Department of Economics, University of Warwick

Monash Business School, Monash University

as part of

Monash Warwick Alliance

Hierarchies of Beliefs for Many Player Games

Venkata Tanay Kasyap Kondiparthy

Warwick-Monash Economics Student Papers

March 2024 No: 2024/72

ISSN 2754-3129 (Online)

https://warwick.ac.uk/fac/soc/economics/research/wmesp/

The Warwick Monash Economics Student Papers (WM-ESP) gather the best Undergraduate and Masters dissertations by Economics students from the University of Warwick and Monash University. This bi-annual paper series showcases research undertaken by our students on a varied range of topics. Papers range in length from 5,000 to 8,000 words depending on whether the student is an undergraduate or postgraduate, and the university they attend. The papers included in the series are carefully selected based on their quality and originality. WM-ESP aims to disseminate research in Economics as well as acknowledge the students for their exemplary work, contributing to the research environment in both departments.

Recommended citation: Kondiparthy, V.T.K (2024). Hierarchies of Beliefs for Many Player Games. *Warwick Monash Economics Student Papers* 2024/72.

[1] Warwick Economics would like to thank Gianna Boero and Samuel Obeng for their contributions towards the selection process.

Hierarchies of Beliefs for Many Player Games

Venkata Tanay Kasyap Kondiparthy †

Abstract

Mertens and Zamir (1985) first provided the universal type space construction for finite player games of incomplete information with a compact state space. Brandenburger and Dekel (1993) complemented it for a Polish state space. This paper extends the construction of Brandenburger and Dekel (1993) to games with infinitely many players for Harsanyi's notion of a type. The extension is formulated by randomly drawing a countably infinite set of actual players from a continuum of potential players, represented by their labels in [0,1]. The random distribution of the countably infinite set of actual players almost surely converges to Lebesgue due to the Glivenko–Cantelli theorem. A coherent type is shown to induce beliefs over other player's types and common knowledge of coherency closes the model of beliefs. Implications of dropping the Polish space assumption are discussed and an informal extension to measurable spaces is provided for future work. The formalisation provided here allows Harsanyi's notion of type to be applied in classes of games with many players such as Morris and Shin (2001).

JEL Classification: C70, D83.
Keywords: Higher Order Beliefs;Universal Type Space; Many Player Games.

†email: v.v.kondiparthy@lse.ac.uk, tanaykondiparthy@gmail.com
Online Appendix:Link. Please note all references to the Online Appendix within the text are hyperlinks.
The author thanks Prof Herakles Polemarchakis for his supervision and support. The author is grateful to Prof Peter Hammond for his invaluable guidance and support throughout the project. The author thanks Joseph Basford for discussions of the paper.

1 Introduction

Harsanyi (1967) proposed for games of incomplete information the concept of a player's type, which summarises all the beliefs and characteristics concerning that player. Harsanyi's formulation helped overcome the issue of considering an infinite hierarchy of beliefs as is common in games of incomplete information by making it more tractable. Players are uncertain about all the aspects which impact their strategies and payoffs. A Harsanyi type profile is an infinite hierarchy of beliefs, as in a strategic game a player needs to form their own beliefs about the parameters of the game, then beliefs about the beliefs of the other players regarding the parameters of the game then beliefs about the other players' beliefs about their own beliefs on the parameters of the game and so on ad infinitum. Furthermore, each type in Harsanyi (1967) implicitly induces a joint belief over the states of nature and the types of other players.

In a seminal paper, Mertens and Zamir (1985) formalised the implicit notion of a type given by Harsanyi (1967) to a corresponding explicit notion in terms of an infinite hierarchy of beliefs, which corresponds to a universal type space for a finite number of players. Brandenburger and Dekel (1993) proved the Mertens and Zamir (1985) results when starting with a Polish space (S)[1] instead of a compact space for the states of nature. There have been many papers in the literature replicating the results of Mertens and Zamir (1985) with various underlying assumptions about the states of nature for a finite set of players, including when S is a general measurable space by Heifetz and Samet (1998). The results of Mertens and Zamir (1985) and others remain yet to be extended in games with an infinite number of players. There are many classes of games with incomplete information in economics which deal with more than a finite set of players, e.g. the global games of Morris and Shin (2001).

The aim of the paper is to extend the construction of Brandenburger and Dekel (1993) to games with infinitely many players. The Brandenburger and Dekel (1993) setup was chosen for this extension because of the generality, applicability and mathematical tractability of Polish spaces. A novel method is proposed, motivated by Hammond (2017), to embed the countably infinite set of actual players randomly into a continuum of potential players, represented by their labels in [0,1]. The embedding serves two purposes: 1) It makes the model tractable by allowing beliefs over $[0,1]$, which is a Polish space, 2) The stochastic labelling mechanism as described in section (3.1) allows non-degenerate individual randomness without getting non-measurable sets. The main results of this paper are proposition (1) and proposition (2) which are analogues of the proposition of Brandenburger and Dekel (1993). Proposition (1) proves that, the space T_1', of coherent types is homeomorphic to $\Delta([0,1] \times S \times T_0')$ and proposition (2) shows that under the common knowledge of coherency, the space T^*, of all coherent types is homeomorphic to $\Delta([0,1] \times S \times T^*)$ thus extending the result of Brandenburger and Dekel (1993), where T^* is the universal type space in this extended framework. The need for the Polish space assumption is discussed – proposition (1) holds when S is weakened to a separable metric space, but this result does not extend to proposition (2).

In the Online Appendix, an extension for future work is given, with S as a measurable space,

[1] A Polish space is a complete, separable metric space, where a metric space is complete if every Cauchy sequence in the space converges, and separable if there is a countable dense subset of the space.

where the Ionescu-Tulcea theorem is used to show how to extend the hierarchy of beliefs to an infinite order in this setup. Although as in Heifetz and Samet (1998) whether a universal type space exists in this extended setup for the measurable case, and that it is a proper subset of the space of infinite order hierarchies in section (3.1) is left to be proved for future research.

The interested reader may refer to the Online Appendix before proceeding, for a mathematical preliminary.

2 Literature review

Mertens and Zamir (1985) formalised the idea mathematically, they wrote when "Starting from a set S of all possible values of the parameters of the game, can one identify a mathematically well defined set Y of the 'states of the world' in which every point contains all characteristics, beliefs and mutual beliefs of all players, such that any infinite hierarchy of beliefs would lead to some point in Y." They begin with a compact set S, so that $\Delta(S)$, the space of all probability distributions over S, is a compact metric space with the weak*topology Conway (2007). The weak*topology is the coarsest topology such that for all functionals in the dual space — linear maps from the original space into the real numbers — a sequence of functionals converges if it converges on each individual point of the original space[2].

Mertens and Zamir (1985) take $\Delta(S)$ as the first space X_1 in their belief hierarchy and then recursively define their k^{th} order belief hierarchies as $X_k = X_{k-1} \times \Delta(S \times (X_{k-1})^{n-1})$, where $\mu_k \in X_k$ is called a belief of order k. They show that each space in this sequence is compact in a metric space. Using their coherency condition, they recursively construct the space of coherent belief hierarchies(Z_k) whose projective limit [3] is the definition of the universal type space(T). The coherency condition stipulates an element in Z_k is a coherent belief hierarchy of order k if $Z_{k+1} := \{[\mu_{k+1} = (\mu_k, \nu_k) \in Z_k \times \Delta(S \times (Z_{k-1})^{n-1})] : \text{marg}_{\nu_k}\Delta(S \times (Z_{k-1})^{n-1}) = \nu_{k-1}\}$, where $\mu_k = (\mu_{k-1}, \nu_{k-1})$. The coherency condition on marginals ensures that successive beliefs in the belief hierarchy do not contradict each other. The pair $\mu_{k+1} = (\mu_k, \nu_k)$ is a coherent belief hierarchy of order $k+1$ if in addition ν_k attaches probability one to $S \times (Z_k)^{n-1}$ i.e. player 1 ascribes probability one to player 2 believing that player 3 believes ... that the belief hierarchy of order $k-1$ of player N is coherent (Maschler et al., 2013). Mertens and Zamir (1985) show that their universal type space is compact and is homeomorphic to the space $\Delta(S \times T^{n-1})$, of joint probability distributions over S and the types of other players(T^{n-1}), thus completing the formalisation and closing their model of beliefs.

A belief space in Mertens and Zamir (1985) for a finite set of players N is an $(N+3)$ tuple $(Y, S, f, (t^i)_{i \in N})$, where the set of states of the world Y is a compact set, S is a compact space, $f : Y \mapsto S$ is a continuous mapping which maps each state of the world to a state of nature and each player's type, $t^i : Y \mapsto \Delta(Y)$ is a continuous mapping from each state of the world to a probability distribution over Y. Equivalently, for any compact non-empty subset of the states of the world $\tilde{Y} \in Y$ a belief subspace can be defined. A general definition of the

[2] Online Appendix

[3] The projective limit of the sequence of the spaces$(Z_k)_{k=1}^{\infty}$ is the space T of all the sequences $(\mu_1, \mu_2, \ldots) \in \times_{k=1}^{\infty} Z_k$, where for every $k \in \mathbb{N}$ the belief hierarchy $\mu_k \in Z_k$ is the projection of the belief hierarchy $\mu_{k+1} \in Z_{k+1}$ on Z Maschler et al. (2013)

belief space which is equivalent to Aumann (1999a) is given in Maschler et al. (2013) with $(Y, \mathcal{Y}, S, f, (t^i)_{i \in N})$, where $(Y, \mathcal{Y})$ is a measurable space of states of the world and f is a measurable function. The universal belief space of Mertens and Zamir (1985), generated by S is Y= $S \times T^n$. Therefore, a state of world $y \in Y$ is a list, whose first coordinate is the state of nature $s(y) \in S$ in that state of the world(y), followed by a list of types $(t_i(w))_{i \in N} \in T^n$, for the N players at that state of the world. Mertens and Zamir (1985) show that every belief space for a set of N players and set of states of nature S is a belief subspace of the universal belief space. Therefore, any belief space can be mapped into the universal belief space in a unique manner, which preserves the manner in which states of nature and types are associated with states of the world.

This result of Mertens and Zamir (1985) has strong implications. It provides a formal proof for the Harsanyi (1967) notion of a type, that there is indeed an infinite hierarchy explicit description for each player's type and given any type, there are corresponding beliefs over the states of nature and types of other players. Further, the Mertens and Zamir (1985) result assures us that in a Harsanyi game of incomplete information, every type space with the same states of nature and set of players can be traced back to a universal type space, i.e. there is a unique belief-morphism, which preserves the beliefs of players and nature. Thus ensuring any relevant types are not excluded in the game.

The Mertens and Zamir (1985) construction of the universal type space was a topological construction in the sense that they started with a compact set S of states of nature, recursively defined k^{th} order belief hierarchies and showed they can be extended to an infinite hierarchy of coherent beliefs on S, which is the universal type space. In a series of companion papers, Heifetz and Samet (1998) and Heifetz and Samet (1999) discuss the construction of a universal type space without any topological assumptions, starting with S as a general measurable space. Aumann (1999a) and Aumann (1999b) takes a semantic approach. Heifetz and Samet (1999) showed that starting from a general measurable space S, an iterative construction of coherent hierarchies of beliefs(C) as in the topological case does not equate to a type in some Harsanyi (1967) game of incomplete information. However, Heifetz and Samet (1998) showed that even in the general measurable case, the unique universal type space(T) exists and as in Heifetz and Samet (1999) the universal type space is a subset or indeed a proper subset of the space C of coherent hierarchies and can be recovered from C in a transfinite process.

Brandenburger and Dekel (1993) provide a similar construction complementary to Mertens and Zamir (1985), with the change in assumption that the underlying space S of uncertainties(common) is compact with S being a Polish space. Brandenburger and Dekel (1993) construction is more elementary and makes the assumption of common knowledge of coherency more explicit. Their recursive construction of the type space is similar to Mertens and Zamir (1985) above. The construction for the two player case in their paper is now provided.

Considering only player i's belief hierarchies:

$$
\begin{aligned}
X_0^i &= S \\
X_1^i &= X_0^i \times \Delta(X_0^i) \\
X_2^i &= X_1^i \times \Delta(X_1^j) \\
&. \\
X_k^i &= X_{k-1}^i \times \Delta(X_{k-1}^j)
\end{aligned}
$$

This requires that a player must first form belief $\Delta(S)$ about the states of nature. As player i only knows his beliefs about S, thus she must form a belief about player j's beliefs about S. Therefore, the second order belief of a player is $\Delta(X_1^j)$ and this process follows iteratively. This construction is similar for player j. Brandenburger and Dekel (1993) state their construction generalises to n players and do not explicitly present it, however for the sake of illustration and to accustom the reader with the aim of the paper, it is presented as follows:

$$
\begin{aligned}
X_0^1 &= S \\
X_1^1 &= X_0^1 \times \Delta(X_0^1) \\
X_2^1 &= X_1^1 \times \prod_{j=2}^{n} \Delta(X_1^j) \\
&. \\
X_k^1 &= X_{k-1}^1 \times \prod_{j=2}^{n} \Delta(X_{k-1}^j)
\end{aligned}
$$

A subtle point to note in the representation of the hierarchies, it is implicit that the beliefs are independent. They can be presented more generally as e.g, $X_2^1 = X_1^1 \times \Delta(\prod_{j=2}^{n}(X_1^j))$ allowing for dependence between beliefs about other player's first order beliefs. The implicit independence representation was presented to be in line with the literature and for ease of notation, it does impact any of the results.

Brandenburger and Dekel (1993) define the type of a player to be an infinite hierarchy of beliefs, $t = (\pi_1, \pi_2, \pi_3, ...) \in T_0$ where $T_0 = \bigtimes_{k=0}^{\infty} \Delta(X_k)$, is the space of all possible type profiles for all players. A player knows their own type but not the types of the other players. Thus, for the model to be closed, a player's type must determine their belief over the states of nature and other player's types — another level of belief hierarchy, a belief over other player's types, beliefs over other players belief about your type and so on. This is as in a standard model of incomplete information where a player forms beliefs about the types of the other players and state of nature given their own type. Brandenburger and Dekel (1993) show that their coherency condition is required for a player to form a belief over $S \times T_0$, given a player's coherent type. However, common knowledge of coherency is required for a player's type to determine higher order beliefs over types — a player's belief over other player's belief about their type. This is because the coherency condition implies a player i can form belief over player j's type, but if player i does not know whether or not player j's type is coherent, she cannot form a belief over j's belief on her type, as without the coherency condition it is not possible to coherently extend the finite order hierarchy to the infinite order.

Brandenburger and Dekel (1993) refer to "common knowledge" in the probabilistic sense of assigning probability one to other player types being coherent. Their coherency condition is similar to Mertens and Zamir (1985) and states that given a type $t = (\pi_1, \pi_2, ..) \in T_0$, the type t is coherent if for every $k \geq 2$, $\text{marg}_{X_{k-2}} \pi_k = \pi_{k-1}$ where $\text{marg}_{X_{k-2}}$ is the marginal on the space X_{k-2}. This coherency condition is what is often referred to as the consistency condition in stochastic processes and Brandenburger and Dekel (1993) utilise Kolmogorov's extension theorem to show there is a homeomorphism from the space of coherent types to the space of joint probability distributions over S and T_0. They then showed that for the space of all coherent types where there is common knowledge of coherency(T), there is a homeomorphism to the space $\Delta(S \times T)$, of joint probability distributions over S and T, thereby making the model closed. It is further shown that the largest belief closed set is equal to the set satisfying common knowledge of coherency, highlighting that common knowledge of coherency is the minimum condition needed for closure[4]. In Brandenburger and Dekel (1993) the space T of all coherent types with common knowledge of coherency, which is a Polish space as well, is the equivalent of the universal type space of Mertens and Zamir (1985). Thus, Brandenburger and Dekel (1993) showed that the results of Mertens and Zamir (1985) hold even when we replace the assumption that the states of nature S is compact to S being Polish, however with finite S, the construction of Brandenburger and Dekel (1993) can be made a compact Polish space.

3 Model

The current literature as covered above is limited to games with a finite set of players. The aim of this paper, is to extend the construction of Brandenburger and Dekel (1993) to games with many players indexed by $\{I = \mathbb{N}\}$, who are randomly selected from a continuum of potential players. Thus, there is a countably infinite set of actual players who are randomly selected from continuum of potential players, given by their labels in [0,1]. The construction provided in the paper thus extends the Brandenburger and Dekel (1993) model to games with a continuum of potential players, allowing the consideration of the special case of a countably infinite set of actual players, randomly selected from the continuum.

There are many models in economics with many players such as the global games framework of Morris and Shin (2001), where there are infinite players and there is an underlying state of fundamental(economy)- θ. The state θ is drawn from a common prior and each player observes it with a noisy signal. Morris and Shin (2001) state that this noisy observation of the underlying state creates strategic uncertainty in the equilibrium, stemming from uncertainty regarding other player's payoffs. Thus players form higher order beliefs— beliefs over the other player's payoffs, then beliefs over the beliefs of other players, and so on. Morris and

[4] In a 1985 working paper version of Brandenburger and Dekel (1993)

Shin (2001) quote that " their analysis using higher order beliefs implicitly assumes common knowledge of the prior distribution of state θ and signalling technology". Morris and Shin (2001) citet Mertens and Zamir (1985) to support the common knowledge of the underlying state space and to form higher order beliefs. However, as stated, Morris and Shin (2001) is not applicable here as the number of players is not finite. No alternative construction is provided in Morris and Shin (2001) or the subsequent many player games literature to amend and apply the results of Mertens and Zamir (1985). The construction provided in this paper is a first. It provides a framework and justification for forming higher order beliefs in games with many players. In section (6) the applicability of the setup to many player games in the literature like Morris and Shin (2001) is discussed in detail.

3.1 Extension to Many Players

In this section, a novel and mathematically tractable construction is proposed to extend the construction, which involves embedding the set of countably infinite actual players(I) into a set of continuum of potential players.

The embedding described in Hammond (2017) is applied. Let the set of actual players be $I = (1, 2, \ldots)$ the countably infinite set of natural numbers. There will however be a continuum of potential players in the form of a non-atomic probability space $(L, \mathcal{L}, \lambda)$, where L is a topological label space, with $\mathcal{L}$ as its completed Borel σ-algebra— it contains not only all open sets, but all subsets of λ-null sets. Now each player $i \in I$ will be given a random label l^i chosen from a non-atomic probability space $(L, \mathcal{L}, \lambda)$. This line on stochastic labelling essentially means there is some underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and for each player $i \in I$, label l^i is a random variable, $l^i : \Omega \mapsto L$ that maps into the topological label space above $(L, \mathcal{L})$ i.e. the distribution of l^i is governed by the non atomic probability measure λ in the label space $(L, \mathcal{L})$. This random labelling embeds a model with a countable set of players into the standard notion of a continuum model with L as the set of potential players. In this continuum model, there is an actual player labelled l if and only if the label l happens to equal one of the countable random draws from $(L, \mathcal{L}, \lambda)$.

For the purposes of the paper, the non-atomic probability space $(L, \mathcal{L}, \lambda)$ is chosen to be $([0, 1], \mathcal{L}([0, 1]), \lambda)$ — the Lebesgue unit interval $[0, 1]$ is the label space, with its σ-algebra as defined above, and the measure λ is the Lebesgue measure, implying labels are drawn uniformly from [0,1]. This set up means the continuum of potential players as represented by their labels, are points in [0,1] and for each player $i \in I$, a label is randomly drawn uniformly from [0,1], so there are countably infinite draws. The uniform draw is the most natural case to consider, implying any label in [0,1] is equally likely to be realised as an actual player, but this set up works for any non atomic distribution. It is important to stress the key feature of the set up is that with the stochastic labelling process, we are randomly drawing countably infinite times from [0,1] uniformly, thus it can be thought of as having a sequence of countably infinite uniform random variables. Another key implication of this setup with a countably infinite set of players drawn from a continuum of potential players, is it allows non-degenerate individual randomness without getting non-measurable sets and measure zero sets (Hammond (2023)). This is illustrated by example (1).

To further illustrate the labelling and embedding, think of a simple case where there are two players- i and j. Each player has two labels x and y. From player i's perspective, she knows her own label(identity) and for example, attaches equal chance to the label of player j being x or y. In the setup above, the number of players is countably infinite, and their labels are a continuum ([0,1]) and thus now probabilities are attached to subsets of [0,1], as any specific point in [0,1] has zero probability of being chosen.

Given the setup, the extension of the Brandenburger and Dekel (1993) finite player framework to the many players is ready to be provided. The S–X pairs in the hierarchy of the Brandenburger and Dekel (1993) framework is replaced with a probability measure over a triple product space whose members consist of : (i) The identities i of other players (drawn from [0, 1]), (ii) The underlying states of uncertainty S, and (iii) recursively defined belief hierarchies similar to the original construction. Further the following initialisation condition is assumed:

Definition 3.1 (Uniform Marginal Condition)**.** $\Delta_\lambda := \{\Delta(S \times [0,1]) : \mathrm{marg}_{[0,1]}\Delta(S \times [0,1]) = \lambda\}$

Another key requirement of the setup as stated in Hammond (2017) is (3.1), which states that the marginal on [0,1] of the measure on the Cartesian product $S \times [0,1]$ must be the Lebesgue measure. This requirement as given in Hammond (2017) provides the existence of a stochastic transition, which is essential for applying the Ionescu-Tulcea theorem in the general measurable cases[5]. Intuitively, this requirement states when forming the first-order belief over states and player labels, the uniform nature of the distribution of player labels should not be violated. The Glivenko–Cantelli theorem in (3.2) provides further support that it is rational to assume this condition.

The condition (3.1) can be stated as a coherency condition if starting with [0,1] as the base space. However, since the base space in the setup, $X'_0 = (S \times [0,1])$ is a product, the coherency conditions in the paper are stated for maintaining consistency as we move to higher order spaces. Thus, for the remainder of the paper this condition holds as an initialisation requirement and is represented by Δ_λ. Note, as mentioned in the setup, the players labels can be drawn from any non-atomic distribution on [0,1], thus in a setting where that is the case, the initialisation condition is updated to reflect the chosen non-atomic distribution.

The hierarchy under this proposed extension would be as follows:

$$\begin{aligned}
X'_0 &= S \times [0,1] \\
X'_1 &= X'_0 \times \Delta_\lambda(X'_0) \\
&. \\
X'_k &= X'_{k-1} \times \Delta_\lambda(X'_{k-1})
\end{aligned}$$

The novelty and beauty of randomly embedding the countably infinite player case into a continuum of potential players in the Lebesgue unit interval is its tractability, as [0,1] is a compact Polish space (with the standard metric) and the space $\Delta([0,1], \mathscr{B})$, of all Borel probability

[5]Online Appendix

measures on [0,1] is also a Polish space given the topology of weak convergence of probability measures or weak*topology(Prohorov theorem). Since [0,1] and S are Polish spaces, X'_0 is a product of two Polish spaces, thus Polish. As stated above the space of probability measures of a Polish space is Polish with the weak*topology thus $\Delta_\lambda([0,1] \times S)$ is a Polish space as well. The belief hierarchies are Polish spaces as any finite and a countable/uncountable Cartesian product of Polish spaces is also a Polish space (Bertsekas and Shreve, 1996; Hammond, 2004). Thus, the type space here would remain a Polish space as in the original Brandenburger and Dekel (1993) construction.

Following Brandenburger and Dekel (1993) notation a type takes the form $t = (\pi_1, \pi_2, \pi_3, \ldots) \in T'_0 = \bigtimes_{k=0} \Delta(X'_k)$, where T'_0 is the analogue of all possible types in Brandenburger and Dekel (1993). The main result of the paper is to establish suitable coherency conditions and show the model is closed as in Brandenburger and Dekel (1993), which would involve showing there is a homeomorphism from the space T'_1 of coherent types to $\Delta_\lambda([0,1] \times S \times T'_0)$. The interpretation here would be that each coherent type is associated with a unique belief measure over the space of player identities, the original states of uncertainty, and all possible extended types. Then showing under the common knowledge of coherency we have a homeomorphism from the space T^* of all coherent types where this is true to $\Delta([0,1] \times S \times T^*)$ extending their result that common knowledge of coherency is the minimum condition for model-closure and T^* is the universal type space of Brandenburger and Dekel (1993) in this extended framework. However, before presenting the coherency conditions the Glivenko–Cantelli Theorem is stated which follows from the works of Parthasarathy (1967) and Varadarajan (1958) and serves as a consistency check for the random labelling in the construction per Hammond (2017) as it provides support to the initialisation condition Δ_λ.

3.2 Glivenko–Cantelli Theorem

In the stochastic labels setup, a countably infinite numbers of labels are drawn randomly and uniformly, constituting a countably infinite sequence of independent uniform random variables. Therefore, there is an infinite sequence of points, with each point in [0,1]. As in the construction a player is forming beliefs over this countably infinite sequence we need to consider the behaviour/ distribution over these points before defining the coherency conditions, to support that the initialisation condition (3.1) was reasonable. The Glivenko-Cantelli theorem thus provides us with a consistency check.

Let (U, d) be any separable metric space, and let α be any probability measure on its Borel σ-algebra. Let α^N denote the infinite product measure on the collection U^N of countably many copies of (U, d), equipped with its product σ-algebra. For each infinite sequence $u^N = (u_k)_{k \in \mathbb{N}} \in U^N$ and for each $n \in \mathbb{N}$, let

$$\alpha_n(u^N) := \frac{1}{n} \sum_{k=1}^{n} \delta_{u_k} \quad (1)$$

denote the empirical measure on U which is generated by $\{u_1, u_2, \ldots, u_n\}$—i.e. the first n elements of u^N. Here δ_{u_k} is per the standard notation the degenerate measure satisfying $\delta_{u_k}(\{u_k\}) = 1$. (Patrick Billingsley, 2012; Hammond, 2017).

The Glivenko Cantelli Theorem: Let (U,d) be any separable metric space, and α be any probability measure on its Borel σ-algebra. Then, for α^N-a.e. $u^N \in U^N$, the empirical measure $\alpha_n(u^N)$ defined by (1) converges weakly to α. (Patrick Billingsley, 2012; Hammond, 2017).

The Glivenko–Cantelli theorem says that the empirical measure converges to the original measure as n increases. After n draws of player labels, consider any sub interval, say [0.25, 0.5]. In the "true" uniform distribution, 25% of the numbers lie in this sub interval. The Glivenko–Cantelli theorem states that as n becomes larger and larger, the proportion of numbers from the draws that fall in the sub-interval [0.25, 0.5] will converge to 25%. Thus, as we start with a finite subset of actual players and extend to the countably infinite subset of actual players, the empirical distribution of the countably infinite set of actual players converges almost surely to the uniform distribution.

Example This is an example of how the Glivenko–Cantelli theorem is applicable to the construction. (U,d) corresponds to $([0,1],d)$ where d is a metric, in this setup. The Glivenko–Cantelli theorem is helpful in the construction as follows: in the first order belief, player 1, forms beliefs over S and who the actual set of players in the game are. Let's assume for simplicity S is binary and player 1 attaches some probability to the two states, and she then attaches probability 0.4 that the label of player 2 is in [0,0.4], 0.6 probability in the rest. With probability 0.6 the label of player 3 is in [0,0.6], 0.4 probability in the rest, and so on for the set of countably infinite players. Now the Glivenko—Cantelli theorem is telling us that as n becomes larger — the number of players about whose identities player 1 forms beliefs becomes larger — the proportion's [0,0.4], [0,0.6] etc. will truly have the proportion of labels as prescribed by the uniform distribution, so 40% of the labels of player 2 lie in [0,0.4], 60% of the labels of player 3 in [0,0.6] etc. The Glivenko–Cantelli theorem therefore provides a consistency check that the label draws for the players are still behaving uniformly.

In a crude sense, the Glivenko–Cantelli theorem implies that it becomes "common knowledge" that player labels are being drawn from a uniform[0,1], every player knows when forming beliefs about the other countably infinite actual players, that their labels indeed follow a uniform distribution in [0,1]. Thus, the result of the Glivenko Cantelli theorem is essential and it supports that it is rational for players to hold the initialisation condition (3.1), as their beliefs about the distribution of other player labels aligns with the true distribution of the labels in the limit.

Note, the theorem is given for any metric space (U,d) and any probability measure, thus this feature of the stochastic labelling setup is not a special case of the Lebesgue measure/ uniform distribution. For instance, if the labels were instead drawn from a $\beta(2,2)$ distribution, then every player knows that the labels of players are drawn from the $\beta(2,2)$ distribution. Further, given the nature of the $\beta(2,2)$ distribution a higher proportion of points are distributed around 0.5 than around 0 or 1, due to the Glivenko–Cantelli theorem whatever skewness is implied by the distribution of the labels will be naturally reflected when the players form first-order beliefs about the identities of the other players. This generality of the stochastic labelling setup and the Glivenko–Cantelli theorem provides a rich framework which is applicable to a wide class of games, where it is not necessary for all player labels to be equally likely and any other non-atomic distribution may be adopted to suit the specific needs of the model. The definition of a coherent type $t \in T_0'$ is stated next.

3.3 Coherency Conditions

Definition 3.2 (Coherent Type)**.** *A type is defined to be coherent if it meets the following two conditions:*

(i)Consistency of higher order beliefs*: For each $k \geq 2$, a belief $\pi_k \in \Delta_\lambda(X'_{k-1})$ is coherent if the marginal of π_k on X'_{k-2} is π_{k-1}. This condition ensures that a player's k^{th}-order belief about lower-order beliefs (as encoded by π_k) is consistent with their $(k-1)^{th}$-order belief π_{k-1}.*

(ii)Consistency of player identity and state beliefs*: For each $k \geq 2$, a belief π_k is coherent if the marginal of π_k on $[0,1] \times S$ is consistent with the marginal of π_{k-1} on $[0,1] \times S$–* $\text{marg}_{[0,1]\times S}\pi_k = \text{marg}_{[0,1]\times S}\pi_{k-1}$. *This is the joint consistency condition and ensures that the measure over the product space is consistent with the player's identity and state of nature.*

Condition (i) is a simple extension of the coherency definition of Brandenburger and Dekel (1993) in this extended framework. It just states that in the hierarchy construction, the higher order beliefs do not contradict lower level beliefs. This is akin to the stochastic consistency condition referred to for stochastic processes. The second condition is the extended joint consistency condition, given the extended framework which includes not just beliefs about the underlying space of uncertainty S, but also beliefs about player identities. Condition (ii) implies that the beliefs about player identities and S should be preserved when you go from one level of the belief hierarchy to the next. To facilitate understanding, a simplification of condition (ii) for the case of independence is provided. The homeomorphism results would trivially be applicable to the independence case, however it is stated as a special case to facilitate understanding and due to its application in the Ionescu-Tulcea theorem[6].

(ii′)Independence of player identity and state beliefs:

For each $k \geq 2$, a belief π_k is coherent if the marginal of π_k on $[0,1] \times S$ is consistent with the product measure of the marginal of π_k on $[0,1]$ and the marginal of π_k on S i.e. $marg_{[0,1]\times S}\pi_k =$ $\text{marg}_{[0,1]}\pi_{k-1} \times \text{marg}_S\pi_{k-1}$. *This is the joint consistency condition and ensures that the measure over the product space is consistent with the player's identity and state of nature.*

Condition (ii′) implies that there is independence in k^{th} order beliefs about player identities and S. When combined with condition (i), as one project's down to lower order beliefs, the independence relation is maintained until at the base level-X'_1, thus, implying that the initial belief of a player, about whom the other player's in the game are and the player's belief about S are independent. The independence is maintained as one goes higher along the construction. The last line is the intuition one expects to be maintained when constructing the type, whatever initial beliefs about player identities and S are formed, must be maintained across the hierarchy.

Condition (ii) is more general and allows for any form of dependence to exist between initial beliefs about player identities and S. This might be useful in games where S might reflect the state of the economy and players are deciding whether to invest in a particular asset class or not. If a player's belief about the state of the economy is bad, then the player would expect the other players in this investment game to be more risk taking (could be done by

[6]Online Appendix.

modelling player identities closer to 1 as risk loving) thus, a player might have an initial joint distribution over [0,1] and S, which reflects an inverse relationship.

4 Examples

In this section examples are provided to highlight the importance of the random labelling for sensible results, the working of the construction and coherency condition. For all examples, restrict to the case where the space S is binary– $S = \{s_1, s_2\}$ for simplicity. The construction is provided from player 1's perspective as it is the same construction for all players.

Example-1: Deterministic Labelling

This example illustrates the importance of random labelling for producing sensible results. For simplicity, assume a player's beliefs about the identities of the other players and states is independent. Let $[0,1]_{\mathbb{Q}}$ denote the set of rational numbers in [0,1], the set of rational numbers is a countably infinite subset of [0,1] and it is natural to think of using it to represent the set of countably infinite actual players in [0,1], as it is a countably dense subset of the reals. Thus, the labelling is deterministic and non-random.

The crudest example where a player 1 with label l=0.5 believes with probability 0.7 the state is s_1 and with probability 0.3 the state is s_2. For simplicity, let's assume regardless of the state she believes with equal probability the set of countably infinite actual players in the game are $[0,0.5]_{\mathbb{Q}} - \lambda([0,0.5]_{\mathbb{Q}})$ – and in $(0.5,1]_{\mathbb{Q}}$–$\lambda((0.5,1]_{\mathbb{Q}})$. The notation $\lambda[a,b]$, corresponds to the uniform distribution—attaching a probability equal to the length of the interval. In this example, we thus have $0.5([0,0.5]_{\mathbb{Q}})$.

The first-order belief of the player is:

$\pi_1 = 0.7(s_1 \otimes \hat{\lambda}) + 0.3(s_2 \otimes \hat{\lambda}), \hat{\lambda} = [\lambda([0,0.5]_{\mathbb{Q}}), \lambda((0.5,1]_{\mathbb{Q}})]$

The second-order belief of the player, is the first-order belief and a belief over the first-order beliefs of all the other players in $[0,0.5]_{\mathbb{Q}}$ and $(0.5,1]_{\mathbb{Q}}$ which is assumed to be the same as her first-order beliefs—the first order beliefs of all the other players in the game is the same as the first-order belief of player 1. Due to notational complexity, π_2 is not explicitly written, but it follows a similar form as π_1 and conveys the above information. The construction follows iteratively for higher order beliefs. However, in this setup, deterministically taking the set of actual players to be the set of rationals in [0,1] has the following issue. The set of rationals has Lebesgue measure zero and in the first-order belief, player 1 ascribes a positive measure to a measure zero set, implying there is a positive probability that a zero probability event occurs, a mathematical impossibility. Thus, the first critique with such deterministic labelling in the belief construction is we end up attaching non degenerate probabilities to measure zero sets. Further, the infinitely many random draws of labels allowed the use of the Glivenko–Cantelli theorem, and in the deterministic setting it is not applicable, and obviously in the limit case the distribution of the actual set of player labels is not almost surely uniform. Therefore, deterministic labelling is not applicable to the class of games such as Morris and

Shin (2001) where one wants in the limit with a continuum of (potential) agents represented by the Lebesgue unit interval, as a canonical non-atomic measure.

With the stochastic labelling mechanism as described in the setup, given the label space of [0,1] and its completed Borel sigma algebra, which includes subsets of λ-null i.e. measure zero sets, allows the possibility for a subset of rationals in [0,1] to be chosen. However, given the non-atomic nature of the measure on the labels in [0,1], no particular point in [0,1] has a positive probability of being chosen, thus the probability a rational number is drawn as a player's label is zero, and given that the set of rationals is countable, rationals in [0,1] can be enumerated as a sequence $q_1, q_2, q_3, ...$ The measure of each q_i, $\lambda(q_i)$=0 from the non atomic nature and the probability of a player's label being a rational in a subset of [0,1] as in the above example $[0, 0.5]_{\mathbb{Q}}$ is equal to zero, as $\lambda([0, 0.5]_{\mathbb{Q}})= q_1 + q_2 + ..= 0$. Thus, under stochastic labels degenerate probabilities can be assigned to measure zero sets.

Example 2: Independence

The example illustrates when the initial belief about states and player-identities is independent— regardless of the state, the beliefs about player identities is the same— how the construction works, checking condition (3.1) and the coherency conditions.

Sub-case: Uniform Marginal Condition:
First, let's illustrate the uniform marginal requirement (3.1). When the first-order belief about S and player identities is independent, (3.1) requires the subjective probabilities a player forms about the labels of other players being in a subset of $[0, 1]$ to be Lebesgue— player 1 believes the labels of all other players is in [0,0.5] with probability 0.5 and in [0.5,1] with probability 0.5, or equivalently, with probability 0.4 in [0,0.4] and probability 0.6 in [0.4,1]. Let's for simplicity, only focus on player 1's belief about the label of player-2(l_2). Let player 1 believe with probability 0.8, the state is s_1 and 0.2 the state is s_2. Regardless of the states she believes with equal probability the label of player 2 is in [0,0.5] and [0.5,1] respectively. The marginal of the first order belief on [0,0.5]= (0.8*0.5) + (0.2*0.5)= 0.5. Similarly the marginal on [0.5,1]=0.5. Thus, the marginal of the first order belief on [0,1] is indeed Lebesgue, satisfying condition (3.1).

The checking of the uniform marginal requirement is similar for the labels of the countably infinite set of all other players. Now, an example on the construction and checking of the coherency conditions is presented. Let player 1 with label $l = 0.6$, believe with probability 0.7 the state is s_1, and 0.3 the state is s_2. Regardless of the state, player 1 believes with certainty the label of player 2 is in [0,1]. Note, the spotlight is on player 2, for notational simplicity[7], to efficiently illustrate the checking of the coherency conditions.

The first-order belief of player 1 is:

$$\pi_1^1 = 0.7(s_1 \times \lambda(l_2)) + 0.3(s_2 \times \lambda(l_2))$$

[7] The example if presented for all players labels would be as follows: $\pi_1 = 0.7(s_1 \otimes \hat{\lambda}) + 0.3(s_2 \otimes \hat{\lambda}), \hat{\lambda} = [\lambda(l_2), \lambda(l_3), \ldots]$. Where, $\hat{\lambda}$ is a list for the labels of all other countably infinite players, with each $\lambda(l_i)$ implying $l_i \in [0, 1]$ with probability-1, $\forall i \in I$.

Where $\lambda(l_2)$ implies with probability-1 the label of player 2 is in [0,1]. This as stated above, satisfies (3.1) as the first-order belief about player 2's label is Lebesgue. The second-order belief of player $1(\pi_2^1)$, is player 1's first-order beliefs about player 2's first-order beliefs. Let player 1 believe after $S = s_1$ and $l_2 \in [0,1]$, that player 2 believes with probability 0.6, the state is s_1 and probability 0.4, the state is s_2. Regardless of state, player 2 believes with certainty the label of player-1 lies in [0,1]. Player 1 has similar second-order beliefs after $S = s_2$ and $l_2 \in [0,1]$.

$$\pi_1^2 = 0.6(s_1 \times \lambda(l_1)) + 0.4(s_2 \times \lambda(l_1))$$

$$\pi_2^1 = 0.7([s_1 \times \lambda(l_2)] \otimes \pi_1^2) + 0.3([s_2 \times \lambda(l_2)] \otimes \pi_1^2)$$

π_1^2 is the first-order belief of player 2, from the perspective of player 1. Coherency condition (i) requires the marginal of π_2^1 on X_0' is equal to π_1^1. To check, integrate/sum out the components from $\Delta(X_0')$ as $X_1' = X_0' \times \Delta(X_0')$ and the marginal of this on (X_0') is being taken. Thus, integrate/sum over the inner beliefs(π_1^2)— ignoring π_1^2— to be left with:

$$\text{marg}_{X_0'}\pi_2^1 = 0.7(s_1 \times \lambda(l_2)) + 0.3(s_2 \times \lambda(l_2)) = \pi_1^1$$

The construction does satisfy the first coherency condition. Coherency condition (i) simply ensures that, as player 1 form's π_2 she is not forgetting or changing her initial beliefs about the state's $\{s_1, s_2\}$ and the label of player 2(other players). Checking condition (i) is similar for higher order beliefs, e.g. for third order beliefs, it requires marginal of π_3 on X_1' is equal to π_2 and so on.

On the contrary, if in forming π_2^1, we had instead:

$$\pi_2^1 = 0.8([s_1 \times \lambda(l_2)] \otimes \pi_1^2) + 0.2([s_2 \times \lambda(l_2)] \otimes \pi_1^2)$$

$$\text{marg}_{X_0'}\pi_2^1 = 0.8(s_1 \times \lambda(l_2)) + 0.2(s_2 \times \lambda(l_2)) \neq \pi_1^1$$

Since, regardless of the state, player 1's beliefs about the label of player 2 was the same, the example was designed to make player 1's beliefs about labels and states independent.

Checking coherency condition(iib):

$$\text{marg}_{[0,1]\times S}\pi_2^1 = \pi_1^1; \text{ from condition(i).}$$

$$\text{marg}_{[0,1]}\pi_1^1 = \lambda(l_2); \text{ focusing only on the [0,1] component of } \pi_1^1$$

$$\text{marg}_S\pi_1^1 = [0.7(s_1) + 0.3(s_2)]$$

$$\text{marg}_S\pi_1^1 \times \text{marg}_{[0,1]}\pi_1^1 = 0.7(s_1 \times \lambda(l_2)) + 0.3(s_2 \times \lambda(l_2)) = \text{marg}_{[0,1]\times S}\pi_2^1$$

The above example while simplistic shows that the belief construction as given would meet the requirements of coherency. When explicitly stated for all players, instead of just player 2 would entail similar calculations, but they can be elicited through intuition, as the coherency conditions state when going across the hierarchy, the previous level beliefs should not change. Further if started with a base belief about independence between player identities and states of the world, this belief must be maintained throughout the hierarchy — once a higher level hierarchy is reached, a player should not suddenly believe their beliefs about the state impacts their belief about player identities.

Example 3: General

This example, illustrates the construction when the beliefs about player labels and states are allowed to be any joint distribution $\Delta_\lambda(S \times [0,1])$. Note, to meet condition (3.1) on the first-order belief here, does not require the subjective beliefs about players to be Lebesgue. For example, let player 1 ascribe a 0.5 probability to state s_1 and given s_1, a 0.4 chance l_2 in [0,0.5] and a 0.6 chance in [0.5,1]. A 0.5 probability to state s_2 and given s_2, a 0.6 chance l_2 in [0,0.5] and a 0.4 chance in [0.5,1]. Then the marginal of l_2 in [0,0.5] is 0.5*0.4+ 0.5*0.6= 0.5, marginal of l_2 in [0.5,1] is 0.5*0.6 + 0.5*0.4= 0.5, thus maintaining the uniform marginal condition. Condition (3.1) made the player form beliefs about s_1 and s_2 in a manner such that this condition is met.

Let the first-order belief of player 1 be:

$$\pi_1^1 = 0.5(s_1 \times \delta_{[0,0.5]}(l_2)) + 0.5(s_2 \times \delta_{[0.5,1]}(l_2))$$

Here $\delta_{[0,0.5]}(l_2)$ implies player 1 believes with certainty the label of player 2 is in [0,0.5]. In this example, player 1's beliefs about l_2 are different under s_1 and s_2, in particular now in s_1, player 1 beliefs the label of player 2 can only be in [0,0.5] and in s_2, believes the label can only be in [0.5,1]. Let the first-order belief of player 2 from the perspective of player 1 be as follows:

$$\pi_1^2 = 0.8(s_1 \times \delta_{[0,0.8]}(l_1)) + 0.2(s_2 \times \delta_{[0.8,1]}(l_1))$$

Let the second order-belief of player 1 be:

$$\pi_2^1 = 0.5([s_1 \times \delta_{[0,0.5]}(l_2)] \otimes \pi_1^2) + 0.5([s_2 \times \delta_{[0.5,1]}(l_2)] \otimes \pi_1^2)$$

In the second-order belief, player 1's belief about the first-order belief of player 2 is the same after s_1, s_2 and player labels as in the previous example. Checking coherency condition (ii) only, checking condition (i) is straightforward.

$\text{marg}_{[0,1]\times S}\pi_2^1 = \text{marg}_{[0,1]\times S}\pi_1^1$; condition(ii).

$\text{marg}_{[0,1]\times S}\pi_2^1 = \pi_1^1$; from condition(i).

$\text{marg}_{[0,1]\times S}\pi_1^1 = \pi_1^1$; as $\pi_1^1 \in \Delta([0,1] \times S)$ and is the marginal of itself.

Unitl the second-order belief, it was trivial to check condition (ii), but for third-order beliefs it requires marginal of π_3 on X_0' is equal to the marginal of π_2 on X_0' and so on.

5 The Universal Type Space

The Kolmogorov extension theorem (KET) allows the construction of a unique probability measure on an infinite family of random variables/ infinite product of measure spaces. The KET is presented as in Cohn (2013):

Theorem 5.1 (Kolmogorov Extension Theorem)**.** *Let I be a nonempty set, and let $\{(\Omega_i, A_i)\}_{i\in I}$ be an indexed family of measurable spaces. Let $\mathcal{I}$ be the collection of all nonempty finite subsets of I. For each $I_0 \in \mathcal{I}$ and each $i \in I$, define the product measurable spaces $\{(\Omega_{I_0}, A_{I_0})\}_{I_0\in I}$ and (Ω, A), plus projections $X_{I_0} : \Omega \to \Omega_{I_0}$ and $proj_{I_2,I_1} : \Omega_{I_1} \to \Omega_{I_2}$, where $I_0, I_1, I_2 \in \mathcal{I}$ and $I_2 \subseteq I_1$. Let $\{P_{I_0}\}_{I_0\in I}$ be an indexed family of probability measures on the spaces $\{(\Omega_{I_0}, A_{I_0})\}_{I_0\in I}$. If*

1. *the measurable spaces $\{(\Omega_i, A_i)\}_{i\in I}$ are all standard, and*
2. *the measures $\{P_{I_0}\}_{I_0\in I}$ are consistent, in the sense that they satisfy*

$$P_{I_2} = P_{I_1} proj^{-1}_{I_2,I_1} \text{ for all } I_1, I_2 \in I \text{ such that } I_2 \subseteq I_1,$$

then there is a unique probability measure P on (Ω, A) such that for each I_0 in $\mathcal{I}$ the distribution of X_{I_0} is P_{I_0}.

A standard measurable space is a Polish space (Cohn, 2013), in the setup Ω_i are equal to $(X'_i) : i \in (0, 1, 2..)$ which are all Polish spaces and thus meet condition (1) of the theorem. Condition (2) of the theorem requires, when you project down(marginal distribution) from a larger space $(\{(\Omega_{I_2}, A_{I_2})\}$ to a smaller space $(\{(\Omega_{I_1}, A_{I_1})\}$ i.e. $(I_2 \subseteq I_1)$, the marginal distribution must be the same as the original distribution P_{I_1}, initialised on the I_1 space. The coherency condition (i) ensures when you project down from a larger space (X'_k) to a smaller space (X'_{k-1}), the marginal distribution on (X'_{k-1}) is same as the true distribution for (X'_{k-1}). However, each (X'_k) contains an $(S\times[0,1])$ component thus when projecting down it must be ensured that the beliefs about $(S \times [0,1])$ are maintained as well. To illustrate via an example, $\pi_3 = \Delta(X'_2) = \Delta_\lambda[((S\times[0,1])\times\Delta_\lambda((S\times[0,1]))\times\Delta_\lambda((S\times[0,1])\times\Delta_\lambda((S\times[0,1]))]$ and when you project it down on X'_1, per coherency condition (i)– $\pi_2 = \Delta_\lambda[((S \times [0,1]) \times \Delta_\lambda((S \times [0,1]))]$, here coherency condition (ii) ensures that the inner belief component $\Delta_\lambda((S \times [0,1]))$ in π_3, when projected down must match the original belief in π_2.

The analogue of the first homeomorphsim result as in Brandenburger and Dekel (1993) is stated. The proof is more descriptive but essentially the same as Brandenburger and Dekel (1993).

Proposition 1. *If $(Z_k)_{k\geq 0}$ are Polish spaces and:*

$D := [\pi = (\pi_1, \ldots) : \pi_k \in \Delta(Z_0, \ldots, Z_{k-1})$ *and π is coherent*], *then there exists a homeomorphism* $(f : D \to \Delta(\times_k Z_k))$.

Proof. For any $\pi \in D$, since π is conherent, by the KET (5.1) there exists a unique probability measure on $(\Delta(\times_k Z_k)$ corresponding to this π. Letting $f : D \to \Delta(\times_k Z_k)$ be defined as the mapping induced by the KET (5.1) used on each $\pi \in D$. The mapping f is injective as (5.1) guarantees the uniqueness of corresponding product measures. Further, for $\pi \in \Delta(\times_k Z_k)$, $f(\text{marg}_{Z_0}\pi, \text{marg}_{Z_0,Z_1}\pi, \ldots) = \pi$ thus f is surjective. Thus by using (5.1) f is a bijection is obtained.

Bi-continuity of f is left to be shown. To show f is continuous, consider a sequence $(\pi_{k_1}, \pi_{k_2}, \ldots) \in D$ such that each probability measure ($\pi_{k_i} \to \pi_i$) weakly. Aim is to show that $f(\pi_{k_1}, \pi_{k_2}, \ldots) \xrightarrow{k\to\infty} f(\pi_1, \pi_2, \ldots)$. Let $\pi_k = f(\pi_{k_1}, \pi_{k_2}, \ldots)$, $\pi = f(\pi_1, \pi_2, \ldots)$, we want to show π_k converges weakly to π. Now note that $f(\pi_{k_1}, \pi_{k_2}, \ldots)$ and $f(\pi_1, \pi_2, \ldots)$ on cylinder sets are entirely determined by $(\pi_{k_1}, \pi_{k_2}, \ldots)$ and $(\pi_1, \pi_2, \ldots)$ as cylinder sets from a basis for the topology and thus form a convergence determining class[8] and we are done.

The inverse mapping $f^{-1}(\pi) = (\text{marg}_{Z_0}\pi, \text{marg}_{Z_0,Z_1}\pi, \ldots)$ is continuous as the marginal maps $\pi \mapsto \text{marg}_{Z_0,Z_1,\ldots,Z_{k-1}}\pi$, $k \geq 1$ are continuous. Since π_k converges to π weakly, it is to be shown that its associated marginals converge weakly as well. A sequence of measures π_k on the product space ($Z = \times_k Z_k$) converging to π weakly or in the weak*topology implies for every bounded continuous function g on Z, we have: $\int g \, d\pi_k \to \int g \, d\pi$.

Let h be a bounded continuous function over the entire product space Z, then its marginal h', over a subset $Z_0, Z_1, \ldots Z_k$, is obtained by integrating out the other factors and is thus as follows: $h'(z_0, z_1, \ldots, z_k) = \int_{z_{k+1},\ldots} h(z_0, z_1, \ldots) dz_{k+1} dz_{k+2} \ldots$. Because π_k's converge to π weakly: $\int h \, d\pi_k \to \int h \, d\pi$ and $\int h' d\text{marg}_{Z_0,Z_1,\ldots,Z_k}\pi = \int h \, d\pi$, $\int h' d\text{marg}_{Z_0,Z_1,\ldots,Z_k}\pi_k = \int h \, d\pi_k$ from the definition of marginals. Therefore it follows that $\int h', d\text{marg}_{Z_0,Z_1,\ldots,Z_k}\pi_k \to \int h', d\text{marg}_{Z_0,Z_1,\ldots,Z_k}\pi$ and thus the sequence of marginals converge weakly. □

Corollary 1.1. *There exists a homeomorphism* $f : T_1' \to \Delta_\lambda(([0,1] \times S) \times T_0')$.

Proof. In the previous proposition, let $Z_0 = X_0' = S \times [0,1]$ and $Z_k = \Delta_\lambda(X_k)$. These are all Polish and $\times_k Z_k = (S \times [0,1]) \times T_0'$. The set of coherent types T' is exactly the set D as it is the set of all coherent hierarchies. Thus, the result immediately follows from the previous proposition. □

A coherent player, given their type is able to form beliefs about the players identities, states of nature and the types of other players from the above corollary, providing the analogue of the first result of Brandenburger and Dekel (1993) in this extended framework. Please note as stated by Brandenburger and Dekel (1993) the closure of the model of belief hierarchies is not a pure measure theoretic result. This is as stated above, to apply the Kolmogorov extension theorem (5.1) requires the space to be a standard measurable space, which is satisfied here, as it was assumed to be Polish. Cohn (2013) states equivalently a standard measurable space is one which is Borel isomorphic to a compact metric space[9]. Dellacherie and Meyer (1978) give the Kolmogorov extension theorem for separable metric spaces only, which can be seen to follow from the standard result that every separable metric space is homeomorphic to a

[8]As cylinder sets form basis for the topology, every open set in that topology can be written as a union of these basis sets. This means that, for measures, if two measures agree on all the basis sets, they will also agree on all the open sets (since every open set can be expressed in terms of these basis sets). In turn, since the Borel sigma-algebra is generated by the open sets, two measures that agree on all open sets will agree on all Borel sets, thus cylinder sets form a convergence determining class.

[9]Online Appendix

subset of the Hilbert Cube, which is a compact metric space. Thus, the above result could be obtained with S as a separable metric space.

However, as detailed, a player knows only their type and not the types of other players and thus a "second level" hierarchy is needed, which as elucidated before requires the "common knowledge" of coherency. As in Brandenburger and Dekel (1993) define a sequence of sets, $T'_l, l \geq 2$ by: $T'_l = \{t \in T'_1 : f(t)(S \times [0,1] \times T_{l-1}) = 1\}$. Each T'_l represents a deeper layer of knowledge or belief. So here for e.g T'_2 corresponds to the set of all coherent types for a player, who believe with probability one that the other player's types are coherent. T'_3 would be a the set of all coherent types for a player, who believes with probability one that all other coherent type players believe with probability one, that the type of the player is coherent, so on for T'_l. Thus as defined by Brandenburger and Dekel (1993) let $T^* = \cap_{l=1}^{\infty} T'_l$ be the set of all coherent types with common knowledge of coherency. As mentioned in Brandenburger and Dekel (1985), for T'_l to be well defined, the space T'_{l-1} must be a Borel set as beliefs(probability measures) over T'_{l-1} are being taken. Each T'_l is in fact a Borel Set and the proof follows from following lemma:

Lemma 5.2. *For each $l \in \mathbb{N}$, the space T'_l is closed.*

Proof. :
The set D defined in proposition (1) is closed since closed since the maps $\pi \mapsto \text{marg}_{Z_0,Z_1,\ldots,Z_{k-1}} \pi$ are continuous, and by (1.1) it follows that T'_1 is closed. Now assume inductively that T'_{l-1} is closed, and consider a sequence $t_m \to t$ where $t_m \in T'_l \; \forall m$. Since f is continuous and $(S \times [0,1]) \times T'_{l-1}$ is closed by assumption, $\limsup_m f(t_m)((S \times [0,1]) \times T'_{l_1}) = f(t)((S \times [0,1]) \times T'_l)$ (this is criterion (iii) of the Portmanteau Theorem in Patrick. Billingsley (1999)). But by assumption $(f(t_m)((S \times [0,1]) \times T'_{l-1}) = 1 \; \forall m)$ so $f(t)((S \times [0,1]) \times T'_l) = 1$ i.e. $t \in T'_l$. So T'_l is closed [10]. □

lemma (5.2) is *lemma A.1* in Brandenburger and Dekel (1985) adapted to the context of this paper. The fact that T'_l is a Borel set also helps in seeing why T^* is well defined. As argued in Hammond (2004) the Borel set T^* where there is common knowledge of coherency can be expressed as an the limiting infinite intersection $T^* = \cap_{l=1}^{\infty} T'_l$ of shrinking sequence of cylindrical set $T'_l := \text{proj}_{T'_l} T^* \times \prod_{j=l+1}^{\infty} \Delta(T'_j)$. Here $\text{proj}_{T'_l} T^*$ is the projection of T^* onto its first l coordinates(or layers of knowledge) and $\prod_{j=l+1}^{\infty} \Delta(T'_j)$ allows for any beliefs deeper than l to be arbitrary. Thus in T'_1, one essentially fixes the first coordinate, and allows the rest to vary. Therefore T'_l, is a shrinking sequence of cylinder sets which implies the sequences of measures $\pi_k(\text{proj}_{T'_l} T^*)$ is non increasing and bounded below, with a well defined limit. Thus, $T^* = \{t \in T'_1 : f(t)(S \times [0,1] \times T^*) = 1\}$ and is well defined.

The following proposition is the analogue of the proposition of Brandenburger and Dekel (1993) in this extended framework, which shows that the space T^* closes the model and is therefore the universal type space

Proposition 2. *There is a homeomorphism $g : T^* \mapsto \Delta((S \times [0,1]) \times T^*)$*

[10] The only additional aspect of the proof was [0,1] which is a closed set.

Proof. The proof trivially follows from Brandenburger and Dekel (1993) and is stated for completeness.

$T^* = \{t \in T_1' : f(t)(S \times [0,1] \times T^*) = 1\}$; or

$f(T^*) = \{\pi \in \Delta((S \times [0,1]) \times T_0') | \pi((S \times [0,1]) \times T^*) = 1\}$ as f is onto

The set on the right hand side $\{\pi \in \Delta((S\times[0,1])\times T_0')|\pi((S\times[0,1])\times T^*) = 1\}$ is homeomorphic to $\Delta((S\times[0,1])\times T^*)$ as for any metric space Z and measurable subset W of Z, $\{\pi \in \Delta(Z)|\pi(W) = 1\}$ is homeomorphic to $\Delta(W)$. Similarly, on left hand side, $f(T^*)$ is homeomorphic to T^* as f is a homeomorphsim from $T_1' \to \Delta(([0,1] \times S) \times T_0'))$ and since $T^* \subset T_1' \subset T_0'$ thus homeomorphic as well. Thus it follows that T^* is homeomorphic to $\Delta((S \times [0,1]) \times T^*)$. □

Remark 1. *When S is a separable metric space only proposition (1) holds.*

The remark follows as in the proof of (5.2) the main property is that $((S \times [0,1]) \times T_{l-1})$ is closed, this follows as S is Polish space i.e. complete, separable and metrizable, where completeness requires that any Cauchy sequence must converge to a point in the space. Since every convergent sequence is Cauchy, this means the space contains all its limit points and so is closed. $[0,1]$ is also closed similarly. If S instead is assumed just to be a separable metrizable space, then the closure of S is not immediate without other assumptions. Thus when S as a separable metrizable space proposition (1) holds, but proposition (2) is not immediate without further assumptions.

Proposition (2) completes the closure of model of beliefs in games with infinite players, demonstrating, a player given her type is able to form beliefs about the labels, states of nature and types of other players. It is to be noted, that the very special finite type-space, common in economics, implies that after a certain k^{th} step in the hierarchy construction, there is no additional information being added. As in the finite type-space case, if the infinite hierarchy of beliefs results in a finite set, only finitely many non trivial expansions are possible.

6 Applications

There are many games in economics with a continuum of players, examples include William S. Vickrey (1945) and Mirrlees (1971) and Aumann (1964), Aumann (1966), and Hildenbrand (1974) where a continuum of economic agents are labelled by the points of the unit interval, equipped with its Lebesgue measure, a framework standard in a spectrum of papers in macroeconomics. There is the series of works in global games culminated in Morris and Shin (2001), which served as the motivation for this paper and being considered as the prime example in this section. Morris and Shin (2001) is the prime example where in the limit, a continuum of (potential) agents represented by the Lebesgue unit interval, as a canonical non-atomic measure.

Morris and Shin (2001) take the set of agents to be a continuum represented by the Lebesgue unit interval. It is implicit in Morris and Shin (2001) that the continuum is of potential agents rather than actual, and thus the construction provided in the paper, covers the global games framework and provides justification to considering higher-order beliefs in games with many players. The implicit realisation that Morris and Shin (2001) refer to a continuum of potential agents, follows from the following issue and proposition as given by Hammond (2023) [11].

Definition 6.1. *Hammond and Sun (2008) A random process with a continuum of random variables is a mapping $L \times \Omega \ni (t,\omega) \mapsto g_t(\omega) \in \mathbb{R}$ defined on the Cartesian product of the Lebesgue unit interval $(L, \mathcal{L}, \lambda)$ with the probability space $(\Omega, \mathcal{F}, P)$, and having the property that for each Borel set $B \in \mathcal{B}(\mathbb{R})$:*

1. *For each $t \in L$, the set $g_t^{-1}(B) := \{\omega \in \Omega \mid g_t(\omega) \in B\}$ is $\mathcal{F}$-measurable. In other words, $\omega \mapsto g_t(\omega)$ is a random variable on $(\Omega, \mathcal{F}, P)$.*
2. *The mapping $L \ni t \mapsto (P \circ g_t^{-1})(B) = P(g_t^{-1}(B))$ is $\mathcal{L}$-measurable.*

Claim: Consider a continuum of a random variables determined by the process $g : L \times \Omega \mapsto \mathbb{R}$. Suppose that the random variables are independent and identically distributed, with mean $m := \int_\Omega g_t(\omega)P(d\omega), \forall t \in L$ then $\int_L g_t(\omega)P(d\omega) \underset{P-a.s.}{=} m$, by the law of large numbers.

Hammond (2023) shows the above claim is false, unless the independent process is essentially deterministic or the integral is interpreted as a Monte-Carlo integral.[12]

Intuitively, you have a process which generates a random number, at each point t in the Lebesgue unit interval, thus an uncountable set of random numbers. In a finite and a countably infinite setting, calculation of the average, is standard or obtained with the law of large numbers. However, it does not readily extend in consideration of averages for an uncountable set of numbers. The following proposition from Hammond (2023) presents this formally.

Proposition 3. *Given an i.i.d. process, $L \times \Omega \ni (t,\omega) \mapsto g_t(t,\omega)$ except when it is essentially deterministic, the sample path $t \mapsto g(t,\omega)$ is $P-a.s.$ non measurable and so $\int_L g_t(\omega)P(d\omega)$ does not exist as a Lebesgue Integral.*

Proof. Theorem 1 in Hammond (2023) □

When applying this to the global games framework of Morris and Shin (2001), the agents are $t \in [0,1]$ with the Lebesgue measure and there is an underlying fundamental(θ) which is drawn from a common prior(uniformly distributed on $\mathbb{R}$ or from a continuously differentiable strictly positive density on $\mathbb{R}$). The agents observe a signal, as θ is observed with noise ϵ_i which is i.i.d. normal. This means an agent observing signal y, considers θ to be normally distributed with mean y and standard deviation of the noise Morris and Shin (2001). The realisations of the signals are independent, conditional on θ Morris and Shin (2001). There is

[11] These results were first presented by Prof. Peter Hammond at SWET, 2018, and the author is thankful to Prof. Peter Hammond for providing a copy of his presentation and the subsequent working paper Hammond (2023).

[12] The process $L \times \Omega \ni (t,\omega) \mapsto g_t(t,\omega)$ is essentially deterministic in case there is non random measurable path $L \ni t \mapsto h(t)$ such that the sample path satisfies $g(t,\omega) \underset{P-a.s.}{=} h(t) \forall t \in L$

a binary action set $a \in \{a_0, a_1\}$ and the choice of the agent depends on their signal and beliefs about the proportion of other agents choosing the action. Relating to the setup of the proposition (3), Ω represents the state space from which the fundamental θ is drawn. The random process $g_t(t,\omega)$, is a function which maps each agent and a realisation ω (of θ) to an action. Thus, $g_t(t,\omega)$ would indicate whether each agent chooses a_0 or a_1 based on their realisation ω. From proposition (3), unless the mapping $t \mapsto g(t,\omega)$ is essentially deterministic—θ is perfectly observable, no noise and each agent observes the same signal— the sample path $t \mapsto g(t,\omega)$ is almost surely non-measurable and thus $\int_L g_t(\omega)P(d\omega)$ does not exist as a Lebesgue Integral. The non existence of $\int_L g_t(\omega)P(d\omega)$ implies, in particular when looking at the average action or the proportion of agents choosing a particular action, is not well defined. Therefore, statements claiming that the best response of an agent, is to choose an action a_0 if it is majoritarian is not well defined. Indeed, the payoffs in the game are also not well defined. Thus, when referring to a continuum of agents in Morris and Shin (2001) it is implicit that there are a potential continuum of agents, from which a countably infinite set of actual agents are randomly selected, whose distribution in the limit converges almost surely to the canonical non-atomic measure of the continuum of (potential) agents represented by the Lebesgue unit interval. With a countably infinite set of actual agents, the claim above can be applied as the law of large numbers can be invoked given the countable nature. Note, $\int_L g_t(\omega)P(d\omega)$ can be calculated by considering its Monte-Carlo Integral as in Hammond and Sun (2008). The discussion of the Monte-Carlo integral is not presented as it is beyond the scope of this paper.

7 Conclusion

The paper provides a formal proof for Harsanyi (1967) notion of a type in games of incomplete information with infinite players, extending from the work of Mertens and Zamir (1985) and Brandenburger and Dekel (1993). A novel extension by considering a continuum of potential players is provided. Suitable conditions are defined for a type to be coherent, and it is shown as in Brandenburger and Dekel (1993) a coherent type with common knowledge of coherency is able to form higher order beliefs over the types of other players, thus closing the model of beliefs. The results of this paper allow the use and provide the framework for considering higher order beliefs in games with many players. The mathematical tractability offered by Polish spaces is justified by considering robustness of the results with separable metric spaces. An informal extension for future work is provided for the general measurable case, to extend the results of Heifetz and Samet (1998) for games with infinite players.

8 References

Aumann, Robert J. (Jan. 1964). "Markets with a Continuum of Traders". In: *Econometrica* 32.1/2, p. 39. ISSN: 00129682. URL: https://www.jstor.org/stable/1913732.

— (Jan. 1966). "Existence of Competitive Equilibria in Markets with a Continuum of Traders". In: *Econometrica* 34.1, p. 1. ISSN: 00129682. URL: https://www.jstor.org/stable/1909854.

— (Aug. 1999a). "Interactive epistemology I: Knowledge". In: *International Journal of Game Theory* 28.3, pp. 263–300. ISSN: 0020-7276. URL: https://rdcu.be/djP9V.

— (Aug. 1999b). "Interactive epistemology II: Probability". In: *International Journal of Game Theory* 28.3, pp. 301–314. ISSN: 0020-7276. URL: https://rdcu.be/djP9W.

Bertsekas, Dimitri and Steven E Shreve (1996). *Stochastic optimal control: the discrete-time case*. Vol. 5. Athena Scientific.

Billingsley, Patrick (2012). *Probability and Measure*. New York, United States: John Wiley & Sons, Incorporated. ISBN: 9781118341896. URL: http://ebookcentral.proquest.com/lib/warw/detail.action?docID=836625.

— (1999). *Convergence of probability measures*. Wiley, p. 277. ISBN: 0471197459.

Brandenburger, Adam and Eddie Dekel (1985). "Hierarchies of Beliefs and Common Knowledge". URL: https://www.gsb.stanford.edu/faculty-research/working-papers/hierarchies-beliefs-common-knowledge.

— (Feb. 1993). "Hierarchies of Beliefs and Common Knowledge". In: *Journal of Economic Theory* 59.1, pp. 189–198. ISSN: 00220531. URL: https://doi.org/10.1006/jeth.1993.1012.

Cohn, Donald L. (2013). *Measure Theory*. New York, NY: Springer New York, Ch 8–Ch 10. ISBN: 978-1-4614-6955-1. URL: https://link.springer.com/book/10.1007/978-1-4614-6956-8.

Conway, John B. (2007). *A Course in Functional Analysis*. Vol. 96. New York, NY: Springer New York. ISBN: 978-1-4419-3092-7. URL: https://link.springer.com/book/10.1007/978-1-4757-4383-8.

Dellacherie, Claude and Paul-André Meyer (1978). "Chapter III Complements to measure theory". In: *Probabilities and Potential A*. Vol. 29, pp. 39–81. URL: https://www.sciencedirect.com/bookseries/north-holland-mathematics-studies/vol/29/suppl/C.

Hammond, Peter (2004). "Expected Utility in Non-Cooperative Game Theory". In: *Handbook of Utility Theory: Volume 2 Extensions*. Boston, MA: Springer US, pp. 979–1063. ISBN: 978-1-4020-7964-1. URL: https://doi.org/10.1007/978-1-4020-7964-1_5.

— (Jan. 2017). "Designing a strategyproof spot market mechanism with many traders: twenty-two steps to Walrasian equilibrium". In: *Economic Theory* 63.1, pp. 1–50. ISSN: 0938-2259. URL: https://rdcu.be/djQcg.

— (2023). "The Monte Carlo Integral of a Continuum of Independent Random Variables". In: *The Warwick Economics Research Paper Series*. The Warwick Economics Research Paper Series (TWERPS) 1479. URL: https://ideas.repec.org/p/wrk/warwec/1479.html.

Hammond, Peter and Yeneng Sun (2008). “Monte Carlo Simulation of Macroeconomic Risk with a Continuum of Agents: The General Case”. In: *Economic Theory* 36.2, pp. 303–325. ISSN: 09382259, 14320479. URL: http://www.jstor.org/stable/40282899.

Harsanyi, John C. (Nov. 1967). “Games with Incomplete Information Played by “Bayesian” Players, Parts I–III.” In: *Management Science* 14.3, pp. 159–182. ISSN: 0025-1909. DOI: 10.1287/mnsc.14.3.159.

Heifetz, Aviad and Dov Samet (Oct. 1998). “Topology-Free Typology of Beliefs”. In: *Journal of Economic Theory* 82.2, pp. 324–341. ISSN: 00220531. URL: https://doi.org/10.1006/jeth.1998.2435.

— (Dec. 1999). “Coherent beliefs are not always types”. In: *Journal of Mathematical Economics* 32.4, pp. 475–488. ISSN: 03044068. URL: https://doi.org/10.1016/S0304-4068(98)00066-4.

Hildenbrand, Werner (1974). *Core and Equilibria of a Large Economy*. Princeton University Press, pp. vii–viii. ISBN: null. URL: http://www.jstor.org/stable/j.ctt13x0tjf.3.

Maschler, Michael, Eilon Solan, and Shmuel Zamir (Mar. 2013). *Game Theory*. Cambridge University Press. ISBN: 9780511794216. URL: https://www.cambridge.org/core/product/identifier/9780511794216/type/book.

Mertens, Jean -François and Shmuel Zamir (Mar. 1985). “Formulation of Bayesian analysis for games with incomplete information”. In: *International Journal of Game Theory* 14.1, pp. 1–29. ISSN: 0020-7276. URL: https://rdcu.be/djQdr.

Mirrlees, J. A. (Apr. 1971). “An Exploration in the Theory of Optimum Income Taxation”. In: *The Review of Economic Studies* 38.2, p. 175. ISSN: 00346527. URL: https://www.jstor.org/stable/2296779.

Morris, Stephen and Hyun Song Shin (2001). *Global Games: Theory and Applications.* Tech. rep. URL: https://cowles.yale.edu/node/145344.

Parthasarathy, K R (1967). *Probability Measures on Metric Spaces*. New York. URL: https://www.sciencedirect.com/book/9781483200224/probability-measures-on-metric-spaces.

Varadarajan, V S (1958). “On the Convergence of Sample Probability Distributions”. In: *Sankhyā: The Indian Journal of Statistics (1933-1960)* 19.1/2, pp. 23–26. ISSN: 00364452. URL: http://www.jstor.org/stable/25048365.

William S. Vickrey (1945). “Measuring Marginal Utility by Reactions to Risks”. In: *Econometrica* 13, pp. 319–333. URL: https://www.jstor.org/stable/1906925.