1(\tilde{\pi}_j^1)} g_i^{B_i}[t_i](\sigma, t_j)\ dt_j \\
&= \int_{t_j \in T_j} g_i^{B_i}[t_i](\sigma, t_j)\ dt_j \\
&= \pi_i^1(\sigma|t_i, B_i) \\
&= q_i(\sigma|B_i).
\end{aligned}$$

Thus, there is some $\tilde{\pi}_i^2 \in \Delta^{\mathcal{B}_i}(\Theta_i^1)$ such that $\pi_i^2(\cdot|t_i) = \tilde{\pi}_i^2$ for all $t_i \in T_i$. Inductively, we obtain $\beta_i^n(\tilde{\pi}_i^n) = T_i$ for all $i \in \{a, b\}$, which implies that there is some $\tilde{\pi}_i^n \in \Delta^{\mathcal{B}_i}(\Theta_i^{n-1})$ such that $\pi_i^n(\cdot|t_i) = \tilde{\pi}_i^n$ for all $t_i \in T_i$, which completes the proof. $\square$

The previous result, though rather straightforward to prove, is a bit surprising, as it essentially states that common knowledge of i's prior leads to common knowledge of i's conditional belief hierarchy. Therefore, all interactive uncertainty about Σ is finally attributed to interactive uncertainty about the prior beliefs.

Notice that the previous result relies on q_i being full-support. This assumption is quite crucial as the following example illustrates.

Example 1. Let $\Sigma := \{\sigma_1, \sigma_2, \sigma_3\}$, and $\mathcal{B}_a := \{\Sigma, \{\sigma_1, \sigma_2\}\}$ and $\mathcal{B}_b := \{\Sigma\}$, and suppose that $T_a = \{t_a^1, t_a^2\}$ and $T_b = \{t_b\}$. Let both types of player a assigns probability 1 to (σ_3, t_b) given the conditioning event $\Sigma \in \mathcal{B}_a$, i.e., $g_a^{\Sigma}[t_a^k](\sigma_3, t_b) = 1$ for $k = 1, 2$. On the other hand, the two types differ in their beliefs given the hypothesis $\{\sigma_1, \sigma_2\}$, in that t_a^1 assigns probability 1 to (σ_1, t_b), whereas t_a^2 assigns probability 1 to (σ_2, t_b), i.e., $g_a^{\{\sigma_1,\sigma_2\}}[t_a^k](\sigma_k, t_b) = 1$ for $k = 1, 2$. Finally, $g_b^{\Sigma}[t_b]$ is uniformly distributed over $\Sigma \times T_a$. Observe that the only prior $q_a \in \Delta(\Sigma)$ such that both t_a^1's and t_a^2's beliefs are derived from q_a, is the one that assigns probability 1 to $\{\sigma_3\}$. Notice that this prior meets all the conditions stated in Theorem 1, except the full-support one. Therefore, the fact that the two types of a yield different first order conditional beliefs is attributed to the failure to satisfy the full-support condition, implying that the latter is rather crucial. $\triangleleft$

In fact, if we relax the requirement about q_i being full-support, all conditional belief hierarchies can be derived from some commonly known prior, as long as the collection of conditioning events does not cover the underlying space of uncertainty.

Theorem 2. *Let $(\Sigma, \mathcal{B}_a, \mathcal{B}_b, T_a, T_b, g_a, g_b)$ be a type space model, such that $\bigcup_{B_i \in \mathcal{B}_i} B_i \subsetneq \Sigma$. Then there are commonly known priors $q_i \in \Delta(\Sigma)$ such that the beliefs of every $i \in \{a, b\}$ are derived from q_i.*

Proof. The proof is straightforward if we consider q_i's such that $\Gamma(q_i) = \Sigma \setminus \left(\bigcup_{B_i \in \mathcal{B}_i} B_i\right)$. The latter is always possible as $\bigcup_{B_i \in \mathcal{B}_i} B_i$ is an open set, implying that the complement $\Sigma \setminus \left(\bigcup_{B_i \in \mathcal{B}_i} B_i\right)$ is closed. Then, obviously $q_i(B_i) = 0$ for all $B_i \in \mathcal{B}_i$ implying that the prior does not impose any restriction on the conditional beliefs given $\mathcal{B}_i$, and therefore i's beliefs can be derived from q_i. $\square$

References

Aliprantis, C. & Border, K. (1994). *Infinite dimensional analysis.* Springer Verlag, Berlin.

Aumann, R.J. (1976). Agreeing to disagree. *Annals of Statistics* 4, 1236-1239.

——— (1998). Common priors: A reply to Gul. *Econometrica* 66, 929–938.

Battigalli, P. & Friedenberg, A. (2010). Forward induction reasoning revisited. *Theoretical Economics*, forthcoming.

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

——— (2002). Strong belief and forward induction reasoning. *Journal of Economic Theory* 106, 356–391.

Brandenburger, A. & Dekel, E. (1993). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory* 59, 189–198.

Gul, F. (1998). A comment on Aumann's Bayesian view. *Econometrica* 66, 923–927.

Harsanyi, J. (1967-68). Games with incomplete information played by Bayesian players, I-III. *Management Science* 14, 159–182, 320–334, 486–502.

Kechris, A. (1995). *Classical descriptive set theory.* Springer Verlag, Berlin.

Mertens, J.F. & Zamir, S. (1985). Formulation of Bayesian analysis for games with incomplete information. *International Journal of Game Theory* 14, 1–29.

Myerson, R. (1986). Multistage games with communication. *Econometrica* 54, 323–358.

Parthasarathy, K.R. (1967). *Probability measures on metric spaces.* AMS Chelsea Publishing, Providence, Rhode Island.

Rênyi, A. (1955). On a new axiomatic theory of probability. *Acta Mathetica Academiae Scientiarum Hungaricae* 6, 285–335.

Srivastava, S.M. (1991). *A course on Borel sets.* Springer Verlag, Berlin.