Rational belief hierarchies[☆]

Elias Tsakas

Department of Economics, Maastricht University

Abstract

We consider agents who attach a rational probability to every Borel event. We call these Borel probability measures rational, and introduce the notion of a rational belief hierarchy, where the first order beliefs are described by a rational measure over the fundamental space of uncertainty, the second order beliefs are described by a rational measure over the product of the fundamental space of uncertainty and the opponent's first order rational beliefs, and so on. Then, we derive the corresponding rational type space model, thus providing a Bayesian representation of rational belief hierarchies. Our main result shows that this type-based representation has the counterintuitive property that some rational types are associated with non-rational beliefs over the product of the fundamental space of uncertainty and the opponent's types, thus implying that the agent may attach an irrational probability to some Borel event even if she has a rational belief hierarchy.

Keywords: Belief hierarchies, type spaces, rational numbers.

1. Introduction

A belief hierarchy is a description of an agent's beliefs about some fundamental space of uncertainty, her beliefs about everybody else's beliefs, and so on. During the past few decades, belief hierarchies have become an integral tool of modern economic theory, often used to analyze games with incomplete information (Harsanyi, 1967-68), as well as in order to provide epistemic characterizations for several solution concepts, such as rationalizability (Brandenburger and Dekel, 1987; Tan and Werlang, 1988), Nash equilibrium (Aumann and Brandenburger, 1995), and correlated equilibrium (Aumann, 1987), just to mention a few.[1]

Belief hierarchies are in general very complex objects, consisting of infinite sequences of probability measures. This makes them in principle very hard to handle and sometimes even to describe, especially

---

[1]For an overview of the epistemic game theory literature we refer to the textbook by Perea (2012) or the review article by Brandenburger (2008).

when it comes to high order beliefs. Having recognized this difficulty, Harsanyi (1967-68) proposed an indirect Bayesian representation of belief hierarchies, known as the type space model. Formally, Harsanyi's model consists of a set of *types* for each agent and a continuous mapping from each type to the corresponding *conditional beliefs* over the product of the fundamental space of uncertainty and the opponent's type space. This structure induces a belief hierarchy for every type, thus reducing the infinite-dimensional regression of beliefs to a single-dimensional type. Mertens and Zamir (1985) and Brandenburger and Dekel (1993) completed the analysis by showing the existence of the universal type space, which represents all belief hierarchies satisfying some standard coherency properties.

In this paper we restrict attention to probabilistic beliefs that can take only rational values, e.g., we have in mind agents who do not hold beliefs of the form "*tomorrow it will rain with probability $\sqrt{2}/2$*". Such beliefs are modeled by Borel probability measures that attach a rational number to every Borel event. Throughout the paper, we call these probability measures *rational*.

Assuming that agents form rational beliefs over some underlying space of uncertainty Θ does not necessarily restrict the language they use in order to describe their beliefs, i.e., we remain within Harsanyi's framework which models the agents' language with the Borel σ-algebra of events in $\Delta(\Theta)$. This implies that infinite conjunctions/disjunctions are expressible, thus inducing a richer language than the ones typically used in logic.[2] As a consequence, our agents understand what it means to assign probability $\sqrt{2}/2$ to a Borel event $E \subseteq \Theta$, as the latter corresponds to the event $\{\mu \in \Delta(\Theta) : \mu(E) = \sqrt{2}/2\}$ which is Borel in $\Delta(\Theta)$. If on the other hand, the language used was finitely generated, similarly to the aforementioned models of logic, our agents would not be able to even understand what "*E occurs with probability $\sqrt{2}/2$*" meant, as such a sentence would not be expressible in the first place.[3] We further discuss the case of a finitely generated language later in the paper.

Even though our agents *can* express subjective beliefs that use non-rational probabilities, they refrain from actually doing so, as these beliefs are very complex. The idea is that agents are sophisticated enough to be able to understand every aspect of the environment, including what it means to form any kind of complex beliefs, but the latter is costly, and therefore they prefer to reason in simpler ways.[4] In fact, recent experimental findings on the understanding of non-rational numbers by students and mathematics teachers indicate that subjects find non-rational numbers very complex (Fischbein et al., 1995; Sirotic and Zazkis,

---

[2]The standard syntactic models of logic typically assume that the language that describes the agents' beliefs is finitely generated by sentences of the form "*E occurs with probability at least p*" where *p* is a rational number (Fagin and Halpern, 1994; Aumann, 1999). The latter induces an algebra of events in $\Delta(\Theta)$, which is obviously coarser than the Borel σ-algebra. Within this framework, Heifetz and Mongin (2001) provided a sound and complete axiomatization of Harsanyi's type-based models, while Zhou (2010) extended their analysis to the case of finitely additive type spaces.

[3]Restricting an agent's language resembles the structure typically considered in models of unawareness (Modica and Rustichini, 1999; Halpern, 2001; Heifetz et al., 2006). More specifically, in these models an agent is aware of a sentence if and only if she can express this sentence within the bounds of her language. Therefore, assuming that the agent's language cannot express non-rational probabilistic assessments is informally equivalent to the agent being unaware of the notion of non-rational numbers.

[4]For instance, in a different framework, Eliaz (2003), Spiegler (2004) and Maenner (2008) study repeated games with players who prefer to form simple beliefs.

2007a,b). For instance, it is found that even though subjects know the concept and the structure of non-rational numbers, they still have a tendency to rely on decimal approximations which they find more intuitive. Thus, we find it plausible to assume that agents may use only rational numbers to express their subjective beliefs.

Supposing, as usual, that an agent thinks that everybody else reasons the same way as she does, it is not only her first order beliefs that are restricted but also her beliefs about everybody else's beliefs, and so on. For instance, besides Alexandra's beliefs not assigning probability $\sqrt{2}/2$ to E, she also does not put positive probability to Barney believing E with probability $\sqrt{5}/5$. In other words, her belief hierarchy is restricted to consist of a sequence of rational probability measures, where the first order rational beliefs are described by a rational measure over the underlying space of uncertainty, the second order rational beliefs are described by a rational measure over the product of the fundamental space of uncertainty and the opponent's space of first order rational beliefs, and so on. We call this infinite regression of probability measures a *rational belief hierarchy*.

Following Mertens and Zamir (1985) and Brandenburger and Dekel (1993), we construct a Harsanyi type space representation of rational belief hierarchies. However, as our main result (Theorem 1) shows, this Bayesian representation has an odd property. Namely, it contains rational types which are represented by *non*-rational probability measures over the product of the fundamental space of uncertainty and the opponent's rational type space. In other words, there is some Borel event in this product space to which this rational type attaches a non-rational probability even though every order of her belief hierarchy involves only rational probabilities. We find this result quite surprising, both from a technical as well as conceptual point of view.

The technical implication is rather straightforward. Namely, it says that, contrary to our intuition, rational belief hierarchies are not necessarily induced by types that are associated with a rational probability measure over the product of the underlying space of uncertainty and the opponents' types.

Regarding the conceptual contribution on the other hand, note that the only Borel events that receive a non-rational probability by a rational type correspond to sentences that describe the opponent's entire belief hierarchy, i.e., these events contain elements of the form "$\theta \in \Theta$ *occurs*", and "*the opponent's first order beliefs are* π_1", and "*the opponent's second order beliefs are* π_2", and so on. Thus, the previous result is relevant only for events that are expressible when the agent's language is modeled by the Borel σ-algebra, as it is the case in Harsanyi's model, but not when the agent has a finitely generated language like for instance in logic. But then, the natural question is whether we should actually care about this type of events. In other words, should we consider agents with a finitely generated language and finitely additive beliefs whose reasoning is exhausted with the formulation of their belief hierarchy, or should we also let agents have a countably generated language and countably additive beliefs who also form beliefs about the opponents' entire belief hierarchy? The answer to this question is far from being straightforward and has attracted the attention of several prominent researchers. For instance, while Savage (1972) postulates that finitely additive subjective beliefs should be used, Harsanyi (1967-68) allows for countably additivity. The aim of this paper is not to provide a general answer to this question, but rather to point out that in the

existence of a countably generated language agents may not be able to form subjective beliefs about every event in their language by using only rational numbers. Later in the paper, we also discuss our main result in the context of a finitely generated language, like the ones typically considered in logic.

The paper is structured as follows: In Section 2 we formally introduce the notion of rational probability measures and we prove some of their properties; Section 3 extends this framework to an interactive setting by introducing rational belief hierarchies; In Section 4 we construct a terminal rational type space model and prove our main result; Section 5 contains a concluding discussion.

2. Rational probability measures

We begin with some definitions and the basic notation. Let X be a Polish space, together with the Borel σ-algebra, $\mathcal{B}$.[5] As usual, $\Delta(X)$ denotes the space of probability measures on $(X, \mathcal{B})$, endowed with the topology of weak convergence.[6] For each $\mu \in \Delta(X)$, let $\mathrm{supp}(\mu)$ denote the support, i.e., the smallest closed subset of X that receives probability 1 by μ.[7]

Consider the Borel probability measures that assign to every Borel event a rational number.

Definition 1. We define the set of *rational probability measures* by

$$\Delta^{\mathbb{Q}}(X) := \left\{ \mu \in \Delta(X) \; : \; \mu(B) \in \mathbb{Q}, \forall B \in \mathcal{B} \right\}. \tag{1}$$

We use rational probability measures to model an agent who does not hold beliefs of the form "*E occurs with probability* $\sqrt{2}/2$". Before moving forward, recall that the language that describes the agent's beliefs is a set of sentences each of which semantically corresponds to a subset of $\Delta(X)$, i.e., the language contains exactly those events with the property that the agent understands their meaning. Notice that a language is typically closed under (finitely or countably many) logical operators. In our case, the agent's language contains all events in the Borel σ-algebra generated by the topology of weak convergence. Hence, the agent *does* understand what it means to put probability $\sqrt{2}/2$ to E, as the latter is countably generated by events of the form $\{\mu \in \Delta(X) : \mu(E) \geq p\}$. Obviously, this language is richer than the one used in logic, where only finitely generated sentences are expressible, and therefore the agent does not even understand the sentence "*E occurs with probability* $\sqrt{2}/2$". In either case, we assume that the agent never uses such beliefs, as they are too complex.

Below, we provide some results on rational probability measures, which we will use later in the paper. Throughout this section, unless stated otherwise, we assume that X is separable and metrizable.

---

[5]A topological space is called Polish whenever it is separable and completely metrizable. Examples of Polish spaces include countable sets endowed with the discrete topology and $\mathbb{R}^n$ together with the usual topology. Closed subsets of Polish spaces endowed with the relative topology are Polish. The countable product of Polish spaces, together with the product topology, is also Polish.

[6]The topology of weak convergence, which is usually denoted by w^*, is the coarsest topology that makes the mapping $\mu \mapsto \int f d\mu$ continuous, for every bounded and continuous real-valued function, f. If X is Polish, then $\Delta(X)$ endowed with the topology of weak convergence is also Polish. For further properties of w^*, we refer to Aliprantis and Border (1994, Ch. 15).

[7]If X is separable and metrizable, the support is unique (Parthasarathy, 1967, Thm. 2.1).

Proposition 1. *Every $\mu \in \Delta^{\mathbb{Q}}(X)$ has a finite support.*

The previous, quite surprising result rules out all probability measures with countably infinite support, even if each singleton in the support receives a rational probability. The following example illustrates such a case.

Example 1. Let $X = \{x_1, x_2, \dots\}$ and suppose that $\mu \in \Delta(X)$ assigns probability 2^{-k} to each x_k. It is straightforward verifying that μ is a probability measure. Now, consider an arbitrary $\xi \in (0, 1)$, and construct the Borel subset $B_\xi \subseteq X$ so that $x_k \in B_\xi$ if and only if $\xi \in [\frac{1}{2^k}, \frac{2}{2^k}) \cup \cdots \cup [\frac{2^k-1}{2^k}, 1)$. Observe that $\mu(B_\xi) = \xi$, implying that for every non-rational ξ there is a Borel event receiving the non-rational probability ξ, and therefore μ is not a rational measure. ◄

The following result proves that rational probability measures form a Borel subset in the space of Borel probability measures. Thus, $\Delta^{\mathbb{Q}}(X)$ can be expressed as an event within the language used by the agents.

Proposition 2. $\Delta^{\mathbb{Q}}(X)$ *is a Borel subset of $\Delta(X)$.*

From a technical point of view, the previous result allows us to relate the framework used in this paper with the existing literature (see Proposition 3 in the next section).

3. Rational belief hierarchies

Let Θ be a Polish space together with the Borel σ-algebra, $\mathcal{B}_0$. For instance, in a game, each $\theta \in \Theta$ corresponds to a payoff vector (Harsanyi, 1967-68), or a strategy profile (Aumann and Brandenburger, 1995; Tan and Werlang, 1988), or a combination of the two. Throughout the paper, we refer to Θ as the underlying – else called fundamental – space of uncertainty. Let $I = \{a, b\}$ be the set of agents, with typical elements i and j.[8] Each agent forms beliefs about Θ (first order beliefs), beliefs about Θ and the opponent's first order beliefs (second order beliefs), and so on. Such a sequence is called a belief hierarchy.

Formally, consider the following sequence of Polish spaces:

$$
\begin{aligned}
\Psi_0 &:= \Theta \\
\Psi_1 &:= \Psi_0 \times \Delta(\Psi_0) \\
&\vdots \\
\Psi_{k+1} &:= \Psi_k \times \Delta(\Psi_k) \\
&\vdots
\end{aligned}
$$

A belief hierarchy is an element of

$$
T_0 := \prod_{k=0}^{\infty} \Delta(\Psi_k). \tag{2}
$$

---

[8] Our analysis can be directly generalized to any finite set of agents, in which case we obviously allow for correlated beliefs, as usual.

For some $(\pi_1, \pi_2, \dots) \in T_0$, the Borel probability measure $\pi_k \in \Delta(\Psi_{k-1})$ denotes the k-th order beliefs.

In this paper, we consider agents who form only rational beliefs. That is, for some $\theta \in \Theta$, agent i never believes that θ *occurs with probability* $\sqrt{2}/2$. The latter implies that we restrict i's first order beliefs to the space of rational probability measures. Furthermore, i is certain that j's does not put positive probability to the event that *"j assigns probability* $\sqrt{3}/3$ *to θ"*. That is, we also restrict i's second order beliefs to rational probability measures over the space of rational probability measures.[9] Likewise, we restrict higher order beliefs.

Formally, consider the sequence

$$
\begin{aligned}
\Theta_0 &:= \Theta \\
\Theta_1 &:= \Theta_0 \times \Delta^{\mathbb{Q}}(\Theta_0) \\
&\quad\vdots \\
\Theta_{k+1} &:= \Theta_k \times \Delta^{\mathbb{Q}}(\Theta_k) \\
&\quad\vdots
\end{aligned}
$$

A *rational belief hierarchy* is a sequence $(\pi_1, \pi_2, \dots)$, with $\pi_k \in \Delta^{\mathbb{Q}}(\Theta_{k-1})$ denoting the k-th order beliefs. Let

$$
T_0^Q := \prod_{k=0}^{\infty} \Delta^{\mathbb{Q}}(\Theta_k) \tag{3}
$$

denote the space of all rational belief hierarchies, endowed with the product topology.

Intuitively, rational belief hierarchies form a strict subset of all belief hierarchies. However, observe that formally T_0^Q is not a subset of T_0, because strictly speaking $\Delta^{\mathbb{Q}}(\Theta_k)$ is not a subset of $\Delta(\Psi_k)$. Therefore, before moving forward, we would like to make sure that the intuitive idea of one being a subset of the other is compatible with our formal model. The following result serves this purpose, by showing that T_0^Q is embedded as a Borel subset of T_0.

Proposition 3. T_0^Q *is homeomorphic to a Borel subset of T_0.*

Throughout the paper, we denote this embedding by

$$
h : T_0^Q \hookrightarrow T_0. \tag{4}
$$

As usual, with slight abuse of terminology, whenever we talk about a rational belief hierarchy $(\pi_1, \pi_2, \dots) \in T_0^Q$ we actually refer to its image $h(\pi_1, \pi_2, \dots) \in T_0$, and therefore we will informally consider T_0^Q to be a Borel subset of T_0.

---

[9]As we have already mentioned, if we assumed instead that i's language was finitely generated similarly to what is done in logic, we could also informally interpret our restriction of rational beliefs as the beliefs of an agent who is unaware of the concept of non-rational numbers. In such a case, i would not even understand the meaning of the sentence "*j assigns probability* $\sqrt{3}/3$ *to E*", and therefore i would not be able to put positive probability to it.

4. Rational types

In general, belief hierarchies are very large and complex objects, and as such it is really hard directly working with them. Harsanyi (1967-68) was the first one to circumvent this problem by proposing a compact way of expressing belief hierarchies, known in the literature as the *type space model*. Formally, this model consists of a tuple $(\Theta, T_a, T_b, g_a, g_b)$, where T_i is a Polish space of types with typical element t_i, and $g_i : T_i \to \Delta(\Theta \times T_j)$ is a continuous function. In a type space, each $t_i \in T_i$ is associated with a unique *belief hierarchy* derived as follows: For each type $t_i \in T_i$, the *first order beliefs*, $\pi_1(t_i) \in \Delta(\Psi_0)$, attach probability

$$\pi_1(t_i)(B_0) = \int_{(\theta, t_j): \theta \in B_0} dg_i(t_i) \tag{5}$$

to every Borel event $B_0 \subseteq \Psi_0$. The *k-th order beliefs*, $\pi_k(t_i) \in \Delta(\Psi_{k-1})$, attach probability

$$\pi_k(t_i)(B_{k-1}) = \int_{(\theta, t_j):(\theta, \pi_1(t_j), \ldots, \pi_{k-1}(t_j)) \in B_{k-1}} dg_i(t_i) \tag{6}$$

for every Borel subset $B_{k-1} \subseteq \Psi_{k-1}$. For a detailed presentation on how the entire belief hierarchy is derived from a type space model, we refer to Siniscalchi (2007).

We say that a type space model induces a *Bayesian representation* of the belief hierarchies that are associated with each type if g_i is injective, i.e., each $t_i \in T_i$ is mapped to a different probability measure $g_i(t_i)$, and therefore can be identified by a conditional belief over $\Theta \times T_j$. We say that the Bayesian representation of T_i is *terminal* whenever for every belief hierarchy there exists a type inducing it. Finally, we call a Bayesian representation *complete*, if g_i is surjective, implying that every measure in $\Delta(\Theta \times T_j)$ is the image of some type in T_i. Mertens and Zamir (1985), and Brandenburger and Dekel (1993) showed that Harsanyi's framework is sufficiently rich to model all instances of interactive uncertainty, in that there is a type space model $(\Theta, T_a^*, T_b^*, g_a^*, g_b^*)$, with $T_a^* = T_b^* = T^*$ and $g_a^* = g_b^* = g^*$, which is both complete and terminal. This construction is called the *universal* type space model.[10]

Brandenburger and Dekel (1993) started by imposing a standard coherency condition, which states that the k-th and $(k + 1)$-th order beliefs cannot contradict each other. Formally, let $T_c := \{(\pi_1, \pi_2, \ldots) \in T_0 : \mathrm{marg}_{\Psi_{k-2}} \pi_k = \pi_{k-1}, \forall k > 1\}$. Then, they showed that there is a homeomorphism

$$f^* : T_c \to \Delta(\Theta \times T_0). \tag{7}$$

This homeomorphism is a natural one, in that for all $(\pi_1, \pi_2, \ldots) \in T_c$,

$$\mathrm{marg}_{\Psi_{k-1}} f^*(\pi_1, \pi_2, \ldots) = \pi_k. \tag{8}$$

Then, they further restricted attention to belief hierarchies that satisfy, not only coherency, but also common

---

[10]Heifetz (1993) generalized this representation result to cases where the underlying space of uncertainty is Hausdorff, while Heifetz and Samet (1998) and Pinter (2005) considered a purely measurable underlying space of uncertainty.

certainty in coherency. Formally, consider the following sequence of subsets of T_c:

$$
\begin{aligned}
T_1 &:= T_c \\
T_2 &:= \{\, t \in T_c : f^*(t)(\Theta \times T_1) = 1 \,\} \\
&\ \ \vdots \\
T_k &:= \{\, t \in T_c : f^*(t)(\Theta \times T_{k-1}) = 1 \,\} \\
&\ \ \vdots
\end{aligned}
$$

Observe that T_1 contains the belief hierarchies satisfying coherency, T_2 those satisfying certainty in everybody's coherency, and so on. Thus,

$$
T^* := \bigcap_{k=1}^{\infty} T_k \tag{9}
$$

contains the belief hierarchies that satisfy both coherency and common certainty in coherency. Finally, Brandenburger and Dekel (1993, Prop. 2) showed that there is a homeomorphism

$$
g^* : T^* \to \Delta(\Theta \times T^*), \tag{10}
$$

implying that there is a universal Bayesian representation of T^*.

The first natural question arising at this point is whether we can extend their result to the case of rational belief hierarchies. In other words, is there a universal type space representation of rational belief hierarchies in the same line as the standard results by Mertens and Zamir (1985) and Brandenburger and Dekel (1993)?

We retain the standard coherency restriction. Formally, let

$$
\begin{aligned}
T_c^Q &:= \{\, (\pi_1, \pi_2, \dots) \in T_0^Q : \operatorname{marg}_{\Theta_{k-2}} \pi_k = \pi_{k-1}, \ \forall k > 1 \,\} \\
&= T_0^Q \cap T_c
\end{aligned} \tag{11}
$$

denote the set of coherent belief hierarchies.

Similarly to Brandenburger and Dekel (1993, Prop. 1), the following result associates each coherent rational belief hierarchy to a probability measure over the product of the underlying space of uncertainty and the space of the opponent's rational hierarchies. This induces an injective mapping, implying that there is no pair of coherent types associated with the same distribution over $\Theta \times T_0^Q$.

Proposition 4. *There is an injection $f : T_c^Q \to \Delta(\Theta \times T_0^Q)$.*

It is rather easy to see that the function f is in fact the same as f^* from Brandenburger and Dekel (1993, Prop. 1), but restricted to rational belief hierarchies, i.e., for every $t \in T_c^Q$ and each Borel subset $B \subseteq \Theta \times T_0^Q$,

$$
f(t)(B) = f^*(h(t))(\{(\theta, t) \in \Theta \times T_0 : (\theta, h^{-1}(t)) \in B\}). \tag{12}
$$

Throughout the paper, we treat f and f^* as the same function. Note that f inherits from f^* the property of being a natural mapping, in that every coherent hierarchy is associated with a probability measure over

$\Theta \times \prod_{k=0}^{\infty} \Delta^{\mathbb{Q}}(\Theta_k)$ that has the property that its marginal distribution over Θ_{k-1} coincides with the k-th order beliefs induced by this hierarchy, i.e., for every $(\pi_1, \pi_2, \dots) \in T_c^Q$

$$\mathrm{marg}_{\Theta_{k-1}} f(\pi_1, \pi_2, \dots) = \pi_k. \tag{13}$$

As usual, we further restrict belief hierarchies so that they satisfy, not only coherency, but also common certainty in coherency. Formally, consider the following sequence of subsets of T_c^Q:

$$\begin{aligned} T_1^Q &:= T_c^Q \\ T_2^Q &:= \{ t \in T_c^Q : f(t)(\Theta \times T_1^Q) = 1 \} \\ &\vdots \\ T_k^Q &:= \{ t \in T_c^Q : f(t)(\Theta \times T_{k-1}^Q) = 1 \} \\ &\vdots \end{aligned}$$

Note that T_1^Q contains the belief hierarchies that satisfy coherency, T_2^Q contains the belief hierarchies that satisfy certainty in everybody coherency, and so on. Thus, the types in

$$T^Q := \bigcap_{k=1}^{\infty} T_k^Q \tag{14}$$

satisfy coherency and common certainty in coherency. Henceforth, whenever we write "rational belief hierarchies" or "rational types", we implicitly refer to elements of T^Q, thus omitting to explicitly say that they satisfy coherency and common certainty in coherency. The following result proves the existence of a *terminal Bayesian representation of rational belief hierarchies*, implying that every rational belief hierarchy is identified by a Borel probability measure on $\Theta \times T^Q$.

Proposition 5. *There is an injection $g : T^Q \to \Delta(\Theta \times T^Q)$.*

Once again, g coincides with the corresponding mapping g^* used by Brandenburger and Dekel (1993, Prop. 2) when restricted on the domain $h(T^Q)$. That is, for every $t \in T^Q$ and each Borel subset $B \subseteq \Theta \times T^Q$,

$$g(t)(B) = g^*(h(t))(\{(\theta, t) \in \Theta \times T^Q : (\theta, h^{-1}(t)) \in B\}). \tag{15}$$

Obviously, the representation induced by g is not complete, in that there are measures in $\Delta(\Theta \times T^Q)$ which are not the image of any rational type. The latter is not surprising, as one can easily see that there exist probability measures $\pi \in \Delta(\Theta \times T^Q)$ with $\mathrm{marg}_{\Theta} \pi \notin \Delta^{\mathbb{Q}}(\Theta)$, e.g., a measure with $\pi(\{\theta\} \times T^Q) = \sqrt{2}/2$.

What is really interesting, as well as far from obvious, is the conclusion of the following theorem. Accordingly, it is shown that there exist rational belief hierarchies that are mapped via g to non-rational probability measures over $\Theta \times T^Q$.

Theorem 1. *There is some $t \in T^Q$ such that $g(t) \notin \Delta^{\mathbb{Q}}(\Theta \times T^Q)$.*

The technical implication of this result is straightforward. Namely, unlike what one would expect, we show that in order to represent some rational belief hierarchies we need to construct a Harsanyi type space model that associates these hierarchies with a non-rational probability measure over the product of the fundamental space of uncertainty and the opponent's types.

The conceptual contribution, on the other hand, relies on the observation that there exists a rational type $t \in T^Q$ associated with an infinite-support probability measure over $\Theta \times T^Q$ (see the proof of Theorem 1 in Appendix C). Then, it follows from Proposition 1 that there exists some Borel event in $\Theta \times T^Q$ that receives a non-rational probability by $g(t)$. That is, such an event describes the underlying space of uncertainty and the opponent's entire belief hierarchy, as it contains elements of the form $(\theta, \pi_1, \pi_2, \dots) \in \Theta \times \Delta^{\mathbb{Q}}(\Theta_0) \times \Delta^{\mathbb{Q}}(\Theta_1) \times \cdots$. The natural question that arises at this point is whether we should be interested in an agent's beliefs about the opponent's entire belief hierarchy or just in the agent's own belief hierarchy. In other words, does the agent reason about the opponent's type, or is her reasoning exhausted by merely forming her own belief hierarchy like in the models of logic where events in $\Theta \times T^Q$ are not expressible in the first place? The answer to this question is not straightforward at all, and goes deep inside philosophical discussions on the nature of subjective beliefs, and whether these should be finitely or countably additive.[11] Though the aim of this paper is not to provide a general answer to this question, our result contributes to this debate by showing that if the language is modeled by the Borel σ-algebra in $\Theta \times T^Q$ – like in Harsanyi's framework – agents may not be able to express all their subjective beliefs by using only rational numbers. In the next section, we discuss whether, in the light of our previous conclusion, it would be more appropriate to study rational beliefs in the context of a finitely generated language.

5. Discussion

5.1. Finitely additive rational beliefs

As we have already mentioned in the previous section, we model rational beliefs by Borel probability measures, implying that we consider agents with a countably generated language described by the Borel σ-algebra. In other words, $\Delta(\Theta)$ is endowed with the σ-algebra that is (countably) generated by events of the form $\{\mu \in \Delta(\Theta) : \mu(E) \geq p\}$ where $E \subseteq \Theta$ is Borel and $p \in [0, 1]$ is rational. This assumption is rather standard within game theory, (explicitly or implicitly) imposed by several authors (e.g., Harsanyi, 1967-68; Brandenburger and Dekel, 1993).

On the other hand, the standard syntactic models of logic typically consider a finitely generated language (Fagin and Halpern, 1994; Aumann, 1999). That is, agents can only express sentences that are finitely generated by events of the form $\{\mu \in \Delta(\Theta) : \mu(E) \geq p\}$, thus implicitly postulating that beliefs are finitely additive, in accordance with Savage's idea that subjective beliefs should be modeled by finitely additive probability measures. In this framework, Heifetz and Mongin (2001) provided an axiomatic foundation for Harsanyi's type spaces. Their setting also allows for types that assign a non-rational probability to a

---

[11]In fact, while Savage (1972) postulates that subjective beliefs should be finitely additive, Harsanyi (1967-68) considers countable additive beliefs.

measurable event in the opponent's type space (though they do not explicitly mention this). In this respect, the technical implication of their result is similar to ours, but in their case such measurable events are not expressible in the first place, and therefore there is no conceptual controversy. More recently, Meier (2006) constructed a universal type space for finitely additive belief hierarchies, and Zhou (2010) extended the Heifetz-Mongin axiomatization to finitely additive type spaces.

One natural question at this point is whether we could also restrict focus to finitely additive rational beliefs, thus embedding our construction to the universal type space of Meier (2006) instead of viewing it as a subspace of the Brandenburger-Dekel universal type space. At first, this might seem straightforward. However, as it turns out, there are several differences between our model and such an alternative construction. For instance, if we focused on finitely additive beliefs, Proposition 1 would not hold any more, implying that we would allow for rational probability measures with an infinite support. To see this, consider a countable fundamental space of uncertainty $\Theta = \{\theta_1, \theta_2, \dots\}$ together with the algebra generated by all finite subsets, and the probability measure introduced in Example 1 which assigns to each singleton θ_k probability 2^{-k}. Now, observe that this probability measure would still be rational, as it would assign a rational number to every event in the algebra that describes the agent's language. In this respect, our result also differs from the implication of the main result in Heifetz and Mongin (2001) that we mentioned above, in that they allow for a richer set of beliefs than we do, even though their language is coarser than ours.

5.2. *Other related literature*

The notion of uncertainty which is modeled by rational probability measures is not new in the literature. For instance, several authors have studied the rational correlated equilibrium[12] and the conditions under which it can be generated though a mediator (e.g., Lehrer, 1996; Krishna, 2007). Other authors have studied utility theory with lotteries that put a rational probability to every outcome (Shepherdson, 1980; Hu, 2009).

The obvious conceptual difference between our work and the aforementioned papers obviously lies on the fact that in our model we consider subjective beliefs, contrary to the objective beliefs that are (only implicitly) present in these papers. Such a difference becomes very significant when we study higher order beliefs, as it would be very difficult to motivate the restriction to rational probabilities in such a case. This is due to the fact that agents cannot observe another agent's state of mind, and consequently they cannot construct an objective lottery representing the agent's uncertainty about the opponent's first order beliefs.

Appendix A. Proofs of Section 2

Proof of Proposition 1. Consider an arbitrary $\mu \in \Delta^{\mathbb{Q}}(X)$, and consider the set of singletons with positive measure,

$$\Gamma := \{x \in X : \mu(\{x\}) > 0\}. \tag{A.1}$$

First, we show that Γ is non-empty. Suppose that μ is a non-atomic measure. Then, it follows from Fremlin (2003, p. 46) that for every $\xi \in (0, 1)$ there is some $B \in \mathcal{B}$ such that $\mu(B) = \xi$, which contradicts

---

[12]A rational correlated equilibrium in a normal form game is one that assigns a rational probability to every strategy profile (Lehrer, 1996).

$\mu \in \Delta^{\mathbb{Q}}(X)$ if we consider some $\xi \in \mathbb{R} \setminus \mathbb{Q}$. Hence, there is at least one atom $A \in \mathcal{B}$. Now, it follows from Aliprantis and Border (1994, Lem. 12.18) that A contains a singleton of positive measure, implying that Γ is non-empty.

Second, we show that Γ is countable. Let $\{\Gamma_n \ ; \ n \geq 1\}$ be the countable partition of Γ, defined by

$$\Gamma_n := \left\{ x \in \Gamma : \frac{1}{n+1} < \mu(\{x\}) \leq \frac{1}{n} \right\}.$$

If Γ is uncountable, there is some $n \geq 1$ such that Γ_n is uncountable, implying that there is a countably infinite $\{x_1, x_2, \dots\} \subseteq \Gamma_n$. Finally, observe that

$$
\begin{aligned}
\mu(X) \ &\geq \ \mu(\Gamma_n) \\
&\geq \ \sum_{k=1}^{\infty} \mu(\{x_k\}) \\
&> \ \sum_{k=1}^{\infty} \frac{1}{n+1} \\
&= \ \infty,
\end{aligned}
$$

which is a contradiction.

Third, we show that $\mu(\Gamma) = 1$. Assume otherwise, implying that $\mu(X \setminus \Gamma) > 0$. Since Γ is countable, it is Borel, implying that $X \setminus \Gamma$ is also Borel. Moreover, recall from the first step of the proof that μ is an atomic measure. Hence, it follows from Aliprantis and Border (1994, Lem. 12.18) that there is $x \in X \setminus \Gamma$ with $\mu(\{x\}) > 0$, implying, by Eq. (A.1), that $x \in \Gamma$, which is a contradiction.

Now, suppose that $\Gamma = \{x_1, x_2, \dots\}$ is infinite. Observe that the sequence of rational numbers $\{\mu(\{x_k\})\}_{k>0}$ satisfies $\sum_{k=1}^{\infty} \mu(\{x_k\}) = 1$. Then, it follows from Badea (1987, Prop., p. 225) that there is a subsequence $\{y_k\}_{k>0}$ of $\{x_k\}_{k>0}$ such that $\sum_{k=1}^{\infty} \mu(\{y_k\})$ is an irrational number, thus contradicting the hypothesis that $\mu \in \Delta^{\mathbb{Q}}(X)$.

Therefore, Γ is necessarily finite. Moreover, it is closed, as it is the finite union of singletons, implying that $\text{supp}(\mu) = \Gamma$, which completes the proof. $\qquad\square$

Proof of Proposition 2. For some finite $N \subseteq \mathbb{N}$, consider the subset of the rational numbers $\mathbb{Q}_N := \{m/n : m = 0, \dots, n \ ; \ n \in N\}$ that can be written as a fraction with the denominator belonging to N. Then, define the set of N-rational probability measures by

$$\Delta^N(X) := \left\{ \mu \in \Delta(X) \ : \ \mu(B) \in \mathbb{Q}_N, \forall B \in \mathcal{B} \right\}. \tag{A.2}$$

Tsakas (2012) showed that $\Delta^N(X)$ is closed in $\Delta(X)$. However, for the sake of completeness, we repeat the proof here. It suffices to show that an arbitrary convergent sequence $\{\mu_k\}$ of elements of $\Delta^N(X)$ has its limit in $\Delta^N(X)$, i.e., if $\mu_k \xrightarrow{w^*} \mu$, then $\mu \in \Delta^N(X)$. Let $\bar{n} := \max_{n \in N} n$. Let also $d : X \times X \to \mathbb{R}$ be a metric compatible with the topology on X, and for every $x \in X$ and $\delta > 0$, define an open neighborhood of x as $B(x, \delta) := \{x' \in X : d(x, x') < \delta\}$. Consider an arbitrary $x \in X$, and suppose there is some $\delta > 0$ such that there are finitely many $k > 0$ with $\mu_k(B(x, \delta)) > 0$. Then, obviously, there are infinitely many $k > 0$ such that $\mu_k(B(x, \delta)) = 0$,

implying that $\liminf \mu_k(B(x,\delta)) = 0$. Hence, it follows from $\mu_k \xrightarrow{w^*} \mu$ that $\mu(B(x,\delta)) \leq \liminf \mu_k(B(x,\delta)) = 0$ (Aliprantis and Border, 1994, Thm. 15.3), implying that $x \notin \text{supp}(\mu)$. If, on the other hand, for every $\delta > 0$ there are infinitely many $k > 0$ such that $\mu_k(B(x,\delta)) > 0$, it follows from $\mu_k \in \Delta^N(X)$ that there are infinitely many $k > 0$ such that $\mu_k(\overline{B(x,\delta)}) \geq 1/\bar{n}$, where $\overline{B(x,\delta)} := \{x' \in X : d(x,x') \leq \delta\}$ is the closure of $B(x,\delta)$. Therefore, $\mu(\overline{B(x,\delta)}) \geq \limsup \mu_k(\overline{B(x,\delta)}) \geq 1/\bar{n}$ (Aliprantis and Border, 1994, Thm. 15.3). Now, consider a sequence of positive reals $\{\delta_n\}$ with $\delta_n \downarrow 0$, which induces a sequence of Borel events $\{\overline{B(x,\delta_n)}\}$ such that $\limsup_{n>0} \overline{B(x,\delta_n)} = \{x\}$. Then, it follows from $\mu(\limsup_{n>0} \overline{B(x,\delta_n)}) \geq \limsup_{n>0} \mu(\overline{B(x,\delta_n)}) \geq 1/\bar{n}$ (Billingsley, 1995, Thm. 4.1) that $\mu(\{x\}) \geq 1/\bar{n}$. Hence, $x \in \text{supp}(\mu)$ if and only if $\mu(\{x\}) \geq 1/\bar{n}$, implying that $\text{supp}(\mu)$ is finite. Let $x \in \text{supp}(\mu)$. It follows from Aliprantis and Border (1994, Thm. 15.3) that for every $\delta > 0$,

$$
\begin{aligned}
\mu(\overline{B(x,\delta)}) &\geq \limsup \mu_k(\overline{B(x,\delta)}) \\
&\geq \limsup \mu_k(B(x,\delta)) \\
&\geq \liminf \mu_k(B(x,\delta)) \\
&\geq \mu(B(x,\delta)).
\end{aligned}
\tag{A.3}
$$

Since $\text{supp}(\mu)$ is finite, there is some $\rho > 0$ such that $x' \notin B(x,\rho)$ for any $x' \in \text{supp}(\mu) \setminus \{x\}$, implying that $\mu(\overline{B(x,\delta)}) = \mu(B(x,\delta)) = \mu(\{x\})$ for every $\delta < \rho$. Hence, it follows from (A.3) that $\mu(\{x\}) = \lim \mu_k(B(x,\delta))$. Finally, since the sequence $\{\mu_k(B(x,\delta))\}$ contains only elements of the finite set $\mathbb{Q}_N$, it follows that $\lim \mu_k(B(x,\delta)) \in \mathbb{Q}_N$.

Now, observe that $\Delta^{\mathbb{Q}}(X) = \bigcup_{n \in \mathbb{N}} \Delta^{\{1,\ldots,n\}}(X)$. It follows from the previous step that $\Delta^{\{1,\ldots,n\}}(X)$ is closed, and therefore Borel in $\Delta(X)$ for every $n \in \mathbb{N}$, which completes the proof. □

Appendix B. Proofs of Section 3

Proof of Proposition 3. We proceed inductively to show that for every $k \geq 0$

- Θ_k is embedded as a Borel subset of Ψ_k, and

- $\Delta^{\mathbb{Q}}(\Theta_k)$ is embedded as a Borel subset of $\Delta(\Psi_k)$.

First, observe that $\Theta_0 = \Psi_0 = \Theta$, implying that Θ_0 is embedded as a Borel subset of Ψ_0 via the identity function. It follows from Proposition 2 that $\Delta^{\mathbb{Q}}(\Theta_0)$ is a Borel subset of $\Delta(\Psi_0)$. Hence, $\Delta^{\mathbb{Q}}(\Theta_0)$ is embedded as a Borel subset of $\Delta(\Psi_0)$ via the identity function.

Now, suppose that Θ_k is embedded as a Borel subset of Ψ_k via $\vartheta_k : \Theta_k \to \Psi_k$, and $\Delta^{\mathbb{Q}}(\Theta_k)$ is embedded as a Borel subset of $\Delta(\Psi_k)$ via $\delta_k : \Delta^{\mathbb{Q}}(\Theta_k) \to \Delta(\Psi_k)$. Define the function $\vartheta_{k+1} : \Theta_{k+1} \to \Psi_{k+1}$ such that for each $(\theta_k, \mu_k) \in \Theta_k \times \Delta^{\mathbb{Q}}(\Theta_k)$, $\vartheta_{k+1}(\theta_k, \mu_k) := (\vartheta_k(\theta_k), \delta_k(\mu_k))$. Obviously, it follows from above that Θ_{k+1} is embedded as a Borel subset of Ψ_{k+1} via ϑ_{k+1}, and therefore Θ_{k+1} is homeomorphic to $\vartheta_{k+1}(\Theta_{k+1})$. Hence, $\Delta(\Theta_{k+1})$ is homeomorphic to $\Delta(\vartheta_{k+1}(\Theta_{k+1}))$. Since $\vartheta_{k+1}(\Theta_{k+1})$ is Borel in Ψ_{k+1}, there is a homeomorphism $\Delta(\vartheta_{k+1}(\Theta_{k+1})) \mapsto \{\mu \in \Delta(\Psi_{k+1}) : \mu(\vartheta_{k+1}(\Theta_{k+1})) = 1\}$, where $\{\mu \in \Delta(\Psi_{k+1}) : \mu(\vartheta_{k+1}(\Theta_{k+1})) = 1\}$ is a Borel

subset of $\Delta(\Psi_{k+1})$. The latter implies that there is a homeomorphism $\nu_{k+1} : \Delta(\Theta_{k+1}) \to \{\mu \in \Delta(\Psi_{k+1}) : \mu(\vartheta_{k+1}(\Theta_{k+1})) = 1\}$. By Proposition 2, $\Delta^{\mathbb{Q}}(\Theta_{k+1})$ is Borel in $\Delta(\Theta_{k+1})$. Let $\delta_{k+1} : \Delta^{\mathbb{Q}}(\Theta_{k+1}) \to \{\mu \in \Delta(\Psi_{k+1}) : \mu(\vartheta_{k+1}(\Theta_{k+1})) = 1\}$ be the same mapping as ν_{k+1} but restricted in the domain $\Delta^{\mathbb{Q}}(\Theta_{k+1})$. Then, it follows directly that δ_{k+1} embeds $\Delta^{\mathbb{Q}}(\Theta_{k+1})$ as a Borel subset on $\{\mu \in \Delta(\Psi_{k+1}) : \mu(\vartheta_{k+1}(\Theta_{k+1})) = 1\}$ and therefore on $\Delta(\Psi_{k+1})$, which completes the proof by induction. $\qquad\square$

Appendix C. Proofs of Section 4

Proof of Proposition 4. Since $\Theta_0 = \Theta$ is Polish, it is by definition separable and metrizable. Supposing that Θ_k is separable and metrizable, $\Delta^{\mathbb{Q}}(\Theta_k)$ is separable and metrizable too, as it is a subspace of $\Delta(\Theta_k)$, which is also separable and metrizable (Aliprantis and Border, 1994, Thm 15.12). Thus, $\Theta_{k+1} = \Theta_k \times \Delta^{\mathbb{Q}}(\Theta_k)$ is also separable and metrizable, and therefore it follows by induction that every Θ_k is separable and metrizable. Since $\pi_{k+1} \in \Delta^{\mathbb{Q}}(\Theta_k)$ has a finite support (by Proposition 1), it follows that it is tight, and therefore, by applying a version of Kolmogorov extension theorem (Aliprantis and Border, 1994, Cor. 15.28), we prove the existence of a unique measure $\pi \in \Delta(\Theta \times T_0^Q)$ that extends every π_k. $\qquad\square$

Proof of Proposition 5. The proof follows directly from $T^Q = \{t \in T_c^Q : f(t)(\Theta \times T^Q) = 1\}$. $\qquad\square$

Proof of Theorem 1. Take two arbitrary $\theta_1, \theta_2 \in \Theta$, and consider the following sequence:

$$P_1 := \{ p_1 \in \Delta^{\mathbb{Q}}(\Theta_0) : p_1(\theta) = 1, \text{ for some } \theta \in \{\theta_1, \theta_2\} \}$$
$$P_2 := \{ p_2 \in \Delta^{\mathbb{Q}}(\Theta_1) : p_2(\theta, p_1) = 1, \text{ for some } (\theta, p_1) \in \{\theta_1, \theta_2\} \times P_1 \}$$
$$\vdots$$
$$P_k := \{ p_k \in \Delta^{\mathbb{Q}}(\Theta_{k-1}) : p_k(\theta, p_1, \ldots, p_{k-1}) = 1, \text{ for some } (\theta, p_1, \ldots, p_{k-1}) \in \{\theta_1, \theta_2\} \times P_1 \times \cdots \times P_{k-1} \}$$
$$\vdots$$

Let T_p be the set of types $(p_1, p_2, \ldots) \in \prod_{k>0} P_k$ that satisfy coherency and common certainty in coherency. Observe that for every $p_1 \in P_1$ there are exactly two measures in P_2 such that (p_1, p_2) does not contradict coherency. Likewise, for every $(p_1, p_2) \in P_1 \times P_2$ that does not contradict coherency, there are exactly two measures $p_3 \in P_3$ such that (p_1, p_2, p_3) does not contradict coherency and 1-fold certainty in coherency. Inductively, for each $k > 1$, for every $(p_1, \ldots, p_{k-1}) \in P_1 \times \cdots \times P_{k-1}$ that does not contradict coherency, 1-fold, $\ldots$, and $(k-3)$-fold certainty in coherency, there are exactly two measures $p_k \in P_k$ such that $(p_1, \ldots, p_k)$ does not contradict coherency, 1-fold, $\ldots$, and $(k-2)$-fold certainty in coherency. Therefore, T_p has the same cardinality as $\{0, 1\}^{\mathbb{N}}$, implying that it is uncountable. Now, consider a belief hierarchy $(\pi_1, \pi_2, \ldots)$ such that

π_1 is uniformly distributed over Θ,

π_2 is uniformly distributed over $\Theta \times \text{proj}_{P_1} T_p$,

$\vdots$

14

π_k is uniformly distributed over $\Theta \times \text{proj}_{P_1 \times \cdots \times P_{k-1}} T_p$

$\vdots$

First observe that $(\pi_1, \pi_2, \dots)$ satisfies coherency and common certainty in coherency. Moreover, by construction $(\pi_1, \pi_2, \dots) \in T^Q$, and therefore $g(\pi_1, \pi_2, \dots) \in \Delta(\Theta \times T^Q)$. However, observe that $g(\pi_1, \pi_2, \dots)$ has an infinite support, and therefore by Proposition 1, $g(\pi_1, \pi_2, \dots) \notin \Delta^{\mathbb{Q}}(\Theta \times T^Q)$, which completes the proof. $\qquad \square$

References

ALIPRANTIS, C. & BORDER, K. (1994). *Infinite dimensional analysis*. Springer Verlag, Berlin.

AUMANN, R.J. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica* 55, 1–18.

——— (1999). Interactive epistemology II: Probability. *International Journal of Game Theory* 28, 301–314.

AUMANN, R.J. & BRANDENBURGER, A. (1995). Epistemic conditions for Nash equilibrium. *Econometrica* 63, 1161–1180.

BADEA, C. (1987). Irrationality of certain infinity series. *Glasgow Mathematical Journal* 29, 221–228.

BILLINGSLEY, P. (1995). *Probability and measure*. John Wiley & Sons, New York.

BRANDENBURGER, A. (2008). Epistemic game theory: complete information. *The New Palgrave Dictionary of Economics*, edited by L. Blume & S. Durlauf, McMillan.

BRANDENBURGER, A. & DEKEL, E. (1987). Rationalizability and correlated equilibria. *Econometrica* 55, 1391–1402.

——— (1993). Hierarchies of beliefs and common knowledge. *Journal of Economic Theory* 59, 189–198.

ELIAZ, K. (2003). Nash equilibrium when players account for the complexity of their forecasts. *Games and Economic Behavior* 44, 286–310.

FAGIN, R. & HALPERN, J. (1994). Reasoning about knowledge and probability. *Journal of the Association for Computing Machinery* 41, 340–367.

FREMLIN, D.H. (2003). *Measure theory: Broad foundations.* Vol. 2, Torres Fremlin, Essex.

GUPTA, A. (2000). *Real and abstract analysis.* Academic Publishers, Calcutta.

FISCHBEIN, E., JEHIAM, R. & COHEN, D. (1995). The concept of irrational numbers in high-school students and prospective teachers. *Educational Studies in Mathematics* 29, 29–44.

HALPERN, J.Y. (2001). Alternative semantics for unawareness. *Games and Economics Behavior* 37, 321–339.

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

HEIFETZ, A. (1993). The Bayesian formulation of incomplete information — The non-compact case. *International Journal of Game Theory* 21, 329–338

HEIFETZ, A., MEIER, M. & SCHIPPER, B. (2006). Interactive unawareness. *Journal of Economic Theory* 130, 78–94.

HEIFETZ, A. & MONGIN, P. (2001). Probability logic for type spaces. *Games and Economic Behavior* 35, 31–53.

HEIFETZ, A. & SAMET (1998). Topology-free typology of beliefs. *Journal of Economic Theory* 82, 324–341.

HU, T.W. (2009). Expected utility theory from the frequentist perspective. *Economic Theory*, forthcoming.

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

KRISHNA, R.V. (2007). Communication in games of incomplete information: Two Players. *Journal of Economic Theory* 132, 584–592.

LEHRER, E. (1996). Mediated talk. *International Journal of Game Theory* 25, 177–188.

MAENNER, E. (2008). Adaptation and complexity in repeated games. *Games and Economic Behavior* 63, 166–187.

MEIER, M. (2006). Finitely additive beliefs and universal type spaces. *Annals of Probability* 34, 386–422.

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

Modica, S. & Rustichini, A. (1999). Unawareness and partitional information structures. *Games and Economic Behavior* 27, 265–298.

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

Perea, A. (2012). *Epistemic game theory: Reasoning and choice*. Cambridge University Press, UK.

Pinter, M. (2005). Type space on a purely measurable parameter space. *Economic Theory* 26, 129–139.

Savage, L. (1972). *The foundations of statistics.* Dover, New York.

Shepherdson, J.C. (1980). Utility theory based on rational probabilities. *Journal of Mathematical Economics* 7, 91–113.

Siniscalchi, M. (2007). Epistemic game theory: Beliefs and types. *The New Palgrave Dictionary of Economics*, edited by L. Blume & S. Durlauf, McMillan, forthcoming.

Sirotic, N. & Zazkis, R. (2007a). Irrational numbers: The gap between formal and intuitive knowledge. *Educational Studies in Mathematics* 65, 49–76.

——— (2007b). Irrational numbers on the line – where are they? *International Journal of Mathematical Education in Science and Technology* 38, 477–488.

Spiegler, R. (2004). Simplicity of beliefs and delay tactics in a concession game. *Games and Economic Behavior* 47, 200–220.

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

Tan, T. & Werlang, S. (1988). The Bayesian foundations of solution concepts of games. *Journal of Economic Theory* 45, 370–391.

Tsakas, E. (2012). Universally rational belief hierarchies. *International Game Theory Review* (forthcoming).

Zhou, C. (2010). Probability logic of finitely additive beliefs. *Journal of Logic, Language and Informatics* 19, 247–282.