Common Priors and Uncommon Priors

Thesis submitted for the degree
Doctor of Philosophy

by

Ziv Hellman

Submitted to the Senate of the Hebrew University of Jerusalem
September 2012

This work was carried out under the supervision of

Professor Dov Samet
and
Professor Sergiu Hart

I would like to extend my gratitude to Professor Dubi Samet, a true *maître à penser*, for the confidence he expressed in me from the beginning, for his great patience with me, his enthusiastic and creative attitude and for his invaluable assistance.

Thanks are due to Professor Sergiu Hart for an abundance of encouragement, excellent explanations of many concepts and generous financial support.

I also thank my mother Gila, and my colleagues Ilan Nehama and Yehuda (John) Levy.

The research here was supported in part by the European Research Council under the European Commission's Seventh Framework Programme (FP7 – 20072013) ERC grant agreement no. 249159.

Abstract

Several topics related to the common prior assumption, common knowledge, common belief and agreement are studied. We first consider how common common priors are, in the sense that we ask what happens when we vary agents beliefs against the background of a fixed knowledge space. Are common priors robust to such perturbations? The answer depends on what we term the tightness of the partition profile. In a tight space a common prior is guaranteed no matter what the posterior beliefs are, but in a non-tight space the set of consistent profiles is nowhere dense. Significantly, these results depend only the structure of the partition spaces, not on any particular posteriors.

We then consider what happens when priors are not common. Can we measure how far a space is from having a common prior, and then use that to find an upper bound on common knowledge disagreements? We show that it is possible to approximate agreement theorems with an appropriate measurement of how far a knowledge space is from having a common prior. Using results in the first essay, we show that based on the partition structure alone one can obtain an upper bound to the number of proper refinements needed to arrive at a common prior, regardless of how far the players are initially from a common prior.

Characterisations of common priors, including the no betting theorem and the iterated expectations characterisation, are studied in the context of infinite state spaces, both countable and uncountable.

Finally, we include an essay that does not directly involve common priors but resolves a question that has been open for a decade. Simon (2003) presented an example of a three-player Bayesian game with no Bayesian equilibrium. That important result left in its wake (at least) two open questions: (1) are there examples of games that have no Bayesian ε-equilibria?; (2) are there examples of two-player games that have no Bayesian equilibria? We show here that the answer to both questions is yes by constructing a two-player Bayesian game with no Bayesian ε-equilibria. As a side-benefit, the example also shows that there there exist strategic-form games with an uncountable number of players and no Nash ε-equilibria and that there exist two-player Bayesian games with no Harsányi equilibria (meaning *ex ante* Nash equilibria over the common prior of a Bayesian game), which had also been open questions.

Contents

Introduction

Ever since the introduction of games with incomplete information by Harsányi in 1967, the assumption that players posterior beliefs in models of differential information are derived from a common prior has been ubiquitous in the literature. It plays an essential role in the no agreements theorem of Aumann (1976) and in the no trade theorems that followed.

It is also a basic building block of the solution concept of correlated equilibrium which was interpreted by Aumann (1987) as the expression of common knowledge of rationality. As pointed out in that paper, the assumption of a common prior, also known as the Harsányi doctrine, is pervasively 'explicit or implicit in the vast majority of the differential information literature in economics and game theory'. It is commonly assumed in a great many models of rational expectations, securities trading, bargaining, auctions, repeated games, Bayesian games, signalling, principal-agent, moral hazard and bankruptcy, to name only a few. Despite its pervasiveness, the justification and the use of the common prior assumption was, and still is, debated and challenged.

The special interest in the common prior assumption leads naturally to many questions. In this work, some of those questions are studied in detail. These include:

(1) How restrictive an assumption is the common prior assumption? It has long been known that the set of type spaces with common priors is measure theoretically 'small'. In Chapter 1 we look at this question from a topological perspective, and find that surprisingly the answer in this case depends critically on the structure of the information partitions, with a 'phase change' in the topological size of the set of common prior spaces occurring when the total number of atoms passes a certain threshold.

(2) What happens if the common prior assumption is not assumed? Can one measure how far a type space is from having a common prior and use that to derive bounds on disagreements? We show in Chapter 2 that the answer is yes.

(3) There are at least two main characterisations of the common prior assumption in the literature. These were first proved in finite state spaces. What happens in infinite spaces?

Chapter Summaries

Each chapter is self-contained, in the sense that each chapter introduces the definitions used in it, numbers sections and theorems relative to the sections and theorems in the same chapter (rather than the other chapters) and contains its own bibliography.

Chapter 1

In Chapter 1, we ask 'how common are common priors'. To answer that question, we vary agents beliefs against the background of a fixed knowledge space, that is, a state space with a partition for each agent. Beliefs are the posterior probabilities of agents, which we call type profiles.

We then ask what is the topological size of the set of consistent type profiles, those that are derived from a common prior (or a common improper prior in the case of an infinite state space). The answer depends on what we term the tightness of the partition profile. A partition profile is tight if in some state it is common knowledge that any increase of any single agent's knowledge results in an increase in common knowledge. We show that for partition profiles that are tight the set of consistent type profiles is topologically large, while for partition profiles that are not tight this set is topologically small.

The results in Chapter 1 (joint work with Dov Samet) have been published in the journal Games and Economic Behavior, 74 (2012), pp. 517–525.

Chapter 2

In Chapter 2, we ask what happens when priors are not common?

We introduce a measure for how far a knowledge space is from having a common prior, which we term prior distance. If a knowledge space has δ prior distance, then for any bet f it cannot be common knowledge that each player expects a positive gain of δ times the sup-norm of f, thus extending no betting results under common priors.

Furthermore, as more information is obtained and partitions are refined, the prior distance, and thus the extent of common knowledge disagreement, can only decrease. We derive an upper bound on the number of refinements needed to arrive at a situation in which the knowledge space has a common prior, which depends only on the number of initial partition elements.

The results in Chapter 2 will be published, forthcoming in the International Journal of Game Theory.

Chapter 3

In Chapter 3, we extend to infinite state spaces that are compact metric spaces a result previously attained by Dov Samet solely in the context of finite state spaces. A necessary and sufficient condition for the existence of a common prior for several players is given in terms of the players' present beliefs only.

A common prior exists if and only if for each random variable it is common knowledge that all Cesàro means of iterated expectations with respect to any permutation converge to the same value; this value is its expectation with respect to the common prior. It is further shown that compactness is a necessary condition for some of the results.

The results in Chapter 3 have been published in the journal Games and Economic Behavior, 72 (2011), 163–171.

Chapter 4

In Chapter 4 we show that the no betting characterisation of the existence of common priors over finite type spaces extends only partially to improper priors in the countably infinite state space context: the existence of a common prior implies the absence of a bounded agreeable bet, and the absence of a common improper prior implies the existence of a bounded agreeable bet.

However, a type space that lacks a common prior but has a common improper prior may or may not have a bounded agreeable bet. The iterated expectations characterisation of the existence of common priors extends almost as is, as a sufficient and necessary condition, from finite spaces to countable spaces, but fails to serve as a characterisation of common improper priors. As a side-benefit of the proofs here, we also obtain a constructive proof of the no betting characterisation in finite spaces.

Chapter 5

In Chapter 5, we resolve a question that has been open for several years: are there are games with no Bayesian approximate equilibria? We present an example of a Bayesian game with two players, two actions, a common prior and an uncountable cardinality of states that possesses no Bayesian approximate equilibria.

As a side benefit we also have for the first time an an example of a 2-player Bayesian game with no Bayesian equilibria and an example of a strategic-form game with no approximate Nash equilibria. The construction makes use of techniques developed in an example by Y. Levy of a discounted stochastic game with no stationary equilibria.

CHAPTER 1

How Common are Common Priors?

Published in Games and Economic Behavior, 74 (2012), pp. 517–525.
The paper appears here on pages 6 – 14.

Games and Economic Behavior 74 (2012) 517–525

Contents lists available at SciVerse ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

How common are common priors? ☆

Ziv Hellman [a,*], Dov Samet [b]

[a] The Department of Mathematics and The Centre for the Study of Rationality, The Hebrew University of Jerusalem, Jerusalem, Israel
[b] Faculty of Management – The Leon Recanati Graduate School of Business Administration, Tel Aviv University, Tel Aviv, Israel

ARTICLE INFO

Article history:
Received 30 November 2007
Available online 31 August 2011

JEL classification:
D82
D83

Keywords:
Common prior
Common knowledge
Knowledge
Belief

ABSTRACT

To answer the question in the title we vary agents' beliefs against the background of a fixed knowledge space, that is, a state space with a partition for each agent. Beliefs are the posterior probabilities of agents, which we call type profiles. We then ask what is the topological size of the set of *consistent* type profiles, those that are derived from a common prior (or a common improper prior in the case of an infinite state space). The answer depends on what we term the tightness of the partition profile. A partition profile is *tight* if in some state it is common knowledge that any increase of any single agent's knowledge results in an increase in common knowledge. We show that for partition profiles that are tight the set of consistent type profiles is topologically large, while for partition profiles that are not tight this set is topologically small.

1. Introduction

Ever since the introduction of games with incomplete information by Harsanyi (1967–1968), the assumption that players' posterior beliefs in models of differential information are derived from a common prior has been ubiquitous in the literature. It plays an essential role in the no agreements theorem of Aumann (1976) and in the no trade theorems that followed. It is also a basic building block of the solution concept of correlated equilibrium which was interpreted by Aumann (1987) as the expression of common knowledge of rationality. As pointed out in that paper, the assumption of a common prior, also known as the Harsanyi doctrine, is pervasively "explicit or implicit in the vast majority of the differential information literature in economics and game theory". Despite its pervasiveness, the justification and the use of the common prior assumption was, and still is, debated and challenged (see Gul, 1998 and Aumann, 1998).

The special interest in the common prior assumption leads naturally to the question how restrictive an assumption it is, or equivalently, how common common priors are. We study this question in a general model of differential information that has two parts, a *knowledge space* and the agents' *posterior beliefs*. The first is given by a finite or countably infinite state space with a partition profile of the state space, one for each agent, which define the agents' knowledge. An agent's posterior beliefs are given by a *type function* which assigns to each element in the agent's partition a probability function on this element. A type profile—one type function for each agent—is *consistent* if all the type functions are derived from one probability on the state space—the *common prior*—by conditioning on the partitions' elements. In the countably infinite state space, we consider a type profile to be consistent if it can be derived from a *common improper prior* via conditioning.

☆ The research of the first author has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement No. 249159. The research of the second author was partially financed by the Henry Crown Institute of Business Research in Israel.

* Corresponding author.
E-mail addresses: ziv.hellman@mail.huji.ac.il (Z. Hellman), dovsamet@gmail.com (D. Samet).

0899-8256/$ – see front matter
doi:10.1016/j.geb.2011.08.008

Against the background of a fixed knowledge space, we vary the type profiles and study the topological size of the set of consistent type profiles. As we show, this size depends on the partition profile of the knowledge space, through its common knowledge structure. We say that knowledge is *tight* at a state when any increase of the agents' knowledge in this state results in increasing common knowledge. We say that the partition profile is *tight* if at some state it is common knowledge that knowledge is tight. We show that when the partition profile is tight the set of consistent type profiles is topologically large, and when it is not tight this set is topologically small. The characterization of tightness in the finite case is particularly simple: the tightness of a type profile, the meet of which consists of one element, can be determined solely by the total number of elements in the partition profile. In the proofs we use another characterization of tightness, which holds for both the finite and the infinite case, in terms of chains which are defined in turn in terms of the accessibility relation on states. The existence of a prior for a given type profile can be characterized by a condition on the relation between chains and the type profile. Such a condition was used in Harsanyi (1967–1968) for Harsanyi type spaces, and was extended by Rodrigues-Neto (2009) for general knowledge spaces.

The precise meaning of large and small depends on whether the state space is finite or countably infinite. For finite knowledge spaces, when the partition profile is tight *each* type profile is consistent, and when it is not tight, the set of consistent type profiles is nowhere dense. For countably infinite knowledge spaces, we endow the set of all type profiles with a topology that makes it a Baire space. When the partition profile is tight the set of consistent type profiles is big, as its complement is of first category, which in a Baire space is a small set. When it is not tight, the set of consistent type profiles is small, being of first category.

2. Preliminaries

2.1. Knowledge spaces

A knowledge space for a nonempty finite set of **agents** I, is a couple $(\Omega, \boldsymbol{\Pi})$, where Ω is a nonempty set called a **state space**, and $\boldsymbol{\Pi} = (\Pi_i)_{i\in I}$ is a **partition profile**, where for each i, Π_i is a partition of Ω. The knowledge space is called finite or countably infinite when Ω is finite of countably infinite, correspondingly. An **event** is a subset of Ω. For a partition Π of Ω and a state ω, $\Pi(\omega)$ is the element of Π that contains ω. We say that agent i **knows** an event E at ω if $\Pi_i(\omega) \subseteq E$. We define for each i a **knowledge operator** $K_i : 2^\Omega \to 2^\Omega$, by $K_i(E) = \{\omega \mid \Pi_i(\omega) \subseteq E\}$. Thus, $K_i(E)$ is the event that i knows E.

For a pair of partitions Π and Π' and state ω, we write $\Pi' \succcurlyeq_\omega \Pi$ when $\Pi'(\omega) \subseteq \Pi(\omega)$. For the partition profiles $\boldsymbol{\Pi}$ and $\boldsymbol{\Pi}'$, $\boldsymbol{\Pi}' \succcurlyeq_\omega \boldsymbol{\Pi}$ means that for each i, $\Pi'_i \succcurlyeq_\omega \Pi_i$. The partition Π' is a **refinement** of Π, denoted $\Pi' \succcurlyeq \Pi$, when $\Pi' \succcurlyeq_\omega \Pi$ for each state ω. The partition profile $\boldsymbol{\Pi}'$ is a **refinement** of $\boldsymbol{\Pi}$, denoted $\boldsymbol{\Pi}' \succcurlyeq \boldsymbol{\Pi}$, if for each i, $\Pi'_i \succcurlyeq \Pi_i$. For each of these four relations, a corresponding relation with $\succ$ instead of $\succcurlyeq$ is obtained by discarding the reflexive part of the relation $\succcurlyeq$. The two irreflexive relations describe an increase of knowledge, while the two reflexive relations describe a weak increase of knowledge. Thus, for example, if $\Pi' \succ_\omega \Pi$, and K and K' are the knowledge operators associated with Π and Π' respectively, then for each event E, if $\omega \in K(E)$, then $\omega \in K'(E)$, but for some events, for example $E = \Pi'(\omega)$, $\omega \in K'(E)$ but $\omega \notin K(E)$.

The **meet** of $\boldsymbol{\Pi}$, denoted $\wedge\boldsymbol{\Pi}$, is the partition which is the finest among all the partitions Π that satisfy $\Pi \succcurlyeq \Pi_i$ for each i. The knowledge operator K_c defined by the meet partition is called the **common knowledge** operator (Aumann, 1976). It can be described in terms of the knowledge operator K_i as follows. Denote by $K(E)$ the event that all agents know E. That is, $K(E) = \bigcap_{i\in I} K_i(E)$. Then $K_c(E) = \bigcap_{n=1}^{\infty} K^n(E)$. For $M \in \wedge\boldsymbol{\Pi}$, the elements of Π_i contained in M form a partition of M. Thus, $(M, \boldsymbol{\Pi}_M)$, where $\boldsymbol{\Pi}_M$ is the restriction of $\boldsymbol{\Pi}$ to M, is a knowledge space.

2.2. Beliefs

The beliefs of an agent in a given state are described by a probability distribution over the state space. These beliefs are related to the agent's knowledge as follows. Denote by $\Delta(\Omega)$ the set of all probability functions on Ω. A **type function** for Π_i is a function $t_i : \Omega \times \Omega \to \mathbb{R}$ that satisfies:

(a) for each ω, $t_i(\omega, \cdot) \in \Delta(\Omega)$,
(b) for each i and $\pi \in \Pi_i$, if $\{\omega, \omega'\} \subseteq \pi$, then $t_i(\omega', \cdot) = t_i(\omega, \cdot)$,
(c) for each i, $\pi \in \Pi_i$, and $\omega \in \pi$, the support of $t_i(\omega, \cdot)$ is π, i.e., $t_i(\omega, \pi) = 1$.

We say that $t_i(\omega, \cdot)$ is i's **type** at ω. By condition (b), the type of i is measurable with respect to Π_i, i.e., the type of i is the same in all states in π which means that i knows her type, or equivalently, knows her beliefs. In light of (b) we sometimes write for i and $\pi \in \Pi_i$, $t_i(\pi, \cdot)$ for the type of i in all the states in π. Condition (c) implies that whenever i knows E at ω she assigns probability 1 to it, i.e., whatever she knows she is certain of.[1]

A **type profile** for $\boldsymbol{\Pi}$ is a vector of type functions, $\mathbf{t} = (t_i)_{i\in I}$, where for each i, t_i is a type function for Π_i. Denote by $\mathcal{T}(\boldsymbol{\Pi})$ the set of all type profiles for $\boldsymbol{\Pi}$. A type profile assigns for each i and ω an element $t_i(\omega, \cdot)$ in $\Delta(\Omega)$. Thus, we may

[1] Conditions (b) and (c) are part of the definition of the space of knowledge and belief in Aumann (1976). The meaning given to them here are expressed as two axioms on the relation between knowledge and belief in Hintikka (1962).

consider $\mathcal{T}(\boldsymbol{\Pi})$ as a subset of $\Delta(\Omega)^{\Omega\times I}$. In particular, for a finite state space we consider $\mathcal{T}(\boldsymbol{\Pi})$ as a topological space with the topology induced by the standard topology of the Euclidean space in which $\mathcal{T}(\boldsymbol{\Pi})$ is embedded.

A **prior** for a type function t_i is a probability function $p \in \Delta(\Omega)$ such that for each $\pi \in \Pi_i$, $p(\pi)t_i(\pi,\omega) = p(\omega)$ for all $\omega \in \pi$. A **common prior** (cp) for the type profile $\mathbf{t}$ is a probability function $p \in \Delta(\Omega)$ which is a prior for each agent i.[2] A type profile $\mathbf{t}$ is **consistent** when it has a common prior.

The model of knowledge space with beliefs used here is the same as the model in Aumann (1976), except that in the latter the assumption is made that there exists a common prior. Our model is also the discrete case of the abstract S-based belief space in Mertens and Zamir (1985), where S is Ω. Although knowledge is not introduced explicitly in their work, the partitions of the space into agents' types makes it a partition model.

3. Main results

3.1. Tightness

The commonness of consistent type profiles for a given knowledge space depends on a property of the knowledge space we call tightness. In both the finite and the infinite case, when the partition profile is not tight, the set of consistent type profiles is topologically small. When the partition profile is tight this set is topologically large.

We say that knowledge is tight at a state if increasing any agents' knowledge at this state must result in increasing common knowledge. A partition profile is tight if at some state there is common knowledge that knowledge is tight. Formally,

Definition 1. For a partition profile $\boldsymbol{\Pi}$, knowledge is **tight at** ω, when for each $\boldsymbol{\Pi}' \succ \boldsymbol{\Pi}$, if $\boldsymbol{\Pi}' \succ_\omega \boldsymbol{\Pi}$ then $\wedge\boldsymbol{\Pi}' \succ \wedge\boldsymbol{\Pi}$. Let T be the event that knowledge is tight. We say that $\boldsymbol{\Pi}$ is **tight**, if $K_c(T) \neq \emptyset$.

In the following example we illustrate the notions of tight knowledge and tight partition profiles.

Example 1. Let $\Omega = \{\omega_1,\omega_2,\omega_3,\omega_4\}$ and $I = \{1,2\}$. Consider the partition profile $\boldsymbol{\Pi} = (\Pi_1,\Pi_2)$, where $\Pi_1 = \{\{\omega_1,\omega_2\},\{\omega_3,\omega_4\}\}$ and $\Pi_2 = \{\{\omega_1,\omega_2,\omega_3\},\{\omega_4\}\}$. Obviously, $\wedge\boldsymbol{\Pi} = \{\Omega\}$. Suppose that $\boldsymbol{\Pi}' \succ \boldsymbol{\Pi}$, and $\boldsymbol{\Pi}' \succ_{\omega_4} \boldsymbol{\Pi}$. The only way that knowledge can increase at ω_4 is by splitting the partition element $\{\omega_3,\omega_4\}$. Thus, $\Pi'_1(\omega_4) = \{\omega_4\}$. Therefore $\{\omega_4\} \in \wedge\boldsymbol{\Pi}'$ which means that $\wedge\boldsymbol{\Pi}' \succ \wedge\boldsymbol{\Pi}$. We conclude that knowledge at ω_4 is tight. Consider now the partition profile $\boldsymbol{\Pi}'$ where $\Pi'_1 = \Pi_1$ and $\Pi'_2 = \{\{\omega_1\},\{\omega_2,\omega_3\},\{\omega_4\}\}$. Then $\boldsymbol{\Pi}' \succ \boldsymbol{\Pi}$, and $\boldsymbol{\Pi}' \succ_{\omega_3} \boldsymbol{\Pi}$. Yet, $\wedge\boldsymbol{\Pi}' = \wedge\boldsymbol{\Pi}$, which shows that knowledge at ω_3 is not tight. Therefore $T \neq \Omega$, and hence $K_c(T) = \emptyset$. We conclude that $\boldsymbol{\Pi}$ is not tight. It is easy to check that for the last partition profile $\boldsymbol{\Pi}'$ knowledge is tight at each state and therefore $\boldsymbol{\Pi}'$ is tight.

The tightness of a partition profile can be expressed without explicit reference to common knowledge, as follows. We say that $\boldsymbol{\Pi}$ is **connected** when $\wedge\boldsymbol{\Pi} = \{\Omega\}$. For each $M \in \wedge\boldsymbol{\Pi}$, $\boldsymbol{\Pi}_M$, the restriction of $\boldsymbol{\Pi}$ to M, is connected.

Proposition 1.

(a) *A connected partition profile $\boldsymbol{\Pi}$ is tight if and only if for any $\boldsymbol{\Pi}' \succ \boldsymbol{\Pi}$, $\boldsymbol{\Pi}'$ is not connected.*
(b) *A partition profile $\boldsymbol{\Pi}$ is tight if and only if there exists $M \in \wedge\boldsymbol{\Pi}$ such that $\boldsymbol{\Pi}_M$ is tight.*

A third characterization of tightness, in terms of chains, is given in Proposition 5 below. In the finite case, there exists yet another simple characterization of the tightness of a connected type profile, in terms of the total number of partition elements.

Proposition 2. *Let Ω be a finite state space and $\boldsymbol{\Pi}$ a connected partition profile. Then $\sum_{i\in I}|\Pi_i| \leqslant (|I|-1)|\Omega|+1$ and equality holds if and only if $\boldsymbol{\Pi}$ is tight.*

Observe, that the dimension of the set of types $\mathcal{T}(\boldsymbol{\Pi})$ is $\sum_{i\in I}(|\Omega|-|\Pi_i|) = |I||\Omega| - \sum_{i\in I}|\Pi_i|$ and the dimension of the set of priors $\Delta(\Omega)$ is $|\Omega|-1$. Thus, Proposition 2 characterizes the connected tight partition profiles $\boldsymbol{\Pi}$ as the ones with minimal dimension of $\mathcal{T}(\boldsymbol{\Pi})$ which equals the dimension of $\Delta(\Omega)$.[3]

[2] Contrasting a prior for t_i with the types $t_i(\omega,\cdot)$, the latter are referred to as the posterior probabilities of i.

[3] This observation suggests a proof for the smallness of the set of consistent type profiles, for partition profiles which are not tight, based on dimensional considerations. We have elected instead to implement an elementary combinatorial proof, which can be applied equally well for the infinite case. Nyarko (2010) states that in a finite Harsanyi type space the set of consistent posteriors has measure zero. The proof requires differential geometry arguments based on dimensionality considerations.

3.2. The size of the set of consistent type profiles

Theorem 1. *Let $(\Omega, \boldsymbol{\Pi})$ be a finite knowledge space.*

1. *If $\boldsymbol{\Pi}$ is tight then each type profile is consistent.*
2. *If $\boldsymbol{\Pi}$ is not tight then the set of consistent type profiles is nowhere dense.*[4]

In order to prove results similar to those of Theorem 1 for countable state spaces, we need to generalize the notion of a common prior. A **common improper prior** (*cip*) for a type profile **t** is a non-negative and non-zero function $p : \Omega \to \mathbb{R}$ such that for each i and $\pi \in \Pi_i$, $p(\pi) < \infty$ and $p(\pi)t_i(\pi, \omega) = p(\omega)$ for all $\omega \in \pi$. Note that although for any $\pi \in \Pi_i$, $p(\pi) < \infty$, the possibility that $p(\Omega) = \infty$ is not ruled out, so that p may not be normalizable. Obviously, a *cp* is in particular a *cip*. Note also that if p is a *cip*, then for any constant $\gamma > 0$, γp is also a *cip*. In particular, if p is a *cip* and $p(\Omega) < \infty$ then $p(\Omega)^{-1}p$ is a common prior. Thus, for a finite space, a profile type has a common prior if and only if it has a common improper prior. In light of this the following definition of consistency for countable spaces generalizes the one given for finite spaces. A type profile **t** is **consistent** when it has a common improper prior and **inconsistent** otherwise.

To measure the topological size of sets in the countable case we use the notion of a set of first category (called also a meager set), namely, a set which is a countable union of nowhere dense sets. A topological space is a Baire space if every set of first category has an empty interior. Therefore, in a Baire space, sets of first category are considered small. We now proceed to define a topology on $\mathcal{T}(\boldsymbol{\Pi})$ for which it is a Baire space.

Consider the complete normed vector space $l^1(\Omega)$ of absolutely summable functions $x : \Omega \to \mathbb{R}$, with the norm $\|x\| = \sum_{\omega\in\Omega} |x(\omega)|$. The set $\Delta(\Omega)$ is closed in $l^1(\Omega)$.[5] Therefore, $\Delta(\Omega)$ with the metric induced on it from $l^1(\Omega)$ is a complete metric space. Hence, the product space $\Delta(\Omega)^{\Omega\times I}$ is a completely metrizable topological space (see Munkres, 1975). Finally, the equalities in the definition of a type guarantee that $\mathcal{T}(\boldsymbol{\Pi})$ is closed in $\Delta(\Omega)^{\Omega\times I}$ and therefore $\mathcal{T}(\boldsymbol{\Pi})$ is a completely metrizable topological space. This implies that $\mathcal{T}(\boldsymbol{\Pi})$ is a Baire space. Obviously, in the finite case the topology just described is the standard topology on finite dimensional Euclidean spaces.

Theorem 2. *Let $(\Omega, \boldsymbol{\Pi})$ be a countable knowledge space.*

1. *If $\boldsymbol{\Pi}$ is tight then the set of inconsistent type profiles is of first category.*
2. *If $\boldsymbol{\Pi}$ is not tight then the set of consistent type profiles is of first category.*

In contrast with the finite case, here the set of inconsistent type profiles of a tight partition profile need not be empty. Example 2 in the next section shows that we cannot even strengthen this part by changing "of first category" to "nowhere dense". Example 3 shows that similar strengthening is also impossible in the second part of the theorem.

3.3. Harsanyi type spaces

Of special interest are Harsanyi type spaces. In such a space $\Omega = \bigtimes_{i\in I} T_i$, where for each i, T_i is a set of types of player i. With each player i we associate the natural partition of Ω, Π_i, into i's types. It is easy to see, using Proposition 1, that the partition profile of a non-trivial Harsanyi type space (one that has more than one state and more than one agent) is connected and not tight. Therefore, if we vary the posterior beliefs of the types on such a finite or countably infinite space, while keeping the sets of types fixed, the set of consistent posterior beliefs is small. The lack of tightness of non-trivial finite Harsanyi type spaces can be also checked using Proposition 2. Obviously, for such a space $|\Pi_i| = |T_i|$, and $|\Omega| = \bigtimes_{i\in I} |\Pi_i|$. It is easy to prove that $\sum_{i\in I} |\Pi_i| < (|I| - 1) \bigtimes_{i\in I} |\Pi_i| + 1 = (|I| - 1)|\Omega| + 1$.

4. Proofs and examples

4.1. Proof of Proposition 1

(a) Let $\boldsymbol{\Pi}$ be a connected partition profile. Then Ω is the only event E such that $K_c(E) \neq \emptyset$. Thus, if $\boldsymbol{\Pi}$ is tight then $T = \Omega$. If $\boldsymbol{\Pi}' \succ \boldsymbol{\Pi}$, then for some ω, $\boldsymbol{\Pi}' \succ_\omega \boldsymbol{\Pi}$, and by tightness, $\wedge\boldsymbol{\Pi}' \succ \wedge\boldsymbol{\Pi}$. Conversely, if $\boldsymbol{\Pi}' \succ \boldsymbol{\Pi}$ implies $\wedge\boldsymbol{\Pi}' \succ \wedge\boldsymbol{\Pi}$, then obviously, each ω is in T, and hence $T = \Omega$, and $K_c(T) = \Omega$.

(b) The partition profile $\boldsymbol{\Pi}$ is tight iff there exists $M \in \wedge\boldsymbol{\Pi}$ such that $M \subseteq T$. It is easy to see that $M \subseteq T$ iff for the knowledge space $(M, \boldsymbol{\Pi}_M)$, knowledge is tight at each $\omega \in M$, which is a necessary and sufficient condition for $\boldsymbol{\Pi}_M$ to be tight. □

[4] A set is nowhere dense if its closure has an empty interior. Such a set is considered topological small.

[5] To see this, consider the linear functional on $l^1(\Omega)$ defined by $f(x) = \sum_{\omega\in\Omega} x(\omega)$. Since $|f(x)| \leqslant \|x\|$, f is continuous. Now, $\Delta(\Omega)$ is the intersection of two closed sets: $f^{-1}(1)$ and the non-negative orthant of $l^1(\Omega)$.

4.2. Chains

We define a **chain** of length $n \geqslant 0$, for the partition profile $\boldsymbol{\Pi}$, from one state to another by induction on n. A state ω_0 is a chain of length 0 from ω_0 to ω_0. A chain of length $n+1$, from ω_0 to ω, is a sequence $c \stackrel{i}{\to} \omega$, where c is a chain of length n from ω_0 to ω', and $\omega \in \Pi_i(\omega')$. Thus, a chain of positive length n is a sequence $c = \omega_0 \stackrel{i_0}{\to} \omega_1 \stackrel{i_1}{\to} \cdots \stackrel{i_{n-1}}{\to} \omega_n$, such that for $s = 0, \ldots, n-1$, $\omega_{s+1} \in \Pi_{i_s}(\omega_s)$.

Obviously, for each i, chains of length 1, $\omega \stackrel{i}{\to} \omega'$, define an equivalence binary relation and Π_i is the partition of Ω into its equivalence classes. We write $\omega \to \omega'$ when there is a chain from ω to ω'. The binary relation $\to$ is the transitive closure of the union of the relations $\stackrel{i}{\to}$, and it is an equivalence relation. We say that ω and ω' are **connected** for $\boldsymbol{\Pi}$, if there is a chain for $\boldsymbol{\Pi}$ from ω to ω'.

Claim 1. *The meet of $\boldsymbol{\Pi}$ is the partition of Ω into the equivalence classes of $\to$.*

To see this, denote by Π_{con} the partition of Ω into equivalence classes of $\to$. Since each of the partitions Π_i is finer than $\wedge\boldsymbol{\Pi}$, it follows by induction on the length of chains that if $\omega \to \omega'$ then $\omega' \in \wedge\boldsymbol{\Pi}(\omega)$. Thus, for each ω, $\Pi_{\text{con}}(\omega) \subseteq \wedge\boldsymbol{\Pi}(\omega)$, i.e., Π_{con} is finer than $\wedge\boldsymbol{\Pi}$. Also, if $\omega' \in \Pi_{\text{con}}(\omega)$ then for all i and $\omega'' \in \Pi_i(\omega')$, $\omega'' \in \Pi_{\text{con}}(\omega)$, i.e., $\Pi_i(\omega') \subseteq \Pi_{\text{con}}(\omega)$. Thus, each of the partitions in Π is finer than Π_{con}. As $\wedge\boldsymbol{\Pi}$ is the finest partition with this property it follows that $\Pi_{\text{con}} = \wedge\boldsymbol{\Pi}$.

Thus, we conclude:

Claim 2. *A partition profile $\boldsymbol{\Pi}$ is connected if and only if every two states are connected.*

We say that a type profile $\mathbf{t}$ is **positive** if for each i, $\pi \in \Pi_i$, and $\omega \in \pi$, $t_i(\pi, \omega) > 0$. Let $\mathbf{t}$ be a positive type profile and (ω_1, ω_2) an ordered pair of states in $\pi \in \Pi_i$. The **type ratio** of (ω_1, ω_2) given i is $\text{tr}_t^i(\omega_1, \omega_2) = t_i(\pi, \omega_2)/t_i(\pi, \omega_1)$. The **type ratio** of a chain $c = \omega_0 \stackrel{i_0}{\to} \omega_1 \stackrel{i_1}{\to} \cdots \stackrel{i_{n-1}}{\to} \omega_n$ of length $n > 0$ is $\text{tr}_{\mathbf{t}}(c) = \bigtimes_{k=0}^{n-1} \text{tr}_{\mathbf{t}}^{i_k}(\omega_k, \omega_{k+1})$. For a chain c of length 0, $\text{tr}_{\mathbf{t}}(c) = 1$. Thus, if $c = c' \stackrel{i}{\to} \omega$ where c' is a chain from ω_0 to ω', $\text{tr}_{\mathbf{t}}(c) = \text{tr}_{\mathbf{t}}(c')\, \text{tr}_{\mathbf{t}}^i(\omega', \omega)$. When we discuss only one type profile we omit the subscript $\mathbf{t}$ in $\text{tr}_{\mathbf{t}}$.

Proposition 3. *If a positive type profile over a connected partition profile has a common improper prior p, then all its common improper priors are of the form γp for some constant $\gamma > 0$. A type profile over a connected partition profile can therefore have at most one common prior.*

Proof. If p is a *cip* for a positive $\mathbf{t}$, then $\text{tr}_{\mathbf{t}}^i(\omega_1, \omega_2) = p(\omega_2)/p(\omega_1)$. Substituting the right-hand side for the left-hand side in the definition of the type ratio of chains, we conclude that for any chain c from ω_0 to ω, $\text{tr}_{\mathbf{t}}(c) = p(\omega)/p(\omega_0)$. Thus, for any *cip*'s for $\mathbf{t}$, p and p', and for any two states ω_0 and ω, $p(\omega)/p(\omega_0) = p'(\omega)/p'(\omega_0)$. $\square$

Proposition 3 was proved in Harsanyi (1967–1968) for Harsanyi type spaces. Samet (1998) noted that for finite spaces the uniqueness of a common prior can be interpreted as the uniqueness of an invariant probability function for an ergodic Markov chain. The simple proof here, for countable spaces, is an extension of the proof in Harsanyi (1967–1968) to general knowledge spaces.

The following proposition is close in its content to the main result in Rodrigues-Neto (2009).

Proposition 4. *Let $\mathbf{t}$ be a positive type profile over a connected partition profile. Then there exists a common improper prior for $\mathbf{t}$ iff for each ω_0 and ω, and chains c and c' from ω_0 to ω, $\text{tr}_{\mathbf{t}}(c) = \text{tr}_{\mathbf{t}}(c')$.*

Proof. As we have shown before, if there exists a common improper prior p for t, then all chains c connecting ω_0 and ω satisfy $\text{tr}(c) = p(\omega)/p(\omega_0)$. Conversely, suppose that for each ω_0 and ω, all the chains from ω_0 to ω have the same type ratio. Fix ω_0 and for each ω let $p(\omega) = \text{tr}(c)$ for some c from ω_0 to ω. To see that p is a *cip* consider $\pi \in \Pi_i$ and $\omega \in \pi$. Let c be a chain from ω_0 to ω. For $\omega' \in \pi$, consider the chain $c' = c \stackrel{i}{\to} \omega'$. Then, by the definitions of tr and p, $p(\omega') = \text{tr}(c') = \text{tr}(c)\,\text{tr}^i(\omega, \omega') = p(\omega) t_i(\pi, \omega')/t_i(\pi, \omega)$. Thus, $p(\pi) = \sum_{\omega' \in \pi} p(\omega') = [p(\omega)/t_i(\pi, \omega)] \sum_{\omega' \in \pi} t_i(\pi, \omega') = p(\omega)/t_i(\pi, \omega) < \infty$, and $p(\omega) = p(\pi) t_i(\pi, \omega)$. $\square$

4.3. Proof of the second parts of Theorems 1 and 2

We first prove our claims for a connected partition profile $\boldsymbol{\Pi}$. Let P be the set of positive types in $\mathcal{T}(\boldsymbol{\Pi})$ and C the set of consistent type profiles in $\mathcal{T}(\boldsymbol{\Pi})$.

We show in the following two results that $C \cap P$ is nowhere dense, that is, that the complement of its closure is dense.

Lemma 1. *If $\boldsymbol{\Pi}$ is connected, then* $\mathrm{cl}(C \cap P) \subseteq (C \cap P) \cup P^c$.

Proof. We need to show that if a sequence of type profiles $\mathbf{t}^n$ in $C \cap P$ converges to $\mathbf{t} \in P$, then $\mathbf{t} \in C$. Let c and c' be chains from ω_0 to ω. By Proposition 4, $\mathrm{tr}_{\mathbf{t}^n}(c) = \mathrm{tr}_{\mathbf{t}^n}(c')$ for each n. Since each chain involves only finitely many states, it follows by continuity that $\mathrm{tr}_{\mathbf{t}}(c) = \mathrm{tr}_{\mathbf{t}}(c')$. Again, by Proposition 4, this implies that $t \in C$. □

Thus, $[\mathrm{cl}(C \cap P)]^c \supseteq [(C \cap P) \cup P^c]^c = C^c \cap P$, and it is enough to show that $C^c \cap P$ is dense.

Lemma 2. *If $\boldsymbol{\Pi}$ is connected and not tight then $C^c \cap P$ is dense in $\mathcal{T}(\boldsymbol{\Pi})$.*

Proof. We show that $C \cap P \subseteq \mathrm{cl}(C^c \cap P)$. Thus, $P \subseteq \mathrm{cl}(C^c \cap P)$, and as P is dense, the claim of the proposition follows.

Since $\boldsymbol{\Pi}$ is connected but not tight, there exists, by Proposition 1 a connected partition profile $\boldsymbol{\Pi}'$ which properly refines $\boldsymbol{\Pi}$. We may assume that $\boldsymbol{\Pi}'$ is obtained from $\boldsymbol{\Pi}$ by splitting one partition element $\pi \in \Pi_i$, for some i, into π^1 and π^2.

For $\mathbf{t} \in P$, define a type profile $\hat{\mathbf{t}}$ for $\boldsymbol{\Pi}$ which agrees with $\mathbf{t}$ except on π. Formally, for each $j \neq i$, $\hat{t}_j = t_j$. For each $\bar{\pi} \neq \pi$ in Π_i, $\hat{t}_i(\bar{\pi}, \cdot) = t_i(\bar{\pi}, \cdot)$. For $\omega \in \pi^1$, $\hat{t}_i(\pi, \omega) = (1+\varepsilon)t_i(\pi, \omega)/c$, and for $\omega \in \pi^2$, $\hat{t}_i(\pi, \omega) = (1-\varepsilon)t_i(\pi, \omega)/c$, where $c = 1 + \varepsilon[t_i(\pi, \pi^1) - t_i(\pi, \pi^2)]$ and $\varepsilon \neq 0$ between -1 and 1. By choosing ε close enough to 0, $\hat{\mathbf{t}}$ can be made arbitrarily close to $\mathbf{t}$.

For any type profile $\mathbf{t}$ in P, let $\mathbf{t}'$ be the type profile for $\boldsymbol{\Pi}'$ which is naturally induced by $\mathbf{t}$ as follows. For each $j \neq i$, $t'_j = t_j$. For each $\bar{\pi} \neq \pi$ in Π_i, $t'_i(\bar{\pi}, \cdot) = t_i(\bar{\pi}, \cdot)$. Finally, for $k = 1, 2$, $t'_i(\pi^k, \cdot) = t_i(\pi, \cdot)/t_i(\pi, \pi^k)$.

Now, let $\mathbf{t} \in P \cap C$ have a *cip* p. We show that $\hat{\mathbf{t}} \in C^c \cap P$. Obviously, p is also a *cip* for $\mathbf{t}'$. Suppose that $\hat{\mathbf{t}}$ has a *cip*, and denote it by $\hat{p}$. Then $\hat{p}$ is also a *cip* for $\hat{\mathbf{t}}'$. But $\hat{\mathbf{t}}' = \mathbf{t}'$, and as $\boldsymbol{\Pi}'$ is connected, it follows, by Proposition 3, that p and $\hat{p}$ differ by a multiplicative constant. Thus, p is a *cip* for $\hat{\mathbf{t}}$ as well. Hence p must satisfy $t_i(\pi, \pi^1) = p(\pi^1)/p(\pi) = \hat{t}_i(\pi, \pi^1)$. But this does not hold as $\hat{t}_i(\pi, \pi^1) = (1+\varepsilon)t_i(\pi, \pi^1)/c$ and $(1+\varepsilon)/c \neq 1$. □

Assume throughout the rest of this section that $\boldsymbol{\Pi}$ is not tight.

One has that $C = (C \cap P) \cup (C \cap P^c) \subseteq (C \cap P) \cup P^c$, and we have shown that $(C \cap P)$ is nowhere dense. In the finite case, P is an open dense set and thus P^c is nowhere dense, so that $(C \cap P) \cup P^c$ is nowhere dense as a finite union of nowhere dense set, and C is nowhere dense as a subset of a nowhere dense set.

For the infinite case, it suffices to show that P^c is of first category. This is indeed the case, because the set T^i_ω of type profiles t for which $t_i(\Pi_i(\omega), \omega) = 0$ is closed and has an empty interior, as its complement contains P which is dense. Thus, T^i_ω is nowhere dense. Finally, $P^c = \bigcup_i \bigcup_\omega T^i_\omega$.

Consider next a partition profile $\boldsymbol{\Pi}$ that is not connected and not tight. For $M \in \wedge\boldsymbol{\Pi}$, denote by $\mathcal{T}_M(\boldsymbol{\Pi}_M)$ the set of type profiles over the knowledge space $(M, \boldsymbol{\Pi}_M)$, and let C_M be the set of consistent type profiles in $\mathcal{T}_M(\boldsymbol{\Pi}_M)$. We can obviously identify $\mathcal{T}(\boldsymbol{\Pi})$ with $\bigtimes_{M \in \wedge\boldsymbol{\Pi}} \mathcal{T}_M(\boldsymbol{\Pi}_M)$.

A type profile $\mathbf{t}$ for $\boldsymbol{\Pi}$ has a *cip* if and only if there exists $M \in \wedge\boldsymbol{\Pi}$ for which $\mathbf{t}_M$, the restriction of $\mathbf{t}$ to $M \times M$, is in C_M, the set of consistent type profiles in $\mathcal{T}_M(\boldsymbol{\Pi}_M)$. Indeed, if $\mathbf{t}_M$ has a *cip* p_M, then the function p on Ω that agrees with p_M on M and vanishes outside M is a *cip* for $\mathbf{t}$. Conversely, if p is a *cip* for $\mathbf{t}$, then for some M, p is not identically 0 on M and thus the restriction of p to M is a *cip* for $\mathbf{t}_M$. We conclude that C, the set of consistent type profiles in $\mathcal{T}(\boldsymbol{\Pi})$, is $\bigcup_{M \in \wedge\boldsymbol{\Pi}} [C_M \times (\bigtimes_{M' \neq M} \mathcal{T}_{M'}(\boldsymbol{\Pi}_{M'}))]$.

Since $\boldsymbol{\Pi}$ is not tight, it follows by Proposition 1 that for each $M \in \wedge\boldsymbol{\Pi}$, $\boldsymbol{\Pi}_M$ is not tight.

In the finite case, this implies that C_M is nowhere dense in $\mathcal{T}_M(\boldsymbol{\Pi}_M)$ and therefore each of the sets in the union is nowhere dense in $\mathcal{T}(\boldsymbol{\Pi})$. Hence, C is nowhere dense as a finite union of nowhere dense sets.

In the infinite case, C_M is of first category and therefore each of the sets in the union is of first category in $\mathcal{T}(\boldsymbol{\Pi})$. Hence, C is of first category as a countable union of sets of first category. We have thus completed the proofs of the second parts of both Theorems 1 and 2. □

4.4. *Proof of the first parts of Theorems 1 and 2*

We say that a chain c is **alternating** if no two consecutive states, ω_s and ω_{s+1}, in c, are the same, and no two consecutive agents, i_s and i_{s+1}, in c, are the same. In particular, any chain of length 0 is alternating and any chain of length 1 from ω_0 to $\omega \neq \omega_0$ is alternating.

Given a connected partition profile $\boldsymbol{\Pi}$, define a distance function d on $\Omega \times \Omega$ such that for each ω and ω', $d(\omega, \omega')$ is the minimal length of a chain from ω to ω'. It is easy to see that d is a metric. A chain from ω_0 to ω of the minimal length $d(\omega, \omega_0)$ is called a **minimal chain**. It is easy to see that if $\omega_0 \ldots \omega_n$ is minimal then $\omega_0 \ldots \omega_s$ is a minimal chain for each $s = 0, \ldots, n$, and therefore $d(\omega_s, \omega_0) = s$. Moreover, the chain must be alternating, because if either $\omega_s = \omega_{s+1}$ or $i_s = i_{s+1}$ we get a shorter chain from ω_0 to ω_n by omitting $\xrightarrow{i_s} \omega_{s+1}$.

Clearly, if for some i, $\omega' \in \Pi_i(\omega)$, then $d(\omega, \omega') \leqslant 1$. Thus, by the triangle inequality, if for some i, $\omega' \in \Pi_i(\omega)$, then for any ω_0, $|d(\omega, \omega_0) - d(\omega', \omega_0)| \leqslant 1$. Thus, on each partition element π, $d(\cdot, \omega_0)$ can have at most two values. In particular, for any chain $c = \omega_0 \xrightarrow{i_0} \omega_1 \xrightarrow{i_1} \cdots \xrightarrow{i_{n-1}} \omega_n$, $|d(\omega_{s+1}, \omega_0) - d(\omega_s, \omega_0)| \leqslant 1$ for $s = 0, \ldots, n-1$.

Proposition 5. *A connected partition profile $\boldsymbol{\Pi}$ is tight if and only if for any states ω and ω' there exists a unique alternating chain for $\boldsymbol{\Pi}$ from ω to ω'.*

Proof. Assume that $\boldsymbol{\Pi}$ is not tight. Then, there exists a connected partition profile $\boldsymbol{\Pi}'$ such that $\boldsymbol{\Pi}' \succ \boldsymbol{\Pi}$. Let ω, ω' and i be such that $\omega' \in \Pi_i(\omega)$ but $\omega' \notin \Pi'_i(\omega)$. Since $\boldsymbol{\Pi}'$ is connected, there exists a minimal chain c for $\boldsymbol{\Pi}'$ from ω to ω', which, as we have shown, is alternating. Since $\boldsymbol{\Pi}'$ is a refinement of $\boldsymbol{\Pi}$, c is also a chain for $\boldsymbol{\Pi}$ and it is alternating. But as $\omega \neq \omega'$, $c' = \omega \xrightarrow{i} \omega'$ is also an alternating chain for $\boldsymbol{\Pi}$ which is different from c, since $\omega' \notin \Pi'_i(\omega)$.

Assume now that $\boldsymbol{\Pi}$ is tight. To show that the condition in the proposition holds we use the following two lemmas.

Lemma 3. *If $\boldsymbol{\Pi}$ is tight then for each ω_0 and ω there exists a unique minimal chain from ω_0 to ω.*

We show that if there are two distinct minimal chains from one state to another then $\boldsymbol{\Pi}$ is not tight. Let $c = \omega_0 \xrightarrow{i_0} \omega_1 \xrightarrow{i_1} \cdots \xrightarrow{i_{n-1}} \omega_n$ and $c' = \omega_0 \xrightarrow{i'_0} \omega'_1 \xrightarrow{i'_1} \cdots \xrightarrow{i'_{n-1}} \omega_n$ be distinct minimal chains, and assume that n is the minimal number for which such a pair exists. Obviously, $n > 0$. It is impossible that both $i_0 = i'_0$ and $\omega_1 = \omega'_1$, because either $n = 1$ in which case $c = c'$, or else $n > 1$ in which case $\omega_1 \xrightarrow{i_1} \cdots \xrightarrow{i_{n-1}} \omega_n$ and $\omega'_1 \xrightarrow{i'_1} \cdots \xrightarrow{i'_{n-1}} \omega_n$ are distinct minimal chains of length $n-1$ contrary to the minimality of n. Thus, either $i_0 \neq i'_0$ or $\omega_1 \neq \omega'_1$.

Consider the refinement $\boldsymbol{\Pi}'$ of $\boldsymbol{\Pi}$ obtained by splitting $\Pi_i(\omega_0)$ into $\{\omega_1\}$ and $\Pi_i(\omega_0) \setminus \{\omega_1\}$. The latter set is not empty since being minimal, c is alternating and thus, $\omega_0 \neq \omega_1$. We will prove that $\boldsymbol{\Pi}$ is not tight by showing that $\boldsymbol{\Pi}'$ is connected. To do so, it suffices to prove that every $\hat{\omega} \in \Pi_i(\omega_0) \setminus \{\omega_1\}$ is connected to ω_1 for $\boldsymbol{\Pi}'$. As $\hat{\omega}$ is connected to ω_0 for $\boldsymbol{\Pi}'$, it suffices to show that ω_0 is connected to ω_1 for $\boldsymbol{\Pi}'$.

Assume first that $\omega_1 = \omega'_1$. Thus, $i_0 \neq i'_0$ and therefore $\omega_0 \xrightarrow{i'_0} \omega_1$, is a chain for $\boldsymbol{\Pi}'$. Now assume that $\omega_1 \neq \omega'_1$, which implies that $n > 1$. Note that all states in $\Pi_i(\omega_0)$ are of distance not greater than 1 from ω_0 and thus each of the states $\omega_2, \ldots, \omega_{n-1}, \omega_n$ and $\omega'_2, \ldots, \omega'_{n-1}, \omega_n$ are not in this set, as their distance from ω_0 is greater than 1. Thus, $\omega_1 \xrightarrow{i_1} \cdots \xrightarrow{i_{m-1}} \omega_m$ and $\omega'_1 \xrightarrow{i'_1} \cdots \xrightarrow{i'_{m-1}} \omega_m$ are chains for $\boldsymbol{\Pi}'$ too. Also, because $\omega_1 \neq \omega'_1$, $\omega_0 \xrightarrow{i'_0} \omega'_1$ is a chain for $\boldsymbol{\Pi}'$ (even if $i_0 = i'_0$). Thus, we have shown that the following relations hold for $\boldsymbol{\Pi}'$: $\omega_0 \to \omega'_1$ $\omega'_1 \to \omega_n$, and $\omega_n \to \omega_1$, which amounts to saying that ω_0 and ω_1 are connected in $\boldsymbol{\Pi}'$.

Lemma 4. *If $\boldsymbol{\Pi}$ is tight then every alternating chain for $\boldsymbol{\Pi}$ is minimal.*

The proof is by induction on n, the length of the chain. The claim is obvious for alternating chains of lengths $n = 0$ and $n = 1$. Suppose the claim holds for alternating chains of length $n = k \geqslant 1$, and assume that c is an alternating chain $c = \omega_0 \xrightarrow{i_0} \omega_1 \xrightarrow{i_1} \cdots \to \omega_k \xrightarrow{i_k} \omega$ of length $k+1$. By the induction hypothesis the alternating chain $\omega_0 \ldots \omega_k$ is minimal and thus for all $s \leqslant k$, $d(\omega_s, \omega_0) = s$. In particular $d(\omega_k, \omega_0) = k$ and thus $d(\omega, \omega_0)$ is either $k+1$, or k, or $k-1$. We only need to prove that the last two values are impossible. Suppose that $d(\omega, \omega_0) < k+1$. Let $c' = \omega_0 \xrightarrow{i'_0} \omega'_1 \ldots \omega$ be a minimal chain from ω_0 to ω. Then it is of length $k-1$ or k.

Consider the refinement $\boldsymbol{\Pi}'$ obtained by splitting $\Pi_{i_k}(\omega_k)$ into $\{\omega_k\}$ and $\Pi_{i_k}(\omega_k) \setminus \{\omega_k\}$. The latter set is not empty, since by the alternation of c, $\omega_k \neq \omega$. We show that for each $\hat{\omega} \in \Pi_{i_k}(\omega_k) \setminus \{\omega_k\}$ there is a chain for $\boldsymbol{\Pi}'$ from ω_k to $\hat{\omega}$. Note, first, that the chain $\omega_0 \ldots \omega_k$ is a chain for $\boldsymbol{\Pi}'$. To see this, observe that by alternation $i_{k-1} \neq i_k$. Thus, $\omega_{k-1} \xrightarrow{i_{k-1}} \omega_k$ is a chain for $\boldsymbol{\Pi}'$. Also the states $\omega_0, \ldots, \omega_{k-2}$ are of distance less than $k-1$ from ω_0 while all states in $\Pi_{i_k}(\omega_k)$ are of distance $k-1$ at least. Thus, $\omega_0 \ldots \omega_{k-2}$ is also a chain for $\boldsymbol{\Pi}'$. We conclude that there is a chain for $\boldsymbol{\Pi}'$ from ω_k to ω_0. We end the proof by showing that there is a chain in $\boldsymbol{\Pi}'$ from $\hat{\omega}$ to ω_0. First, $\hat{\omega} \xrightarrow{i_k} \omega$ is a chain for $\boldsymbol{\Pi}'$ by the construction of the latter. Moreover, none of the states in c' is ω_k, which implies that c' is also a chain for $\boldsymbol{\Pi}'$. Indeed, $\omega \neq \omega_k$ by the alternation of c, and the distance from ω_0 to all of the states in c' that precede ω is less than k, while ω_k is of distance k from ω_0. □

The characterization of tightness in terms of alternating chains is used to prove the following.

Lemma 5. *If $\boldsymbol{\Pi}$ is connected and tight then every positive type profile has a common improper prior.*

Proof. We show that the condition in Proposition 4 holds. By Proposition 5 for any states ω_0 and ω there exists a unique alternating chain $c(\omega_0, \omega)$ from ω_0 to ω. We show that for any ω_0 and ω and any chain c from ω_0 to ω, $\mathrm{tr}(c) = \mathrm{tr}(c(\omega_0, \omega))$. The proof is by induction on the length of c. The claim trivially holds for chains of length 0. Assume that it holds for all chains of length $n-1$ for $n > 0$ and let $c = c' \xrightarrow{i} \omega$ be a chain from ω_0 to ω of length n, where c' is a chain from ω_0 to ω'. By the induction hypothesis, $\mathrm{tr}(c') = \mathrm{tr}(c(\omega_0, \omega'))$. Consider the chain $\hat{c} = c(\omega_0, \omega') \xrightarrow{i} \omega$. Then, $\mathrm{tr}(c) = \mathrm{tr}(c')\,\mathrm{tr}^i(\omega', \omega) = \mathrm{tr}(c(\omega_0, \omega'))\,\mathrm{tr}^i(\omega', \omega) = \mathrm{tr}(\hat{c})$, and therefore it suffices to show that $\mathrm{tr}(\hat{c}) = \mathrm{tr}(c(\omega_0, \omega))$. If $\hat{c}$ is alternating then $\hat{c} = c(\omega_0, \omega)$, and we are done. If $\hat{c}$ is not alternating, then, since $c(\omega_0, \omega')$ is alternating, one of the following two cases holds. Case 1: $\omega' = \omega$. In this case $\mathrm{tr}(\hat{c}) = \mathrm{tr}(c(\omega_0, \omega))\,\mathrm{tr}^i(\omega, \omega) = \mathrm{tr}(c(\omega_0, \omega))$. Case 2: $\omega \neq \omega'$, and for $c(\omega_0, \omega') = \omega_0 \ldots \omega_{n-2} \stackrel{i_{n-2}}{\to} \omega'$, $i_{n-2} = i$. Here, $\mathrm{tr}(\hat{c}) = \mathrm{tr}(\omega_0 \ldots \omega_{n-2})\,\mathrm{tr}^i(\omega_{n-2}, \omega')\,\mathrm{tr}^i(\omega', \omega) = \mathrm{tr}(\omega_0 \ldots \omega_{n-2})\,\mathrm{tr}^i(\omega_{n-2}, \omega)$. But the latter is the type ratio of the chain $\omega_0 \ldots \omega_{n-2} \stackrel{i}{\to} \omega$, which, being alternating is $c(\omega_0, \omega)$. □

The proof of part 1 in Theorems 1 and 2 follows readily in case $\boldsymbol{\Pi}$ is connected. As before, P is the set of positive type profiles and C the set of type profiles that have a *cip*. By Lemma 5, $C^c \subseteq P^c$.

In the infinite case, we have shown that P^c is of first category, and thus, C^c is of first category. In the finite case, C is closed. Indeed, let $\mathbf{t}^n$ be a sequence of type profiles in C that converges to $\mathbf{t}$. For each n, $\mathbf{t}^n$ has a common prior p^n that satisfies for each i and $\pi \in \Pi_i$, $p^n(\pi) t_i^n(\pi, \cdot) = p^n(\cdot)$. By the compactness of $\Delta(\Omega)$, a subsequence of p^n converges to a probability function p. By continuity, $p(\pi) t_i(\pi, \cdot) = p(\cdot)$ for each i and $\pi \in \Pi_i$. Thus, p is a common prior for $\mathbf{t}$ and $\mathbf{t} \in C$. By Lemma 5, $P \subset C$ and thus, $\mathcal{T}(\boldsymbol{\Pi}) = \mathrm{cl}(P) \subseteq \mathrm{cl}(C) = C$.

Suppose that $\boldsymbol{\Pi}$ is not connected, then $C^c = \bigtimes_{M \in \wedge \boldsymbol{\Pi}} C^c_M$. As $\boldsymbol{\Pi}$ is tight, there is an $\hat{M} \in \wedge \boldsymbol{\Pi}$ such that $\boldsymbol{\Pi}_{\hat{M}}$ is tight. In the finite case, $C^c_{\hat{M}} = \emptyset$ and therefore $C^c = \emptyset$. In the infinite case, $C^c_{\hat{M}} \subseteq P^c_{\hat{M}}$, and as $P^c_{\hat{M}}$ is of first category, so is C^c. □

4.5. Proof of Proposition 2

Let $\boldsymbol{\Pi}$ be a connected tight partition profile. The proof is by induction on the size on Ω. If Ω is a singleton the equality in the proposition is obvious. Suppose the equality is proved for all state spaces smaller than $n > 1$ and let $|\Omega| = n$. Since $n \geqslant 2$ and $\boldsymbol{\Pi}$ is connected, there must be i and ω_0 such that $\Pi_i(\omega_0)$ is not a singleton. Consider the refinement of $\boldsymbol{\Pi}$, $\hat{\boldsymbol{\Pi}}$, obtained by splitting $\Pi_i(\omega_0)$ into $\{\omega_0\}$ and $\Pi_i(\omega_0) \setminus \{\omega_0\}$. By the tightness of $\boldsymbol{\Pi}$, $\hat{\boldsymbol{\Pi}}$ is not connected.

Let Ω_0 consist of all states ω such that there is a chain for $\hat{\boldsymbol{\Pi}}$ from ω_0 to ω. Fix ω_1 in the set $\Pi_i(\omega_0) \setminus \{\omega_0\}$ and let Ω_1 be the set of all ω such that there is a chain for $\hat{\boldsymbol{\Pi}}$ from ω_1 to ω. Each of Ω_0 and Ω_1 is an element of the meet of $\hat{\boldsymbol{\Pi}}$. They are disjoint because if they shared a state then ω_0 would be connected to ω_1 which would make $\hat{\boldsymbol{\Pi}}$ connected. Each state ω is in either Ω_0 or Ω_1. Indeed, let c be a minimal chain for $\boldsymbol{\Pi}$ from ω_0 to an arbitrary ω. If c does not contain a state $\hat{\omega} \in \Pi_i(\omega_0) \setminus \{\omega_0\}$, then c is a chain for $\hat{\boldsymbol{\Pi}}$ and $\omega \in \Omega_0$. If c does contain such a $\hat{\omega}$, then no state that follows $\hat{\omega}$ in c is ω_0 (because the distance of each state in c to ω_0, other than ω_0 itself, is positive). Thus, there is a chain for $\hat{\boldsymbol{\Pi}}$ from $\hat{\omega}$ to ω, and trivially there is a chain for $\hat{\boldsymbol{\Pi}}$ from $\hat{\omega}$ to ω_1, so that $\omega \in \Omega_1$. Thus, the meet of $\hat{\boldsymbol{\Pi}}$ is exactly the set $\{\Omega_0, \Omega_1\}$, and each of $\hat{\boldsymbol{\Pi}}^0$ and $\hat{\boldsymbol{\Pi}}^1$, respectively the restriction of $\hat{\boldsymbol{\Pi}}$ to Ω_0 and Ω_1, is connected.

By the induction hypothesis, for $k = 0, 1$, $\sum_{i \in I} |\hat{\boldsymbol{\Pi}}^k_i| = (|I| - 1)|\Omega_k| + 1$. By adding the two equations and noting that $\sum_{i \in I} |\hat{\boldsymbol{\Pi}}^0_i| + \sum_{i \in I} |\hat{\boldsymbol{\Pi}}^1_i| = \sum_{i \in I} |\Pi_i| + 1$ we get the desired equality.

If $\boldsymbol{\Pi}$ is not tight, then it must have a refinement which is tight, and therefore it satisfies the inequality of the proposition. □

4.6. Examples

Example 2. We construct an infinite knowledge space with a tight partition profile, such that the set of inconsistent type profiles is dense. Therefore it is not nowhere dense, since the complement of a nowhere dense set contains a nonempty open set. To show that the set of inconsistent type profiles is dense, it is enough to show that it is dense in the set of positive type profiles, since the latter is dense.

Consider a knowledge space for two agents, where Ω is the set of integers $\mathbb{Z}$, and the partitions are $\Pi_1 = \{\pi_1^n \mid n \in \mathbb{Z}\}$, where $\pi_1^n = \{2n, 2n+1\}$, and $\Pi_2 = \{\pi_2^n \mid n \in \mathbb{Z}\}$, where $\pi_2^n = \{2n-1, 2n\}$. The partition profile $\boldsymbol{\Pi} = (\Pi_1, \Pi_2)$ is tight, since it is connected and any proper refinement of $\boldsymbol{\Pi}$ is not. Let $\mathbf{t} = (t_1, t_2)$ be a positive type profile over $\boldsymbol{\Pi}$. We construct a sequence of inconsistent type profiles $\mathbf{t}^k$ such that $\mathbf{t}^k$ converges to $\mathbf{t}$ as $k \to -\infty$. For $n \leqslant k$, $t_1^k(\pi_1^n, 2n) = 1$ and $t_2^k(\pi_2^n, 2n) = 0$. For $n > k$, $t_1^k(\pi_1^n, \cdot) = t_1(\pi_1^n, \cdot)$ and $t_2^k(\pi_2^n, \cdot) = t_2(\pi_2^n, \cdot)$. Obviously, $\mathbf{t}^k$ converges to $\mathbf{t}$ as $k \to -\infty$.

To show that $\mathbf{t}^k$ is inconsistent we prove that if p is a *cip* for $\mathbf{t}^k$, then it must be identically 0, which is impossible for a *cip*. By the definition of *cip* whenever for some i, $\pi \in \Pi_i$ and $\omega \in \pi$, $t_i(\pi, \omega) = 0$, then $p(\omega) = 0$. Now, $t_1^k(\pi_1^k, 2k+1) = 0$ and therefore $p(2k+1) = 0$. Also, for each $m \leqslant 2k$, either $t_1^k(\Pi_1(m), m) = 0$ or $t_2^k(\Pi_2(m), m) = 0$. Thus, $p(m) = 0$ for all $m \leqslant 2k+1$. We prove now by induction on m that $p(m) = 0$ for all $m \geqslant 2k+1$. This holds as we have shown for $m = 2k+1$. Suppose that for $m = 2n+1$, $p(m) = 0$. Since $t_2^k(\pi_2^{2n+1}, 2n+1) = t_2(\pi_2^{2n+1}, 2n+1) > 0$ it follows by the definition of *cip* that $p(\pi_2^{2n+1}) = 0$. This implies that $p(2n+2) = 0$. The induction step when $m = 2n$ is similar.

Example 3. We construct an infinite knowledge space for two agents with a partition profile which is not tight, such that the set of consistent type profiles is dense, which shows that it is not nowhere dense. To show this we prove that the set of consistent type profiles is dense in the set of positive type profiles.

Let Ω be the set $\mathbb{N} \times \mathbb{N}$. Player 1's partition consists of the rows and 2's the columns. That is, $\Pi_1 = \{\pi_1^i \mid i \in \mathbb{N}\}$, where $\pi_1^i = \{(i, j) \mid j \in \mathbb{N}\}$, and $\Pi_2 = \{\pi_2^j \mid j \in \mathbb{N}\}$, where $\pi_2^j = \{(i, j) \mid i \in \mathbb{N}\}$.

Let $\mathbf{t}$ be a positive type profile for this partition profile. We define a sequence of consistent type profiles $\mathbf{t}^n$ that converge to $\mathbf{t}$ as $n \to \infty$. Fix a consistent type profile $\hat{\mathbf{t}}$ with *cip* p. For each $i \leqslant n$ and $j \leqslant n$, let

$$t_1^n\big(\pi_1^i, (i, j)\big) = t_1\big(\pi_1^i, (i, j)\big) \Big/ \sum_{k=1}^{n} t_1\big(\pi_1^i, (i, k)\big)$$

and

$$t_2^n\big(\pi_2^j, (i, j)\big) = t_2\big(\pi_2^j, (i, j)\big) \Big/ \sum_{k=1}^{n} t_2\big(\pi_2^j, (k, j)\big).$$

For $i \geqslant n+1$ and $j \geqslant n+1$, $t_1^n(\pi_1^i, (i, j)) = \hat{t}_1(\pi_1^{i-n}, (i-n, j-n))$, and $t_2^n(\pi_2^j, (i, j)) = \hat{t}_2(\pi_2^{j-n}, (i-n, j-n))$. For $i \leqslant n$ and $j \geqslant n+1$, or $i \geqslant n = 1$ and $j \leqslant n$, $t_1^i(\pi_1^i, (i, j)) = t_2^j(\pi_2^j, (i, j)) = 0$.

It is easy to check that for each i, $\|t_1^n(\pi_1^i, \cdot) - t_1(\pi_1^i, \cdot)\| \to 0$ when $n \to \infty$, and a similar convergence holds for agent 2. Thus, $\mathbf{t}^n \to \mathbf{t}$. To see that $\mathbf{t}^n$ is consistent, we define p^n by $p^n(i, j) = p(i-n, j-n)$ for $i \geqslant n+1$ and $j \geqslant n+1$ and $p^n(i, j) = 0$ otherwise. It is easy to see that p^n is a *cip* for $\mathbf{t}^n$.

References

Aumann, R.J., 1976. Agreeing to disagree. Ann. Stat. 4 (6), 1236–1239.
Aumann, R.J., 1987. Correlated equilibrium as an expression of Bayesian rationality. Econometrica 55, 1–18.
Aumann, R.J., 1998. Common priors: A reply to Gul. Econometrica 66, 929–938.
Gul, F., 1998. A comment on Aumann's Bayesian view. Econometrica 66, 923–927.
Harsanyi, J.C., 1967–1968. Games with incomplete information played by Bayesian players. Manage. Sci. 14, 159–182, 320–334, 486–502.
Hintikka, J., 1962. Knowledge and Belief. Cornell University Press, Ithaca, NY.
Mertens, J.F., Zamir, S., 1985. Formulation of Bayesian analysis for games with incomplete information. Int. J. Game Theory 14, 1–29.
Munkres, J.R., 1975. Topology. Prentice–Hall Inc., Englewood Cliffs, NJ.
Nyarko, Y., 2010. Most games violate the common priors doctrine. Int. J. Econ. Theory 6, 189–194.
Rodrigues-Neto, J.A., 2009. From posteriors to priors via cycles. J. Econ. Theory 144, 876–883.
Samet, D., 1998. Iterated expectations and common priors. Games Econ. Behav. 24, 131–141.

CHAPTER 2

Almost Common Priors

Forthcoming in International Journal of Game Theory.
The paper appears here on pages 16 – 27.

Int J Game Theory
DOI 10.1007/s00182-012-0347-5

Almost common priors

Ziv Hellman

Accepted: 1 August 2012

Abstract What happens when priors are not common? We introduce a measure for how far a type space is from having a common prior, which we term prior distance. If a type space has δ prior distance, then for any bet f it cannot be common knowledge that each player expects a positive gain of δ times the sup-norm of f, thus extending no betting results under common priors. Furthermore, as more information is obtained and partitions are refined, the prior distance, and thus the extent of common knowledge disagreement, can only decrease. We derive an upper bound on the number of refinements needed to arrive at a situation in which the knowledge space has a common prior, which depends only on the number of initial partition elements.

Keywords Common prior · Agreeing to disagree · No betting and no trade · Knowledge and beliefs

1 Introduction

What happens when priors are not common? Can one measure 'how far' a belief space is from having a common prior, and use that to approximate standard results that apply under the common prior assumption?

Surprisingly, there has been relatively little published to date in the systematic study of situations of non-common priors. One of the justifications for the common prior assumption that is often raised is the claim that, once we begin to relax the common priors assumption, 'anything is possible', in the sense that heterogeneous priors allow

Z. Hellman (✉)
Department of Mathematics and the Centre for the Study of Rationality,
Hebrew University, Jerusalem, Israel
e-mail: ziv.hellman@mail.huji.ac.il

Published online: 24 August 2012

'sufficient freedom as to be capable of generating virtually any outcome'. (The quote is from Samuelson (2004). A similar argument appears in Morris (1995b)). Perhaps this is one reason that the assumption that players' posterior beliefs in models of differential information are derived from a common prior has been ubiquitous in the literature since Harsányi (1967–1968) introduced the concept in his groundbreaking work on games with incomplete information. Indeed, as pointed out in Aumann (1987), the assumption of a common prior (also known as the Harsányi doctrine) 'is pervasively explicit or implicit in the vast majority of the differential information literature in economics and game theory'. Although more than a score of years have passed since those lines were published, they retain their full force.

We show in this paper that it is not true that by moving away from common priors one can generate 'virtually any outcome'. To the contrary, the common prior assumption is actually quite robust in terms of no betting and agreeing to disagree results, which constitute the main characterisation of common priors in the literature. The central result here, Theorem 1, shows that in models in which priors are 'almost common' there is 'almost no betting'.

Moreover, we can continuously measure how far a type space is from having a common prior and use that to bound common knowledge betting and disagreements. More precisely, we show that every type space can be associated with a value which we term the prior distance that is the intuitive measure of 'how far' the type space is from a common prior. Letting $E_i f(\omega)$ denote player i's posterior expected value of a random variable f at state ω, if there is a common prior then the well-known No Betting theorem for two players states that

$$\neg\exists f \in \mathbb{R}^{\Omega}, \quad \forall \omega \in \Omega, \quad E_1 f(\omega) > 0 \wedge E_2(-f)(\omega) > 0.$$

Setting $S = \{f \in \mathbb{R}^{\Omega} \mid \|f\|_{\infty} \leq 1\}$, the results in this paper state that if the type profile has δ prior distance, then

$$\neg\exists f \in S, \quad \forall \omega \in \Omega, \quad E_1 f(\omega) > \delta \wedge E_2(-f)(\omega) > \delta.$$

If $\delta = 0$ there is a common prior and the common prior result is recapitulated.

In the n-player case, if the prior distance is δ then there is no n-tuple of random variables $f = (f_1, \ldots, f_n)$ such that $\sum_i f_i = 0$ and it is common knowledge that in every state $E_i f_i(\omega) > \delta \|f\|_{\infty}$.

By scaling f, $\|f\|_{\infty}$ can be made as large as desired, and hence the upper bound on common knowledge disagreements can also be raised without limit unless $\delta = 0$. Never the less, when taking into account budget constraints and possible risk aversion on the part of the players, the upper bound can have bite, even when $\delta \neq 0$. For example, if δ is sufficiently small (say, less than 10^{-9}), then in order to have common knowledge that each player expects a positive gain of more than a few pennies, a bet may need to be scaled so much that in some states of the world billions of dollars are at stake.

Finally, we show that after sufficiently many proper refinements of the partition space, the resulting type space is one in which the players must have a common prior, and in fact we derive an upper bound on the number of such refinement steps needed based solely on the structure of the partition profile. This may justify, in some models, supposing that the analysis of the model begins after the players have received sufficient signals to have refined their partitions to the point where they might as well have started with a common prior. This supposition may be more philosophically acceptable than the stark assertion of the common prior assumption, an assumption that has been much debated in the literature.

2 Preliminaries

2.1 Knowledge and belief

Denote by $\|\cdot\|_1$ the L^1 norm of a vector, i.e., $\|x\|_1 := \sum_{i=1}^{m} |x_i|$. Similarly, denote by $\|\cdot\|_\infty$ the L^∞ norm of a vector, i.e., $\|x\|_\infty := \max(|x_1|, \dots |x_m|)$. More generally, for any real number $1 < p < \infty$, $\|\cdot\|_p := (\sum_{i=1}^{m} |x_i|^p)^{1/p}$. Two norms $\|\cdot\|_p$ and $\|\cdot\|_q$ are dual norms if p and q are dual conjugates, i.e., if $1/p + 1/q = 1$ when $1 < p < \infty$ and $1 < q < \infty$. $\|x\|_1$ and $\|x\|_\infty$ are also a pair of dual norms.

For a set Ω, denote by $\Delta(\Omega) \subset \mathbb{R}^\Omega$ the simplex of probability distributions over Ω. An *event* is a subset of Ω. A *random variable* f over Ω is an element of $\mathbb{R}^\Omega$. Given a probability distribution $\mu \in \Delta(\Omega)$ and a random variable f, the *expected value of* f with respect to μ is defined by

$$E_\mu f := \sum_{\omega \in \Omega} f(\omega)\mu(\omega) \tag{1}$$

The probability of an event H is the expected value of the random variable 1^H, which is the standard characteristic function defined as:

$$1^H(\omega) = \begin{cases} 1 & \text{if } \omega \in H \\ 0 & \text{if } \omega \notin H \end{cases}$$

A *knowledge space* for a nonempty, finite set of *players* I is a pair (Ω, Π). In this context, Ω is a nonempty set called a *state space* (and each $\omega \in \Omega$ is called a state), and $\Pi = (\Pi_i)_{i \in I}$ is a *partition profile*, where for each $i \in I$, Π_i is a partition of Ω. We will assume throughout this paper that the state space Ω satisfies $|\Omega| = m$, where m is a positive integer, and that $|I| = n$, where $n > 1$.

Π_i is interpreted as the information available to player i; $\Pi_i(\omega)$ is the set of all states that are indistinguishable to i when ω occurs. Give a partition Π_i of Ω of player i, the number of partition elements in Π is denoted $|\Pi_i|$ (we will call it the *size* of Π). For a partition profile $\Pi = (\Pi_1, \dots, \Pi_{|I|})$, the total number of partition elements, i.e. $\sum_{i=1}^{|I|} |\Pi_i|$, is denoted $|\Pi|$ (the *size* of Π).

A *type function* for Π_i is a function $t_i : \Omega \to \Delta(\Omega)$ that associates with each state ω a distribution in $\Delta(\Omega)$, in which case the latter is termed the *type* of i at ω. Each type function t_i further satisfies the following two conditions:

(a) $t_i(\omega)(\Pi_i(\omega)) = 1$, for each $\omega \in \Omega$;
(b) t_i is constant over each element of Π_i.

Given a type function t_i and $f \in \mathbb{R}^\Omega$, $E_i f$ denotes the random variable defined by $(E_i f)(\omega) = t_i(\omega) \cdot f$ (considering both $t_i(\omega)$ and f as vectors in $\mathbb{R}^\Omega$ and taking the standard dot product). We will sometimes relate to $E_i f$ as a vector in $\mathbb{R}^\Omega$, enabling us to use standard vector notation. For example, $E_i f > 0$ will be short-hand for $E_i f(\omega) > 0$ for all $\omega \in \Omega$.

A *type profile*, given Π, is a set of type functions $(t_i)_{i \in I}$, where for each i, t_i is a type function for Π_i, which intuitively represents the player's beliefs. A *type space* τ is then given by a knowledge space and a type profile, i.e., $\tau = \{\Omega, \Pi, (t_i)_{i \in I}\}$, where the t_i are defined relative to Π. A type space τ is called *positive* if $t_i(\omega)(\omega) > 0$ for all $\omega \in \Omega$ and each $i \in I$.

2.2 The meet

A partition Π' is a *refinement* of Π if every element of Π' is a subset of an element of Π. Π' is a *proper refinement* of Π if for at least one $\omega \in \Omega$, $\Pi'(\omega)$ is a proper subset of $\Pi(\omega)$. Refinement intuitively describes an increase of knowledge.

A partition profile Π' is a *(proper) refinement* of Π if for at least one player i, Π'_i is a (proper) refinement of Π_i. That Π' is a proper refinement of Π is denoted $\Pi \prec \Pi'$.

If τ is a type profile over Ω and Π, then a refinement Π' of Π induces a refinement τ' of τ defined by assigning to each $t'_i(\omega)(\omega')$ the probability of $t_i(\omega)(\omega')$ conditional on the event $\pi'_i(\omega)$. Then τ' thus defined is a proper refinement of τ if Π' is a proper refinement of Π. Denote $\tau \prec \tau'$ if τ' is a proper refinement of τ. The *size* $|\tau|$ of a type profile τ is defined to be the size $|\Pi|$ of the partition profile Π over which τ is defined.

If Π' is a refinement of Π, we also say that Π' is *finer* than Π, and that Π is *coarser* than Π'. The *meet* of Π is the finest common coarsening of the players' partitions. Each element of the meet of Π is called a *common knowledge component* of Π. Denote by $C(\Pi)$ the number of common knowledge components in the meet of Π. A type profile Π is called *connected* when its meet is the singleton set $\{\Omega\}$.

2.3 Common priors

A *prior* for a type function t_i is a probability distribution $p \in \Delta(\Omega)$, such that for each $\pi \in \Pi_i$, if $p(\pi) > 0$, and $\omega \in \pi$, then $t_i(\omega)(\cdot) = p(\cdot \mid \pi)$. Denote the set of all priors of player i by $P_i(\tau)$, or simply by P_i when τ is understood.[1] In general, P_i is a set of probability distributions, not a single element; as shown in Samet (1998), P_i is the convex hull of all of i's types.

A *common prior* for the type profile τ is a probability distribution $p \in \Delta(\Omega)$ which is a prior for each player i.[2]

[1] Strictly speaking, the set of priors of a player i depends solely on i's type function t_i, not on the full type profile τ. However, since we are studying connections between sets of priors of different players, we will find it more convenient to write $P_i(\tau)$, as if P_i is a function of τ.

[2] Contrasting a prior for t_i with the types $t_i(\omega, \cdot)$, the latter are referred to as the posterior probabilities of i.

3 Motivating the definition of almost common priors

We will assume in this section and the next that all type profiles are connected. In a connected type profile, stating that $E_i f(\omega) > 0$ for all ω for a random variable f is equivalent to saying that the positivity of $E_i f$ is common knowledge among the players. This simplifying assumption eases the exposition without loss of generality, because all the results here can be extended to non-connected type spaces by taking convex combinations of functions and probability distributions over the common knowledge components.

3.1 Characterization of the existence of a common prior

The main characterization of the existence of a common prior in the literature is based on the concept of agreeable bets.

Definition 1 Given an n-player type space τ, an n-tuple of random variables $(f_1, \ldots, f_n)$ is a *bet* if $\sum_{i=1}^{n} f_i = 0$.

Definition 2 A bet is an *agreeable bet* if[3] $E_i f_i > 0$ for all i.

In the special case of a two-player type space, Definition 2 implies that we may consider a random variable f to be an agreeable bet if $E_1 f > 0 > E_2 f$ (by working with the pair $(f, -f)$).

The standard characterization of the existence of common priors is then:

> A finite type space has a common prior if and only if there does not exist an agreeable bet.

This establishes a fundamental and remarkable two-way connection between posteriors and priors, relating beliefs in prior time periods with behaviour in the present time period. The most accessible proof of this result is in Samet (1998). It was proved by Morris (1995a) for finite type spaces and independently by Feinberg (2000) for compact type spaces. Bonanno and Nehring (1999) proved it for finite type spaces with two agents.

In the special case in which the bet is conducted between two players over the probability of an event H occurring, i.e., the bet is the pair $(1^H, -1^H)$, this result yields the No Disagreements Theorem of Aumann (1976): if there is a common prior and it is common knowledge that player 1's ascribes probability η_1 to H and player 2 ascribes probability η_2 to H, then $\eta_1 = \eta_2$.

3.2 Almost common priors and prior distance

We now wish to generalise the characterization in Sect. 3.1 to situations in which there is no common prior. For some motivating intuition, consider the following scenario.

[3] Recall that we adopted vector notation for $E_i f$, hence $E_i f > 0$ means that $E_i f(\omega) > 0$ for all $\omega \in \Omega$.

The state space is $\Omega = (1, 2, 3, 4, 5, 6, 7, 8)$. There are two players. At time t^0, the knowledge of both players is given by the trivial partition; i.e., $\Pi_1^0 = \Pi_2^0 = \{\Omega\}$, where Π_i^0 is player i's initial partition.

We suppose that the players start with different priors: player 1 has prior

$$\mu_1 = \left(\frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}, \frac{1}{6}, \frac{1}{6}\right),$$

and player 2 has prior

$$\mu_2 = \left(\frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{3}{16}, \frac{3}{16}, \frac{3}{16}, \frac{3}{16}\right).$$

At time t^1, the players receive asymmetric information. As a result, player 1's partition Π_1^1 at t^1 is given by

1 2 3	4 5 6	7 8

and player 2's partition Π_2^2 at t^2 is given by

1 2 3 4	5 6 7 8

.

Using standard belief revision given the priors μ_1 and μ_2 and the partitions Π_1^1 and Π_2^1 yields the type functions t_1 and t_2 of players 1 and 2, respectively, given by

$$t_1(1) = (1/3, 1/3, 1/3, 0, 0, 0, 0, 0)$$
$$t_1(4) = (0, 0, 0, 1/3, 1/3, 1/3, 0, 0)$$
$$t_1(7) = (0, 0, 0, 0, 0, 0, 1/2, 1/2).$$

and

$$t_2(1) = (1/4, 1/4, 1/4, 1/4, 0, 0, 0, 0)$$
$$t_2(5) = (0, 0, 0, 0, 1/4, 1/4, 1/4, 1/4).$$

Under a naïve and erroneous reading of Aumann's No Disagreements Theorem, one might be led to conclude that given the type profile $\tau = \{t_1, t_2\}$ the players would be able to find a disagreement, because the type functions were derived above from the non-equal priors μ_1 and μ_2. However, this is not the case, because τ *could have been* derived instead from a common prior, namely $(1/8, 1/8, 1/8, 1/8, 1/8, 1/8, 1/8, 1/8)$. It therefore satisfies the condition of Aumann's Theorem: since the type profile has a common prior, there can be no common knowledge disagreement.

Definition 3 When we consider a given type space to have been dynamically derived over two or more time periods from a set of priors $(\mu_1, \ldots, \mu_n)$, one prior per each player i, we will call $(\mu_1, \ldots, \mu_n)$ the set of *historical priors* of the type space.

As we have just seen, consideration of the historical priors alone is insufficient for making conclusions regarding agreement and disagreement; what counts is the collection of the sets of priors, $P_i(\tau)$ for each $i \in I$. More explicitly, there is a common prior if and only if $\bigcap_{i=1}^{n} P_i(\tau) \neq \emptyset$. Each historical prior satisfies $\mu_i \in P_i(\tau)$, but whether or not the historical priors are all equal to each other is irrelevant for disagreement theorems, because the historical priors can be entirely disparate even when $\bigcap_{i=1}^{n} P_i(\tau) \neq \emptyset$.

In the two-player case, instead of the historical priors what we need to concentrate on are the points in the players' sets of priors P_1 and P_2 that are of 'minimal distance' from each other in an appropriate metric. In a sense, we are considering an 'alternative history' in which the players derived the same posterior profile but started out with priors that are as close as possible to each other. Since P_1 and P_2 are closed, convex and compact, such a distance is well-defined. There is a common prior if and only if this distance is 0, if and only if there is non-empty intersection of P_1 and P_2.

This leads to the idea that we may measure how far a two-player type space is from having a common prior by measuring a distance between the nearest points of the sets P_1 and P_2. The greater this distance, the farther the type space is from a common prior. A point equidistant along the line between the nearest points may be regarded as an 'almost common prior'. In the n-player case matters are a bit more involved, but the basic idea is similar.

Given a type space τ and its associated sets of priors $(P_1, \ldots, P_n)$, one for each player i, consider the following bounded, closed, and convex subsets of $\mathbb{R}^{mn}$.

$$X = P_1 \times P_2 \times \cdots \times P_n, \tag{2}$$

and the 'diagonal set'

$$D = \{(p, p, \ldots, p) \in \mathbb{R}^{nm} | p \in \Delta^m\}. \tag{3}$$

Clearly, there is no common prior if and only if $\bigcap_{i=1}^{n} P_i = \emptyset$, if and only if X and D are disjoint.

Definition 4 Let τ be an n-player type space. The *non-normalised prior distance* of τ is $\min_{x \in X, p \in D} \|x - p\|_1$, i.e., the minimal L^1 distance between points in X and points in D. The *prior distance* of τ is $\delta = \frac{\gamma}{n}$, where γ is the non-normalised prior distance of τ,

Definition 5 A probability distribution $p \in \Delta(\Omega)$ is a δ*-almost common prior* of a type space τ of prior distance δ if there is a point $x \in X$ such that the L^1 distance between p and x is the non-normalised prior distance $n\delta$.

It is straightforward that a type space has 0 prior distance if and only if there is a common prior. δ-almost common priors then serve as a 'proxy' for common priors when there is no common prior.

4 The main theorem

We make use of the following generalization of Definition 2.

Definition 6 If a bet $(f_1, \dots, f_n)$ satisfies $\max_\omega f_i(\omega) - \min_\omega f_i(\omega) \le 2$ and $E_i f_i > \delta$ for all i, then it is a *δ-agreeable bet.*

4.1 Proof of the main theorem

Theorem 1 *Let τ be a finite type space. Then the prior distance of τ is greater than δ if and only if there is a δ-agreeable bet.*

Proof When $\delta = 0$ the statement of the theorem reduces to the standard no betting theorem of common priors. We therefore assume in the proof that $\delta > 0$. Notationally, throughout this proof e will denote the vector in $\mathbb{R}^m$ whose every coordinate is 1.

We will need the following generalisation of the Minimum Norm Duality Theorem (see, for example, Dax (2006)): given two disjoint convex sets C_1 and C_2 of $\mathbb{R}^{nm}$,

$$\min_{x_1 \in C_1, x_2 \in C_2} \|x_1 - x_2\|_p = \max_{\{f \in \mathbb{R}^{nm} \mid \|f\|_q \le 1\}} \inf_{x_1 \in C_1} f \cdot x_1 - \sup_{x_2 \in C_2} f \cdot x_2, \qquad (4)$$

where $\|\cdot\|_p$ and $\|\cdot\|_q$ are dual norms, with the maximum in the right-hand side of Equation (4) attained by some $f \in \mathbb{R}^{nm}$ such that $\|f\|_q \le 1$. In broad outline, this theorem is needed for both directions of the proof; if there exists a δ-agreeable bet we use it deduce a lower bound on the distance between X and D, and if the prior distance is greater than δ then the fact that the maximum in the right-hand side of Equation (4) is attained by some f is exploited to construct a δ-agreeable bet.

In one direction, suppose that there exists $f = (f_1, \dots, f_n)$ that is a δ-agreeable bet. Since $\delta > 0$, X and D are separated. Choose a pair of points $(x_1^*, \dots, x_n^*) \in X$ and $(p^*, \dots, p^*) \in D$ at which the minimal L^1 distance between X and D is attained. As f is a δ-agreeable bet, $f_i \cdot x_i^* > \delta$ for each i, hence $\sum_i f_i \cdot x_i^* > n\delta$. Since $\sum_i f_i = 0$, $\sum_i f_i \cdot p^* = 0$, and we have $\sum_i f_i \cdot (x_i^* - p^*) > n\delta$.

Next, note that by assumption $\max_\omega f_i(\omega) - \min_\omega f_i(\omega) \le 2$ for each i. It follows that there is a real number c_i such that, setting $g_i = f_i + c_i e$, $\|g_i\|_\infty \le 1$.

By definition of g_i,

$$\begin{aligned} g_i \cdot (x_i^* - p^*) &= (f_i + c_i e) \cdot (x_i^* - p^*) \\ &= f_i \cdot x_i^* + c_i e \cdot x_i^* - f_i \cdot p^* - c_i e \cdot p^*. \end{aligned}$$

Since both x_i^* and p^* are elements of the simplex, $\sum_\omega x_i^*(\omega) = \sum_\omega p^*(\omega) = 1$. This yields

$$\begin{aligned} c_i e \cdot x_i^* &= c_i (e \cdot x_i^*) \\ &= c_i \left(\sum_{\omega \in \Omega} x_i^*(\omega) \right) \\ &= c_i \end{aligned}$$

and using exactly similar reasoning, $c_i e \cdot p^* = c_i$, yielding $g_i \cdot (x_i^* - p^*) = f_i \cdot (x_i^* - p^*)$ for each i.

We deduce that $\sum_i g_i \cdot (x_i^* - p^*) = \sum_i f_i \cdot (x_i^* - p^*) > n\delta$. At this point we invoke the Minimum Norm Duality Theorem: since $\|g_i\|_\infty \leq 1$ and $\sum_i g_i \cdot (x_i^* - p^*) > n\delta$, the theorem implies that the minimal distance between X and D is greater than $n\delta$, hence the prior distance is greater than δ.

In the other direction, suppose that the prior distance is greater than δ. We need to show the existence of a δ-agreeable bet. By the Minimum Norm Duality Theorem, there exists $f = (f_1, \ldots, f_n) \in \mathbb{R}^{nm}$ such that $\|f_i\|_\infty \leq 1$ for all i at which the maximal distance is attained, i.e., $fx - fy > n\delta$ for all $x = (x_1, \ldots, x_n)$ in X and $y = (p, \ldots, p) \in D$.

The condition $fx - fy > n\delta$ for all $x \in X$ and all $y \in D$ is equivalent to the existence of $b, c, d \in \mathbb{R}$ such that $b - d > n\delta$ and $fx \geq b > d \geq fy$ for all $x \in X$ and all $y \in D$. Rearranging terms, we have $fx - d \geq b - d > n\delta$ and $fy - d < 0$ for all $x \in X$ and all $y \in D$.

Define $g_i := f_i - d/n$ for each i. This yields, for all $x \in X$,

$$g \cdot x = \sum_{i \in I} x_i \cdot (f_i - d/n) = \sum_{i \in I} x_i f_i - \sum_{i \in I} (d/n) = \sum_{i \in I} x_i f_i - d = f \cdot x - d > n\delta.$$

Similarly, for all p in the simplex, with $y = (p, \ldots, p)$,

$$g \cdot y = p \cdot \left(\sum_{i \in I} g_i \right) = p \cdot \left(\sum_{i \in I} (f_i - d/n) \right) = f \cdot y - d < 0.$$

Since the last inequality holds for all p in the simplex, we deduce that $\sum_i g_i < 0$. Furthermore, since the coordinates of x_i are non-negative, increasing the coordinates of the g_i does not change the inequality $\sum_i g_i x_i > n\delta$, hence we may assume that g satisfies $\sum_i g_i = 0$.

The fact that $\sum_i g_i x_i > n\delta$ still leaves the possibility that $g_i x_i < \delta$ for some i. But let x_i^* be the point that minimises $g_i x_i$ over P_i, for each i. Since $\sum_{i=1}^n g_i x_i^* > n\delta$, there are constants c_i guaranteeing $x_i^* g_i + c_i > \delta$ for each i, satisfying $\sum_i c_i = 0$. Define $h_i = g_i + c_i e$. Then $\max_\omega h_i(\omega) - \min_\omega h_i(\omega) \leq 2$ for each i, $\sum_i h_i = \sum_i g_i = 0$, $\sum_i h_i x_i > n\delta$, and for each $x_i \in P_i$, $h_i x_i > \delta$. □

Corollary 1 *If the prior distance of a finite type space is δ, then there is no bet $(f_1, \ldots, f_n)$ satisfying $\|f_i\|_\infty \leq 1$ and $E_i f_i > \delta$ for all i.*

Remark 1 Although we have chosen to use the L^1 norm to measure prior distance, from a purely mathematical perspective, this choice of norm is arbitrary. The proof of the Theorem 1 is based on a form of the Minimal Norm Duality Theorem that holds for any pair of conjugate real numbers p and q. With appropriate changes to Definitions 4 and 5, the proof and the statement of Theorem 1 could be rewritten in terms of any pair of dual norms.

4.2 Agreeing to disagree, but boundedly

Corollary 2 *Let τ be a finite, two-player type profile with δ prior distance, and let $\omega^* \in \Omega$. Let $f \in \mathbb{R}^\Omega$ be a random variable, and let $\eta_1, \eta_2 \in \mathbb{R}$. If it is common knowledge at ω^* that player 1's expected value of f is greater than or equal to η_1, and player 2's expected value of f is less than or equal to η_2, then*

$$|\eta_1 - \eta_2| \leq 2\delta \, \|f\|_\infty .$$

Proof Suppose that $|\eta_1 - \eta_2| > 2\delta \, \|f\|_\infty$. Let $g = \frac{f}{\|f\|_\infty}$. Then g satisfies $\|g\|_\infty \leq 1$, hence $\max_\omega f_i(\omega) - \min_\omega f_i(\omega) \leq 2$, yet $|E_1(g(\omega)) + E_2(-g(\omega))| > 2\delta$ for all ω, contradicting the assumption that the prior distance of τ is δ. □

This also leads to a generalisation of the No Disagreements Theorem of Aumann (1976), to which it reduces when $\delta = 0$.

Corollary 3 *Let τ be a finite, two-player type profile of δ prior distance, and let $\omega^* \in \Omega$. Let H be an event. If it is common knowledge at ω^* that $E_1(H \mid \omega) = \eta_1$ and $E_2(H \mid \omega) = \eta_2$, then $|\eta_1 - \eta_2| \leq \delta$.*

Proof Let $f \in \mathbb{R}^\Omega$ satisfy $0 \leq f(\omega) \leq 1$ for all $\omega \in \Omega$. Then it cannot be the case that $|E_1(f(\omega)) + E_2(-f(\omega))| > \delta$ for all ω. Suppose by contradiction that this statement holds. Let $g = 2f - 1$. Then $|E_1(g(\omega)) + E_2(-g(\omega))| > 2\delta$ for all ω, contradicting the assumption that the prior distance of τ is δ.

Consider the standard characteristic function 1^H. Since $0 \leq 1^H(\omega) \leq 1$ for all ω and the expected value of 1^H at every state is the probability of the event H at that state, the conclusion follows. □

5 Getting to agreement

What happens when partition profiles are refined? It is straight-forward to show that such refinements can only increase the set of priors, and therefore the prior distance can only decrease. More formally, let $\tau_0 \prec \tau_1 \prec \cdots \prec \tau_k$ be a sequence of successive proper refinements of type spaces with $(\delta_0, \delta_1, \ldots, \delta_k)$ the corresponding sequence of prior distances of the refinements, Then $\delta_0 \geq \delta_1 \geq \cdots \geq \delta_k$. In words, 'increasing information can never increase (common knowledge) disagreements'.

This naturally leads to the questions: can refinements always lead to a common prior, and if so, how many refinements are needed to attain a common prior? Theorem 2 answers these questions.

Theorem 2 *Let τ_0 be a positive type profile with δ_0 prior distance, and let $\tau_0 \prec \tau_1 \prec \cdots$ be a sequence of successive proper refinements, with $\delta_0 \geq \delta_1 \geq \cdots$ the corresponding sequence of prior distances of the refinements. Then there is a $k \leq (|I| - 1)|\Omega| - |\tau_0| + 1$ such that $\delta_k = 0$, i.e., after at most k proper refinements, there is a common prior.*

Proof We make use of the following result from Hellman and Samet (2012): for a partition profile Π, if

$$|\Pi| = (|I| - 1)|\Omega| + C(\Pi), \tag{5}$$

then any type profile over Π has a common prior.[4]

Suppose that we start with a type profile τ_0 of size $|\tau_0|$. In the 'worst case', the size of the type spaces in the sequence of refinements $\tau_0 \prec \tau_1 \prec \cdots$ increases by only one at each step; i.e., $|\tau_{j+1}| = |\tau_j| + 1$. If a given refinement in this sequence increases the number of common knowledge components, it can only add 1 to the total number of common knowledge components while at the same time the number of partition elements has increased by 1, hence this makes no difference for the number of steps remaining towards the attainment of a common prior.

It follows that for $q = (|I| - 1)|\Omega| - |\tau_0| + 1$, the partition profile of τ_q will satisfy the condition in Equation 5, and hence τ_q will have a common prior. If the size increases by more than one in some steps, the condition will be satisfied at some $k \leq q$, and then τ_k will be guaranteed to have a common prior. □

Note that:

1. It does not matter what the prior distance δ_0 is when we start the process of successive refinements; we will always get to a common prior.
2. Perhaps surprisingly, the upper bound on the number of successive refinements needed to attain a common prior is entirely independent of δ_0 and depends only on the total number of initial partition elements. This means that no matter how far apart the players start out and what their initial type functions are, they are guaranteed to attain a common prior within k refinements.

Geanakoplos (1994) presents a version of the well-known envelopes problem in which the players refrain from betting, not because their posteriors are derived from common priors, but because they know that the posteriors could have been derived from a common prior and hence they know they cannot disagree. What we have here is an extension of this principle to all type profiles: what count for bounding disagreements are the almost common priors. All the data needed for bounds on disagreements can be known from the posterior probabilities, without reference to a prior stage. Indeed, even if there was historically a prior stage, one is better off ignoring the historical priors and considering instead the almost common priors.

Acknowledgments This research comprises part of the author's Ph.D. thesis conducted under the supervision of Professor Dov Samet of Tel Aviv University and Professor Sergiu Hart of the Hebrew University of Jerusalem. It was supported in part by funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013/ERC Grant Agreement No. 249159). The author would like to thank his Ph.D. advisers for their guidance in the composition of this research, and two anonymous referees for many helpful comments. The author is also grateful to the Centre for the Study of Rationality at the Hebrew University for its support and assistance.

[4] Recall that $C(\Pi)$ denotes the number of common knowledge components in the meet of Π.

References

Aumann RJ (1976) Agreeing to disagree. Ann Stat 4(6):1236–1239
Aumann RJ (1987) Correlated equilibrium as an expression of Bayesian rationality. Econometrica 55:1–18
Bonanno G, Nehring K (1999) How to make sense of the common prior assumption under incomplete information. Int J Game Theory 28:409–434
Dax A (2006) The distance between two convex sets. Linear Algebra Appl 416:184–213
Feinberg Y (2000) Characterizing common priors in the form of posteriors. J Econ Theory 91:127–179
Geanakoplos J (1994) Common knowledge, handbook of game theory, vol 2. Elsevier Science B.V., The Netherlands, pp 1439–1496
Harsányi JC (1967–1968) Games with incomplete information played by Bayesian players. Manag Sci 14, 159–182, 320–334, 486–502
Hellman Z, Samet D (2012) How common are common priors. Games Econ Behav 74:517–525
Morris S (1995a) Trade with heterogenous prior beliefs and asymmetric information. Econometrica 62:1327–1347
Morris S (1995b) The common prior assumption in economic theory. Econ Philos 11:227–253
Samet D (1998) Common priors and separation of convex sets. Games Econ Behav 24:172–174
Samuelson L (2004) Modeling knowledge in economic analysis. J Econ Lit 42:367–403

CHAPTER 3

Iterated Expectations, Compact Spaces and Common Priors

Published in Games and Economic Behavior, 72 (2011), pp. 163 – 171.
The paper appears here on pages 29 – 37.

Games and Economic Behavior 72 (2011) 163–171

Contents lists available at ScienceDirect

Games and Economic Behavior

www.elsevier.com/locate/geb

Iterated expectations, compact spaces, and common priors ☆

Ziv Hellman

The Department of Mathematics and The Centre for the Study of Rationality, The Hebrew University of Jerusalem, Giv'at Ram, Jerusalem 91904, Israel

ARTICLE INFO

Article history:
Received 20 June 2007
Available online 31 July 2010

JEL classification:
C70
D82

Keywords:
Common priors
Iterated expectations

ABSTRACT

Extending to infinite state spaces that are compact metric spaces a result previously attained by D. Samet solely in the context of finite state spaces, a necessary and sufficient condition for the existence of a common prior for several players is given in terms of the players' present beliefs only. A common prior exists if and only if for each random variable it is common knowledge that all Cesàro means of iterated expectations with respect to any permutation converge to the same value; this value is its expectation with respect to the common prior. It is further shown that compactness is a necessary condition for some of the results.

1. Introduction

The common prior assumption, ever since it was introduced into the study of games with incomplete information by Harsányi (1967–1968), posits that all women and men are "created equal" with respect to probability assessments in the absence of information — hence the term common prior — and all differences in probabilities should, in principle, be traced to asymmetries in information received over time. The idea has become very pervasive, and in most applications of type spaces to economics it is assumed that players' beliefs can indeed be derived from a common prior by Bayesian updating. A prior probability can be interpreted as the beliefs of a player in a previous period. In many models, however, any previous period is either fictional or irrelevant to the matter being studied. It is also clear that there are many plausible models of type spaces in which it is impossible for the players to have arrived at their current beliefs via updating from a common prior. This leads naturally to the question of whether a criterion can be identified by which one can tell, using the current beliefs of the players, that they have a common prior.

Aumann (1976), in his celebrated No Disagreements theorem, presented a necessary condition for the existence of a common prior in terms of present beliefs: if there is a common prior, then it is impossible to have common knowledge of difference in the beliefs of any given event. Numerous authors extended this result and applied it to interactions between agents in various situations. The typical result is a "no-bet" or "no-trade" theorem (for example, Milgrom and Stokey, 1982, or Sebenius and Geanakoplos, 1983) — agents who start with common prior distributions will never agree to engage in speculative trade based on differences in private information that they subsequently receive. As soon as it becomes common knowledge that they wish to trade, their expectations for the value of assets in question become identical.

In the 1990s, it was shown, by several researchers independently, that the converse statement also holds, thus leading to a characterization of the existence of common priors that may be termed the No Betting characterization: a necessary and sufficient condition for the existence of a common prior is that there is no bet for which it is always common knowledge

☆ The author is grateful to Dov Samet and Ariel Rubinstein for their encouragement and assistance. Thanks are also due to an anonymous referee, and an associate editor of this journal, whose comments vastly improved the presentation here. This paper contains material from the author's doctoral thesis. An earlier version of the paper appeared as Discussion Paper 522 of the Hebrew University's Centre for the Study of Rationality.

E-mail address: ziv.hellman@mail.huji.ac.il.

0899-8256/$ – see front matter
doi:10.1016/j.geb.2010.06.012

that all players expect a positive gain. The most accessible proof of this result is in Samet (1998b). It was proved by Morris (1994) for finite type spaces and independently by Feinberg (2000) for compact type spaces. Bonanno and Nehring (1999) proved it for finite type spaces with two agents.

As Samet (1998a) points out, this characterization satisfactorily solves the question of how one can tell when players have a common prior, but it fails to express the common prior in a meaningful way; the fact that disagreement regarding the outcome of a bet cannot be common knowledge may guarantee the existence of a common prior, but it tells us nothing about this common prior. He then proceeds to present a very different necessary and sufficient condition for the existence of a common prior that not only identifies the common prior when it exists, but also provides an epistemically meaningful interpretation to it.

This condition is expressed intuitively in Samet (1998a) in a colorful story. Imagine that Adam and Eve, who have both excelled in their studies at the same school of economics, are asked what return they expect on IBM stock. Having been exposed to different sources of information, we should not be surprised if the two provide different answers. But we can then go on to ask Eve what she thinks Adam's answer was. Being a good Bayesian, she can compute the expectation of various answers and come up with Adam's expected answer. Likewise, Adam can provide us with what he expects was Eve's answer to that question. This process can continue, moving back and forth between Eve and Adam, theoretically forever. There are, in this example, two possible infinite sequences of alternating expectations, one that starts with Eve and one that starts with Adam. Samet calls this process "obtaining an iterated expectation", and shows that, when the relevant state space is finite, there exists a common prior if and only if both of these sequences converge to the same limit.

He achieves this result by representing Adam's beliefs[1] by a type matrix M_1 and Eve's beliefs by a type matrix M_2. These then form two permutation matrices, $M_{\sigma_1} = M_2 M_1$, which is intended to be used for the process of obtaining iterated expectations starting with Adam, and $M_{\sigma_2} = M_1 M_2$ which does the same for the iterated expectations starting with Eve. It then turns out to be the case that both M_{σ_1} and M_{σ_2} are ergodic Markov matrices, and therefore each of them has a unique invariant probability measure, which may be labelled respectively p_1 and p_2. It is then shown that if $p_1 \neq p_2$, Adam and Eve cannot share a common prior. On the other hand, if $p_1 = p_2$, then not only is there a common prior, it has positively been identified — it is precisely $p := p_1 = p_2$.

We can term the criterion by which a common prior is ascertained to exist, by the identities of the invariant probability measures associated with permutation matrices, the *iterated expectations characterization*. Samet (1998a), however, proves it only in the context of finite state spaces. Given the results in Feinberg (2000) and Heifetz (2006), which extend the No Betting characterization to compact state spaces, it is natural to wonder whether an analogue of Samet's characterization can also be shown to hold in compact state spaces.

It is the goal of this paper to show that there is an affirmative answer to that question. The significance of such a result is clear, given that there are many models of interest which involve infinite state spaces and cannot be reduced to a finite space — we therefore extend the application of the iterated expectations characterization to many models to which it previously could not be applied.

It is also shown here, by way of an example, that compactness is necessary in the sense that if one does not assume compactness, the infinite dimensional analogue of the permutation matrix M_σ may not have a well-defined invariant probability measure — and without that, the subsequent propositions do not follow, and indeed in that case there is no guarantee that the iterated expectations characterization for checking the existence of a common prior can even be applied intelligibly, as there may not be invariant probability measures that can be compared against each other.

It should also be noted that the iterated expectations characterization is significant because it provides, in principle, a way of calculating a common prior given a type space. In the finite state space context, one can form the type matrices and apply numerical solutions for calculating invariant probability measures in Markov chains — a subject of active research — in order to ascertain whether or not there is a common prior and if one exists, to identify it. Similarly, with the extension here of the iterated expectations characterization to the more general compact spaces, it now becomes possible, given knowledge of the players' type spaces, in principle to estimate the expected values of random variables by use of numerical solutions, such as those appearing in e.g. Hernández-Lerma and Lasserre (2003).

The following rough correspondences exist between results in this paper and those that appear in Samet (1998a), save for the fact that the results in that paper are strictly limited to finite state spaces, whereas that restriction is lifted here: Proposition 1 here is (roughly) an infinite state space version of Proposition 4 of Samet (1998a); Proposition 2 here corresponds to Proposition 5 of Samet (1998a); and similarly Proposition 3 corresponds to Proposition 2′ and Proposition 4 to Theorem 1′.

2. Preliminary definitions and results

2.1. Type spaces

A *type space* for a set of players is a tuple $\langle I, \Omega, \mathcal{F}, \varphi, (\Pi_i, t_i)_{i \in I} \rangle$. The set of players is denoted by $I = \{1, \ldots, n\}$, where $n \geqslant 2$. Ω is a measurable space of arbitrary cardinality, whose elements are called states. $\mathcal{F}$ is a σ-field of *events* (subsets

[1] For the sake of simplicity here, we will make the mild technical assumption that the entire relevant state space is the meet of the Adam and Eve type space.

of Ω), and φ a non-trivial σ-finite measure. For each player $i \in I$, Π_i is a partition of Ω, which may be termed player i's *knowledge partition*, and $t_i(\cdot,\omega)$ denotes a belief — or probability measure on $(\Omega,\mathcal{F})$ — associated with each player i at each state. We further assume that each element of each partition Π_i is an element of $\mathcal{F}$ (and therefore that the atoms of the knowledge partitions of each player are $\mathcal{F}$-measurable), and that $\varphi(\Pi_i) > 0$. The collection $(\Pi_i)_{i\in I}$ is termed a *partition profile*, and will sometimes be denoted here by $\boldsymbol{\Pi}$.

The probability measures $t_i(\cdot,\omega)$ for each player i and each state ω are required to satisfy:

1. $t_i(\Pi_i(\omega)|\omega) = 1$.
2. For all $\omega' \in \Pi_i(\omega)$, $t_i(A,\omega') = t_i(A,\omega)$.

The *meet* of $\boldsymbol{\Pi}$, denoted $\wedge\boldsymbol{\Pi}$, is the partition of Ω that is the finest among all partitions that are coarser than Π_i for each i. For each ω, $\wedge\boldsymbol{\Pi}(\omega)$ denotes the element of the meet containing ω. A somewhat more constructive way to define the elements of the meet utilizes the concept of reachability. A state ω' is *reachable* from ω if there exists a sequence $\omega_0 = \omega, \omega_1, \omega_2, \ldots, \omega_m = \omega'$ such that for each $k \in \{0,1,\ldots,m-1\}$, there exists a player i_k such that $\Pi_{i_k}(\omega_k) = \Pi_{i_k}(\omega_{k+1})$. It is well known that $\omega' \in \wedge\boldsymbol{\Pi}(\omega)$ iff ω' is reachable from ω, and therefore the relation of reachability can be used to define the meet. This characterization of the meet will be used in proofs in the body of this paper. For an event A, the event that A is *common knowledge* is the union of all the elements of $\wedge\boldsymbol{\Pi}$ contained in A.

A *random variable* is a real-valued function on Ω. For a probability measure ν and a random variable f on Ω, the expectation of f with respect to ν is $\nu f := \int_\Omega f(\omega)\,d\nu(\omega)$. For each player i and random variable f, i's expectation of f, denoted $E_i f$ is the random variable

$$(E_i f)(\omega) := \int_\Omega f(\overline{\omega})\,dt_i(\overline{\omega}|\omega).$$

Given a type space, one can ask whether the space might have come to exist, in its current state, from a space with no information at all, by the players acquiring new information over time and updating their beliefs in a Bayesian manner. Each player's possible initial belief on the no-information primeval space is called a prior. In general, given player i's current type, there will not be a single prior, but a set of possible priors. A main question is then whether or not the players have a common prior.

More formally, a probability measure μ over $(\Omega,\mathcal{F})$ is a *prior* for player i if for every event $A \in \mathcal{F}$,

$$\mu(A) = \int_\Omega t_i(A|\omega)\,d\mu(\omega).$$

A probability measure is a *common prior* if it is a prior for each of the players $i \in I$.

2.2. Markov transitions

When working with a finite state space, a Markov chain is typically represented by a series of random variables $\{X_1, X_2, \ldots\}$ along with a transition matrix M, such that the (i,j)-th element of M is the probability that $X_{n+1} = j$ given that $X_n = i$.

In transferring this idea to a more general state space, we cannot always expect to measure the probability that the value of a random variable in a successive time period will be a specific state, but we can ask what the probability is that it will be in an event. In formulae, if $(\Omega,\mathcal{F})$ is a measurable space, $(X,\mathcal{B},\mathcal{P})$ a probability space, E an event in $\mathcal{F}$, and $\{\zeta_1,\zeta_2,\ldots\}$ is a sequence of Ω-valued random variables defined on X, our analogue of the transition matrix will be given by $M(E|\omega) := \mathcal{P}(\zeta_{n+1} \in E|\zeta_n = \omega)$. This motivates the standard definition of a general Markov transition probability function:

A *stochastic kernel* or *Markov transition probability function* on $(\Omega,\mathcal{F})$ is a function M such that

1. $M(\cdot|\omega)$ is a probability measure for each fixed $\omega \in \Omega$.
2. $M(E|\cdot)$ is an $\mathcal{F}$-measurable function on Ω for each fixed event $E \in \mathcal{F}$.

One of the most important aspects of finite state Markov transitions is the interpretation of the n-th power of a Markov transition matrix as representing the n-th step of iterating the transition probabilities encoded in the matrix. The analogue in general state spaces iterates a Markov transition probability function M by the following recursive definition:

$$M^n(E|\omega) = \int_\Omega M^{n-1}(E|\omega')\,dM(\omega'|\omega) = \int_\Omega M(E|\omega')\,dM^{n-1}(\omega'|\omega)$$

for all $E \in \mathcal{F}$ and $\omega \in \Omega$.

In the rest of this section, fix a Markov transition probability function M.

Let $\Delta(\Omega)$ denote the space of probability measures on Ω, with this space naturally outfitted with the induced weak* topology. It is possible to regard M as a function from $\Delta(\Omega)$ to $\Delta(\Omega)$, as follows: For each $\nu \in \Delta(\Omega)$, let

$$(\nu M)(E) := \int_{\Omega} M(E|\omega)\, d\nu(\omega).$$

Then M acts on $\Delta(\Omega)$ by way of $\nu \mapsto \nu M$. Using this notation, a probability measure ν is *invariant* with respect to M if $\nu = \nu M$. If such a measure exists, M is said to admit an invariant probability measure.

The transition probability function M can also be considered as operating on bounded functions in the following way. For each bounded integrable function $f : \Omega \to \mathbb{R}$, let Mf be the bounded function

$$Mf(\omega) := \int_{\Omega} f(\widehat{\omega})\, dM(\widehat{\omega}|\omega).$$

If ν is an invariant probability measure with respect to M, then M can also be considered to be a linear operator on $L_1(\nu) := L_1(\Omega, \mathcal{F}, \nu)$ into itself. We can then define, for any k and $f \in L_1(\nu)$

$$M^k f(\omega) := \int_{\Omega} f(\widehat{\omega})\, dM^k(\widehat{\omega}|\omega).$$

We have in addition the concept of the Cesàro mean, defined as

$$M^{(n)} f(\omega) := \frac{1}{n} \sum_{k=0}^{n-1} M^k f(\omega).$$

If Ω has a topology τ, denote the class of bounded continuous functions with respect to τ from Ω to $\mathbb{R}$ by $C(\Omega)$. Then M satisfies the *weak Feller property* if M maps $C(\Omega)$ to $C(\Omega)$.

Given the underlying space $(\Omega, \mathcal{F}, \varphi)$, a Markov transition function M is *φ-irreducible* if

$$\sum_{n=1}^{\infty} M^n(E|\omega) > 0$$

for all $\omega \in \Omega$ whenever $\varphi(E) > 0$ for $E \in \mathcal{F}$.

We will make use of the following important theorems from the theory of Markov chains. These three theorems appear in Hernández-Lerma and Lasserre (2003) respectively as Theorem 7.2.3, Proposition 4.2.2, and an amalgam of Theorem 2.3.4, Proposition 2.4.2 and Proposition 2.4.3:

THEOREM (Existence of invariant probability measure). Let Ω be a compact metric space, and let M be a Markov transition function on Ω. Then M admits an invariant probability measure.

THEOREM (Uniqueness of invariant probability measure). Let M be a φ-irreducible Markov transition function and suppose that M admits an invariant probability measure ν. Then ν is the unique invariant probability measure for M.

THEOREM (Birkhoff's Ergodic Theorem for Markov processes). Let M be a Markov transition function that admits an invariant probability measure ν. For every $f \in L_1(\nu)$ there exists a function $f^* \in L_1(\nu)$ such that

$$P^{(n)} f \to f^* \nu \text{ almost everywhere}$$

and

$$\int f^*\, d\nu = \int f\, d\nu.$$

In addition, if ν is the unique invariant probability measure of M, then f^* is constant ν-almost everywhere, and $f^* = \int f\, d\nu$, ν-almost everywhere, so

$$\text{the time-average } \lim_{n\to\infty} M^{(n)} f = \text{the space-average } \int f\, d\nu,\ \nu\text{-a.e.}$$

3. Type spaces within the Markov framework

3.1. Relating type spaces to Markov processes

In this section, we relate the concepts of type spaces and Markov processes (similarly to the way this is accomplished in Samet, 2000).

First, note that by definition, the probability measure $t_i(\cdot|\cdot)$ associated with each player i satisfies the conditions for being a Markov transition probability function, hence we can relate to it as such. We will relabel $t_i(\cdot|\cdot)$ as M_i in the sequel when we wish to emphasize we are treating it as a Markov transition.

In general, given any two probability measures P_1 and P_2, one can define a new probability measure $P_2P_1(E|\omega)$ by

$$P_2P_1(E|\omega) = \int_\Omega P_2(E|\widehat{\omega})\, dP_1(\widehat{\omega}|\omega).$$

This obviously can be iterated any number of times. In particular, given a measure P, we can construct an infinite sequence of measures $\{P^n(\cdot|\omega)_{n \geqslant 1}\}$. In our specific context, given any two players i and j and a measurable event E, the probability measure $t_it_j(E|\omega)$ based on t_i and t_j is similarly defined by

$$t_it_j(E|\omega) = \int_\Omega t_i(E|\widehat{\omega})\, dt_j(\widehat{\omega}|\omega).$$

In particular, given an element σ in *Sym*(I), the set of all permutations of the elements of I, define

$$t_\sigma := t_{\sigma(1)} \ldots t_{\sigma(n)}$$

iteratively, by using the above to define $t_{\sigma(n-1)}t_{\sigma n}$, then $t_{\sigma n-2}(t_{\sigma(n-1)}t_{\sigma n})$, and so on.

We can now re-interpret various notions relating to a type space within the Markov framework. First, note that for any function f on the state space, $M_i f$ is precisely the expectation of f in player i's estimation (cf. Samet, 2000). This expectation is what is usually considered of economic significance and importance, as players choose their actions by comparing the relative expectations of functions.

Complementary to this, an invariant probability measure ν with respect to the Markov chain M_i is precisely a prior of player i. A common prior is a probability measure that is simultaneously invariant with respect to all $\{M_i\}_{i \in I}$.

A sequence $s = (i_1, i_2, \ldots)$ of elements of I is called an I-sequence if for each player j, $i_k = j$ for infinitely many ks. The *iterated expectation* of a random variable f with respect to the I-sequence s is the sequence of random variables $\{M_{i_k} \ldots M_{i_1} f\}_{k=1}^{\infty}$.

Given the identification of $E_i f$ with $M_i f$, we can write, given a permutation σ of I,

$$M_\sigma := E_\sigma := t_\sigma := E_{\sigma_1} \ldots E_{\sigma_n} = M_{\sigma_1} \ldots M_{\sigma_n} = t_{\sigma_1} \ldots t_{\sigma_n}$$

and term this a *permutation chain.*

The *iterated expectation of f with respect to σ* is the sequence $\{E_\sigma^k f\}_{k=1}^{\infty}$, and the *Cesàro iterated expectation of f with respect to σ* is $\{E_\sigma^{(k)} f\}_{k=1}^{\infty}$. The iterated expectation of f with respect to σ is the iterated expectation of f with respect to the I-sequence

$$\sigma_1, \ldots, \sigma_n, \sigma_1, \ldots, \sigma_n, \ldots$$

as defined above.

It should be noted here that the results in this paper do not extend all the results of Samet (1998a) to compact metric spaces. To be precise, the claims of that paper, in the finite type space context, show that the existence of a common prior implies that for each random variable f it is common knowledge in each state that *all* the iterated expectations of f, with respect to all I-sequences s, converge to the same limit. The claims of this paper show that, in the context of a compact e.m.p. type space (as defined in the next section), the existence of a common prior implies that for each random variable f it is common knowledge in each state that the Cesàro iterated expectations of f *with respect to each permutation* converge to the same limit, but not with respect to all I-sequences.

3.2. Everywhere mutually positive type space

A type space $\langle I, \Omega, \mathcal{F}, \varphi, (\Pi_i, t_i)_{i \in I} \rangle$ with a topology τ over Ω will be termed *everywhere mutually positive (e.m.p.)* if it satisfies the conditions:

1. For each state $\omega \in \Omega$, there exists an event $A(\omega) \ni \omega$ such that $A(\omega) \subseteq \cap_i \Pi_i(\omega)$, and $\varphi(A(\omega)) > 0$.
2. For all $i \in I$, each state $\omega \in \Omega$, and every event $A \ni \omega$ such that $\varphi(A) > 0$, the inequality $t_i(A|\omega) > 0$ is satisfied.
3. The correspondence $\omega \mapsto t_i(\cdot|\omega)$ is continuous with respect to the topology τ and the weak topology on $\Delta(\Omega)$, for every $i \in I$.

The intuitive reason for working with everywhere mutually positive spaces is that we wish to relate together the main elements with which we are working, namely the partitional structure, the type probabilities, the topology, and the underlying measure on the space. Consider, for example, three states $\omega_1, \omega_2, \omega_3$, that are reachable one from the other by way of $\omega_1, \omega_2 \in \Pi_i(\omega_1)$ and $\omega_2, \omega_3 \in \Pi_j(\omega_3)$, for players $i \neq j$. As these states are connected from the perspective of the partitional structure, we want them to be "connected" also in the sense of Markov transitions, that is, intuitively speaking, we want there to be a positive probability of transitioning to ω_3 from ω_1. This means avoiding situations in which, e.g., there is non-zero probability of transitioning from ω_2 to ω_3 according to t_j, but the transition from ω_1 "gets stuck" because t_i assigns zero probability to transitioning from ω_1 to ω_2, or because the underlying measure φ assigns zero probability to every event in $\Pi_i(\omega_2) \cap \Pi_j(\omega_2)$. The first two properties in the definition of an e.m.p. are together intended to avoid various such difficulties. The third property relates the topology τ and the type spaces in a standard continuity requirement.

Note that when Ω is finite and Π is positive, meaning that $t_i(\omega|\omega) > 0$ for all i and all ω, the corresponding type space, using the standard counting measure, trivially satisfies the conditions of being everywhere mutually positive. Also note that from previous definitions it follows that

$$\int_\Omega f(\widehat{\omega})\, dt_i(\widehat{\omega}|\omega)$$

is always continuous in ω for every $f \in C(\Omega)$.

If in addition to the above conditions, (Ω, τ) is compact metric space, the type space $\langle I, \Omega, \varphi, \mathcal{F}, \tau, (\Pi_i, t_i)_{i \in I}\rangle$ will be called a *compact e.m.p. space* for short. Nearly all the results in this paper will henceforth assume a compact e.m.p. type space. For notational ease, $\langle I, \Omega, \varphi, \mathcal{F}, \tau, (\Pi_i, t_i)_{i \in I}\rangle$ will be written simply as (Ω, τ).

4. Common priors and compact e.m.p. type spaces

Given any $Q \in \Pi$, the restriction of M_i to Q, for any player i, will be written as M_i^Q. Given a permutation σ in $Sym(I)$, the restriction of M_σ to Q is similarly denoted by M_σ^Q.

Lemma 1. *Given a type space satisfying properties* (1) *and* (2) *of e.m.p. type spaces, for any permutation σ of I and player i, and for any arbitrary pair of states $\omega, \widehat{\omega} \in \Pi_{\sigma(i)}(\omega)$, there is a φ-non-null event $A(\omega)$ such that $t_\sigma(A(\omega)|\widehat{\omega}) > 0$.*

Proof. Let $\omega, \widehat{\omega} \in \Pi_{\sigma(i)}(\omega)$. By property (1) of e.m.p. type spaces, there exist φ-non-null events $A(\omega) \ni \omega$ and $A(\widehat{\omega}) \ni \widehat{\omega}$ such that $A(\omega) \subseteq \cap_j \Pi_j(\omega)$ and $A(\widehat{\omega}) \subseteq \cap_j \Pi_j(\widehat{\omega})$. Let $i < j \leqslant n$; by property (2), $t_{\sigma(j)}(A(\widehat{\omega})|\widehat{\omega}) > 0$, and similarly, for any $1 \leqslant k < i$, we have $t_{\sigma(k)}(A(\omega)|\omega) > 0$ for $1 \leqslant k < i$. From $t_{\sigma(i)}(A(\widehat{\omega})|\widehat{\omega}) > 0$ and $t_{\sigma(i)}(A(\omega)|\widehat{\omega}) = t_{\sigma(i)}(A(\widehat{\omega})|\widehat{\omega})$ (which holds since $\omega, \widehat{\omega} \in \Pi_{\sigma(i)}(\omega)$), it follows that $t_{\sigma(i)}(A(\omega)|\widehat{\omega}) > 0$.

We now unravel the recursive definition of $t_{\sigma(1)} \ldots t_{\sigma(n)}$. For $i < j \leqslant n$, suppose that $t_{\sigma(j)} \ldots t_{\sigma(n)}(A(\widehat{\omega})|\widehat{\omega}) > 0$ (which is certainly true when $j = n$). Then

$$t_{\sigma(j-1)} t_{\sigma(j)} \ldots t_{\sigma(n)}\big(A(\widehat{\omega})|\widehat{\omega}\big) = \int_\Omega t_{\sigma(j-1)}\big(A(\widehat{\omega})|\omega'\big)\, d(t_{\sigma(j)} \ldots t_{\sigma(n)})\big(\omega'|\widehat{\omega}\big)$$

But the facts that $t_{\sigma(j)} \ldots t_{\sigma(n)}(A(\widehat{\omega})|\widehat{\omega}) > 0$, that $A(\widehat{\omega}) \subseteq \Pi_{\sigma(j-1)}(\widehat{\omega})$, and that $t_{\sigma(j-1)}(A(\widehat{\omega})|\omega') = t_{\sigma(j-1)}(A(\widehat{\omega})|\widehat{\omega}) > 0$ for all $\omega' \in \Pi_{\sigma(i)}(\widehat{\omega})$, taken all together, imply that $t_{\sigma(j-1)} \ldots t_{\sigma(n)}(A(\widehat{\omega})|\widehat{\omega}) > 0$.

Similar reasoning can be applied at the transition point from $t_{\sigma(i+1)} \ldots t_{\sigma(n)}$ to $t_{\sigma(i)} \ldots t_{\sigma(n)}$, and the transition point from $t_{\sigma(k)} \ldots t_{\sigma(n)}$ to $t_{\sigma(k-1)} \ldots t_{\sigma(n)}$ for $1 \leqslant k < i$, to conclude that $t_\sigma(A(\omega)|\widehat{\omega}) > 0$. □

Proposition 1. *For any permutation σ of I, and for any element Q of the meet of a compact e.m.p. type space (Ω, τ), M_σ^Q has a unique invariant probability measure π_σ^Q.*

Proof. By the assumed properties of an e.m.p. type space, $M_{\sigma(i)}^Q$, for any i, satisfies the weak Feller property. The weak Feller property of M_σ^Q follows readily from the concatenation formation via $M_{\sigma(1)}^Q \ldots M_{\sigma(n)}^Q$. The compactness of the metric topology τ then guarantees the existence of at least one invariant probability measure π_σ^Q over M_σ^Q by application of a theorem cited in Section 2.2.

Next, select an event $E \subseteq Q$ such that $\varphi(E) > 0$, and a state $\omega' \in E$. Let $\omega \in Q$ be selected arbitrarily. Since $\omega \in \wedge \boldsymbol{\Pi}(\omega')$, there exists a sequence $\{\omega = \omega_0, \omega_1, \omega_2, \ldots, \omega_m = \omega'\}$ such that for each $k \in \{0, 1, \ldots, m-1\}$, there exists a player i_k such that $\Pi_{i_k}(\omega_k) = \Pi_{i_k}(\omega_{k+1})$.

We can now define the following iterative process: by definition, there is a player i_0 such that $\Pi_{i_0}(\omega_0) = \Pi_{i_0}(\omega_1)$. At step 0 of the iterative process, we conclude from Lemma 1 the existence of a φ-non-null event $A(\omega_1)$ such that $t_\sigma(A(\omega_1)|\omega_0) > 0$.

At step $j > 0$, there is a player i_j such that $\Pi_{i_j}(\omega_j) = \Pi_{i_j}(\omega_{j+1})$. Applying the same reasoning as above, there is a φ-non-null event $A(\omega_{j+1}) \subseteq \Pi_{i_j}(\omega_{j+1}) \cap \Pi_{i_{j+1}}(\omega_{j+1})$ such that $t_\sigma(A(\omega_{j+1})|\omega_j) > 0$, and furthermore by concatenating the results of steps previous to step j, $t_\sigma^{j+1}(A(\omega_{j+1})|\omega_0) > 0$.

At the end of the process, the conclusion is $t_\sigma^m(A(\omega_m)|\omega_0) > 0$, with $\omega_m \in A(\omega_m) \subseteq \Pi_{i_{m-1}}(\omega_m)$. Finally, from $\varphi(E) > 0$ and the assumption that the type space is an e.m.p. space, we have $t_{\sigma(k)}(E|\omega' = \omega_m) > 0$ for all $k \in I$. By a slight tweaking of the proof of Lemma 1, using the fact that $\omega' \in A(\omega') \cap E$ and that for all k, $t_{\sigma(k)}(A(\omega')|\omega')) > 0$, we can show that $t_\sigma(E|\omega') > 0$ and then that $t_\sigma^{m+1}(E|\omega_0) > 0$. We thus conclude that M_σ^Q is φ-irreducible, hence π_σ^Q is unique. □

Proposition 2. *For a compact e.m.p. type space Ω, the following are equivalent.*

1. *π is a common prior on Ω.*
2. *π is an invariant probability measure of M_i for each $i \in I$.*
3. *π is an invariant probability measure of M_σ for every permutation σ.*

Proof. This is the compact-space equivalent result to Proposition 5 of Samet (1998a), and the proof is nearly identical. Almost immediately from the definitions, 1) and 2) are equivalent. That 2) implies 3) is quite readily seen: if $\pi t_i = \pi$ for each player, then one can successively calculate $\pi(t_{\sigma(1)} \ldots t_{\sigma(n)}) = \pi(t_{\sigma(2)} \ldots t_{\sigma(n)}) = \cdots = \pi t_{\sigma(n)} = \pi$, for any permutation σ.

It remains to show that 3) implies 2). Suppose 3), and let π be the invariant probability measure. Thus

$$\pi(t_1 t_2 \ldots t_n) = \pi.$$

Multiplying from the right by t_1 gives

$$\pi(t_1 t_2 \ldots t_n t_1) = \pi t_1.$$

So πt_1 is an invariant probability measure of $t_2 \ldots t_n t_1$. But by 3), π is an invariant probability measure of $M_2 M_n \ldots M_1$, and by the previous proposition, $M_2 M_n \ldots M_1$ has a unique invariant probability measure. Thus, $\pi M_1 = \pi$, and by entirely similarly arguments $\pi M_i = \pi$ for all i. □

Corollary 1. *For every $Q \in \wedge \boldsymbol{\Pi}$, there exists at most one common prior on Q.*

5. Permutations, iterated expectations and common priors

5.1. Main results

Proposition 3. *Given a compact e.m.p. type space Ω, for each random variable f on Ω and permutation σ, $\lim_{n\to\infty} E_\sigma^{(n)} f$ exists, and on each element Q in $\wedge \boldsymbol{\Pi}$ it is constant and is equal to $\pi_\sigma^Q f$, π_σ^Q-almost everywhere.*

Proof. This follows from the previous propositions and Birkhoff's Ergodic Theorem, cited in Section 2.2. □

Proposition 4. *Given a compact e.m.p. type space Ω satisfying $\wedge \boldsymbol{\Pi} = \{\Omega\}$, a common prior π exists iff for each random variable f, the elements of*

$$\left\{ \lim_{n\to\infty} E_\sigma^{(n)} f \;\middle|\; \sigma \in Sym(I) \right\}$$

converge π_σ-almost everywhere to the same limit. Moreover, if π is the common prior, then this limit is πf, π-almost everywhere.

Proof. As above, $\lim_{n\to\infty} E_\sigma^{(n)} f$ is constantly $\pi_\sigma f$, π_σ-almost everywhere, where π_σ is the unique invariant probability measure of M_σ on Ω. Thus, for each f, the limits for all σ are respectively π_σ-a.e. equal to each other iff for each f, $\pi_\sigma f$ are π_σ-a.e. constantly equal to the same value for all σ.

Clearly, if there is a probability measure π such that $\pi_\sigma = \pi$ for all σ, then the $\pi_\sigma f$ are all equal to each other. In the other direction, if in particular for each $A \in \mathcal{F}$, the $\pi_\sigma \chi_A$ are all equal, then there is a probability measure π such that $\pi_\sigma = \pi$ for all σ. This amounts, given previous propositions, to saying that π is a common prior. □

We can summarize these results as follows:

Main Theorem. *Given a compact e.m.p. type space whose meet is a single element, for each random variable f and permutation σ of the players, the Cesàro iterated expectation of f with respect to σ converges, and the value of its limit is common knowledge. Moreover, there exists a common prior if and only if for each random variable it is common knowledge that all its Cesàro iterated expectations with respect to all permutations converge to the same value.*

5.2. On the use of Cesàro means

In Samet (1998a), in the finite state space case, results are stated in terms of iterated expectations, i.e. $\lim_{n\to\infty} E_\sigma^n f$, whereas the results here work with Cesàro limits, i.e. $\lim_{n\to\infty} E_\sigma^{(n)} f$. It may be natural to inquire what is gained and/or lost in this distinction.

At the intuitive level, returning to the story of Eve and Adam in Samet (1998a), we again have the iterated operations of Eve computing the expectation of Adam's expectation of Eve's expectation … and so on. But now the sequences we concentrate on, $E^n_\sigma f$, are the running averages of these iterated expectations, rather than the expectations themselves, and the question is whether or not these averages converge to the same value.

From one perspective, Proposition 4 can be regarded as pointing to a "test" for establishing whether a common prior exists — in words, check if $\lim_{n\to\infty} E^{(n)}_\sigma f$ converges a.e. to the same value for each $\sigma \in Sym(I)$ and each random variable f. But because the operation of taking Cesàro means preserves convergent sequences and their limits — i.e. if $\lim_{n\to\infty} E^n_\sigma f = a$ then certainly $\lim_{n\to\infty} E^{(n)}_\sigma f = a$ — in one direction it suffices to replace the Cesàro iterated expectation $\lim_{n\to\infty} E^{(n)}_\sigma f$ with the simple iterated expectation $\lim_{n\to\infty} E^n_\sigma f$, and from this point of view we have an "exact" extension of the finite Samet result. However, if one has identified a random variable f and $\sigma \in Sym(I)$ such that $\lim_{n\to\infty} E^n_\sigma f$ diverges, that is not sufficient to conclude, in the infinite state space case, that there is no common prior, because in that case one needs to check in addition whether $\lim_{n\to\infty} E^{(n)}_\sigma f$ also diverges.

Similarly, Proposition 4 asserts that if there is a common prior, then for each random variable f and each $\sigma \in Sym(I)$, $\lim_{n\to\infty} E^{(n)}_\sigma f$ converges a.e. to the same value — but from this it cannot be concluded that the same can be said of the iterated expectation $\lim_{n\to\infty} E^n_\sigma f$, because it is possible for the latter to diverge when the Cesàro sums converge.

Nevertheless, this result may still be useful for certain applications. To take one example, consider the model of utilitarian preference aggregation under incomplete information presented in Nehring (2004), in which social preferences amongst a set of agents I is calculated as $E_\mu(\sum_{i\in I} U^f_i)$, where the random variable U^f_i is agent i's utility derived from "social act" f (represented as a random variable over a finite set of states Ω) and E_μ denotes the expectation with respect to a common prior μ. Without going into details here, the key point of the model in that paper of interest here is a result that asserts that act f is "socially preferred" to g, written $f \succ_I g$, if and only if $E_\mu(\sum_{i\in I} U^f_i) > E_\mu(\sum_{i\in I} U^g_i)$, where the common prior μ is assumed to exist. The common prior therefore plays the role of a "group" valuation. Nehring (2004) seeks to characterize this group valuation in terms of the beliefs of the individual agents, and appeals to Samet's result to do so.

With the results of this paper, extending Nehring's model to an infinite compact e.m.p. type space Ω with a common prior μ, it can be shown that the group valuation may be related to the (potentially finite iterations of) beliefs of individual agents. In Nehring (2004), $f \succ_I g$ if and only if for some finite sequence $\{i_1, \ldots, i_k\}$, it is common knowledge that $E_{i_k\ldots i_1}(\sum_{i\in I} U^f_i) > E_{i_k\ldots i_1}(\sum_{i\in I} U^g_i)$. Using the Cesàro mean approach in the infinite compact case, we can recapitulate this result in the "if" direction. In place of Nehring's supposition of common knowledge that $E_{i_k\ldots i_1}(\sum_{i\in I} U^f_i) > E_{i_k\ldots i_1}(\sum_{i\in I} U^g_i)$ for some finite sequence, suppose (in the infinite compact e.m.p. case) that for some finite k and permutation $\sigma \in Sym(I)$, there is common knowledge amongst the agents in I that $E^k_\sigma(\sum_{i\in I} U^f_i) > E^k_\sigma(\sum_{i\in I} U^g_i) + \varepsilon$, where $\varepsilon > 0$ is arbitrary. This last inequality may be rephrased as

$$E^k_\sigma\left(\sum_{i\in I} U^f_i - \sum_{i\in I} U^g_i\right)(\omega) > \varepsilon.$$

It then follows from the definitions that there is common knowledge amongst the agents that $E^n_\sigma(\sum_{i\in I} U^f_i - \sum_{i\in I} U^g_i) > \varepsilon$ for all integers $n > k$. But from this it readily follows that

$$\lim_{n\to\infty} E^{(n)}_\sigma\left(\sum_{i\in I} U^f_i\right) > \lim_{n\to\infty} E^{(n)}_\sigma\left(\sum_{i\in I} U^g_i\right) + \varepsilon,$$

or

$$E_\mu\left(\sum_{i\in I} U^f_i\right) > E_\mu\left(\sum_{i\in I} U^g_i\right) + \varepsilon,$$

since by Proposition 4 and the assumption of the existence of a common prior μ, $E_\mu(\sum_{i\in I} U^f_i)$ is given by $\lim_{n\to\infty} E^{(n)}_\sigma(\sum_{i\in I} U^f_i)$ for each permutation σ. We can then conclude that $f \succ_I g$.

6. The necessity of compactness

In this section we demonstrate that the conclusion of Proposition 1 above, namely that each M_σ has an invariant probability measure, does not hold when the assumption of compactness is relaxed. As the proofs of the propositions subsequent to Proposition 1 are ultimately dependent on the conclusion of Proposition 1, they cannot be conducted without compactness.

The example we use here is a variant of the famous "electronic mail" games. Consider two individuals, Anna and Ben, and a denumerable state space $\Omega = \{1, 2, 3, \ldots\}$. Anna's partition is $\{\{1\}, \{2, 3\}, \{4, 5\}, \ldots\}$ and Ben's partition is $\{\{1, 2\}, \{3, 4\}, \{5, 6\}, \ldots\}$. The meet is $\{\Omega\}$.

Ben's beliefs are always equal probabilities to the two states in each of his partition members. Anna's beliefs are also equal probabilities to the two states in her partition members, save for the probability 1 which is necessary for the single partition containing one state.

We can depict the beliefs of each of the two players in the form of infinite matrices:

$$\text{Anna} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \dots \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & \dots \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \frac{1}{2} & \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \frac{1}{2} & \dots \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 & \dots \end{bmatrix}, \qquad \text{Ben} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & \dots \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \dots \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \dots \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & \dots \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 & \dots \\ \cdot & \cdot & \cdot & \cdot & 0 & 0 & \dots \end{bmatrix}.$$

Form the permutation matrix $M_\sigma := \text{Ben} \times \text{Anna}$

$$\text{Ben} \times \text{Anna} = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \dots \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \dots \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \dots \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \dots \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \dots \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & \dots \\ \cdot & \cdot & \cdot & 0 & 0 & \dots \\ \cdot & \cdot & \cdot & 0 & 0 & \dots \end{bmatrix}$$

and note that it forms the following pattern: letting O stand for the set of positive odd integers, and regarding M_σ as a mapping on the domain $\mathbb{N} \times \mathbb{N}$, we start with $M_\sigma(1,1) = \frac{1}{2}$, $M_\sigma(2,1) = \frac{1}{2}$, and for each $j \in O$, $\frac{1}{4} = M_\sigma(j, j+1) = M_\sigma(j+1, j+1) = M_\sigma(j+2, j+1) = M_\sigma(j+3, j+1) = M_\sigma(j, j+2) = M_\sigma(j+1, j+2) = M_\sigma(j+2, j+2) = M_\sigma(j+3, j+2)$. For all other values of k and l, $M_\sigma(k,l) = 0$.

Suppose now that there is an invariant probability measure π with respect to M_σ. Let $\pi(1) = \alpha$. Then by the definition of invariant probability, it must also be the case that $\pi(2) = \alpha$, because $\pi(1) = \frac{1}{2}(\pi(1) + \pi(2))$. Similar reasoning leads to the conclusion that $\pi(3) = \alpha$, $\pi(4) = \alpha$, $\dots, \pi(k) = \alpha, \dots$.

Now, $\alpha \in [0,1]$, so either $\sum_{k=1}^{\infty} \pi(k) = 0$, or $\sum_{k=1}^{\infty} \pi(k) = \infty$. In either case, π cannot be a normalized probability.

7. Conclusion

As stated in the introduction, in this paper we have extended most of the results of Samet (1998a) to compact e.m.p. type spaces and shown that compactness is necessary for the proofs of the results. As noted in Section 3, whether our results on compact e.m.p. type spaces also apply with respect to all I-sequences remains an open question.

References

Aumann, R.J., 1976. Agreeing to disagree. Ann. Stat. 4 (6), 1236–1239.
Bonanno, G., Nehring, K., 1999. How to make sense of the common prior assumption under incomplete information. Int. J. Game Theory 28 (3), 409–434.
Feinberg, Y., 2000. Characterizing common priors in the form of posteriors. J. Econ. Theory 91 (2), 127–179.
Harsányi, J.C., 1967–1968. Games with incomplete information played by Bayesian players. Manage. Sci. 14, 159–182, 320–334, 486–502.
Heifetz, A., 2006. The positive foundation of the common prior assumption. Games Econ. Behav. 56 (1), 105–120.
Hernández-Lerma, O., Lasserre, J.B., 2003. Markov Chains and Invariant Probabilities. Birkhäuser-Verlag, Basel.
Milgrom, P., Stokey, N., 1982. Information, trade and common knowledge. J. Econ. Theory 26 (1), 17–27.
Morris, S., 1994. Trade with heterogeneous prior beliefs and asymmetric information. Econometrica 62 (6), 1327–1347.
Nehring, K., 2004. The veil of public ignorance. J. Econ. Theory 119 (2), 247–270.
Samet, D., 1998a. Iterated expectations and common priors. Games Econ. Behav. 24 (1–2), 131–141.
Samet, D., 1998b. Common priors and the separation of convex sets. Games Econ. Behav. 24 (1–2), 173–175.
Samet, D., 2000. Quantified beliefs and believed quantities. J. Econ. Theory 95 (2), 169–185.
Sebenius, J., Geanakoplos, J., 1983. Don't bet on it: Contingent agreements with asymmetric information. J. Amer. Statistical Assoc. 78 (382), 424–426.

CHAPTER 4

Countable Spaces and Common Priors

Abstract

We show that the no betting characterisation of the existence of common priors over finite type spaces extends only partially to improper priors in the countably infinite state space context: the existence of a common prior implies the absence of a bounded agreeable bet, and the absence of a common improper prior implies the existence of a bounded agreeable bet. However, a type space that lacks a common prior but has a common improper prior may or may not have a bounded agreeable bet. The iterated expectations characterisation of the existence of common priors extends almost as is, as a sufficient and necessary condition, from finite spaces to countable spaces, but fails to serve as a characterisation of common improper priors. As a side-benefit of the proofs here, we also obtain a constructive proof of the no betting characterisation in finite spaces.

Countable Spaces and Common Priors

1. Introduction

The common prior assumption (as first introduced in Harsányi (1967-1968)) is taken as an integral assumption in the vast majority of models of incomplete information. It asserts that the beliefs of individuals in different states of the world are the posteriors that they form, after each is given private information, from a prior that is common to them all.

Despite its pervasiveness, the common prior assumption was, and still is, debated and challenged (see Gul (1998) and Aumann (1998)). It has been noted that, in many cases of interest, all that observers have are profiles of posteriors, not priors, and that there are examples of posteriors that could not possibly have been derived from common priors.

Given the importance of the common prior assumption, intense interest has been focussed on fully characterising the existence of a common prior in terms of the posterior profiles — since we are interested in the players at the present time, it is desirable to express the assumption of a common prior in present-time terms only. Aumann (1976) in his agreement theorem, gave a necessary condition for the existence of a common prior in terms of present beliefs: if there is a common prior, then it is impossible to agree to disagree, i.e., to have common knowledge of differences in the beliefs of any given event. By extending the notion of disagreement to differences in the expectation of a general random variable, several researchers (Morris (1995); Feinberg (2000); Samet (1998b)) were able to show that the impossibility of there being common knowledge of disagreement is not only a necessary, but also a sufficient condition for the existence of a common prior. Since this characterisation is based on the criterion of whether or not there exists a bet such that the players take opposite sides of the bet, yet each player ascribes positive expected value to the bet at every state of the world (we will henceforth term such a bet an *agreeable bet*), it is often termed the 'no betting' characterisation. It has also been proved that this characterisation obtains for type spaces over compact, continuous state spaces (see Feinberg (2000) and Heifetz (2006)).

That left open the question of characterising the existence of common priors in type spaces over countable state spaces, a major lacuna given the many models of incomplete information in the game theory and economics literature that involve countable state spaces. That the no betting characterisation cannot be extended 'as is' to countable spaces was shown in Feinberg (2000), which presents an example of a type space over a countable state space that has no common prior, yet also admits no bounded agreeable bet (in fact, even no agreeable bet bounded from only above or from below).

Several researchers, however, noted that the counter-example in Feinberg (2000), and several other counter-examples (see, for example, Heifetz (2006) and Lehrer and Samet (2011)) admit no common prior but satisfy the property of having a common *improper* prior. An improper prior for a player is a measure over the state space that may not be normalisable, i.e., the measure of the entire space may be infinite. On encountering this idea for the first time, it may seem strange to consider measures that are not probability measures in the context of then deriving posterior probabilities, but there is a sense in which an improper prior is an entirely natural construction. Consider, for example, the standard uniform probability distribution in finite spaces, which represent the intuitive idea that 'any state is equally likely'. Clearly, there is no equivalent probability distribution over a countable space. The closest thing would be a non-normalisable measure μ that assigns equal weight to each state ω, say $\mu(\omega) = 1$. If now E is a finite partition element and one derives a posterior probability for an element $\omega \in E$ by defining it to be $\mu(\omega)/\mu(E)$, the result is indistinguishable from the posterior probability that would result if E were a subset of a finite space with the uniform distribution. Extending this idea to arbitrary measures gives the intuition behind improper priors.

Thus, an improper prior does not enable one to consider the probabilities that a player assigns to events at the *ex ante* stage, but it still enables discussion of the relative likelihood that he ascribes to pairs of events, as well as the interim probability assessments of the player — i.e. his types — constitute a disintegration of the improper prior. If one is interested in interim probability assessments and how they may be derived from *ex ante* considerations, without necessarily demanding that a player have a full-fledged *ex ante* probability measure, an improper prior can serve much the same purpose as a proper prior. An improper prior common to all the players is then a common improper prior.

There has been an open conjecture for several years, due to Heifetz (2006), that the no betting characterisation or a close variant of it might obtain with respect to common improper priors over countable state spaces. In this paper, we directly address this conjecture, and prove that the absence of a common improper prior is a sufficient condition for the existence of a bounded agreeable bet among players. It is not, however, a necessary condition; we exhibit a simple example of a type spaces over a countable state spaces that has both a common improper prior and a bounded agreeable bet. We also show that the existence of a (proper) common prior is a sufficient condition for the absence of a bounded agreeable bet.

These results, along with the example in Feinberg (2000), indicate that the no betting criterion is rather weak in the countable state space case. In particular, the 'intermediate' case of a type space with no common prior but a common improper prior is consistent both with the existence of a bounded agreeable bet and the absence of a bounded agreeable bet. We can present this schematically as follows:

- Common prior $\Rightarrow$ No betting
  No betting $\not\Rightarrow$ Common prior
- No betting $\Rightarrow$ Common improper prior
  Common improper prior $\not\Rightarrow$ No betting

One reason that the study of no betting in countable state spaces remained an open problem for several years was due to the fact that the known proofs of the no betting characterisation, in both the finite case and the compact, continuous case, were all based on one or another variant of the Separation Theorem for convex sets. As Heifetz (2006) notes, this theorem is inapplicable in non-compact type spaces. This dictated seeking a different approach in this paper.

The proof of Theorem 1(b) presented here, which takes up most of Section 4, is entirely combinatorial and constructive. If the sole intention were to obtain the shortest possible proof of the claim of the theorem this paper could be shortened considerably, making use of the well known no betting result in finite spaces with common priors along with Lemma 3 and Proposition 7. Despite this, there is much to be gained from the constructive proof in Section 4. Aside from the intrinsic value of a constructive proof that does not rely on the convex separation theorem, the details of the proof reveal interesting aspects of structures of partition profiles in knowledge spaces, using chains and cycles. This is in line with a spate of recent papers, including Rodrigues-Neto (2009), Lehrer and Samet (2011) and Hellman and Samet (2012), that have used these concepts to study the fine-grained structure of partition profiles, yielding new insights into common knowledge and related subjects.

Since the same construction for countable spaces applies in the finite state space case, we have as a side-effect the first constructive proof of the no betting characterisation for finite spaces. In fact, putting together the elements of the proofs in Section 4 essentially yields an algorithm which, given a finite type space, determines whether or not it has a common prior; if it does have a common prior, the algorithm then constructs the common prior; if it does not have a common prior, the algorithm constructs an agreeable bet, thus indicating a random variables about whose expected values the players 'agree to disagree'.

In the last section, Section 5, we turn our attention to studying, in the countable state space context, a second characterisation of the existence of a common prior in finite state spaces, the iterated expectations characterisation introduced in Samet (1998a) (extended to the case of compact, continuous spaces in Hellman (2011); Morris (2002) also considers iterated expectations in infinite state spaces). The main result in Samet (1998a) is that each iterated expectation of a random variable converges, and the value of its limit is common knowledge. Moreover, there exists a common prior if and only if for each random variable it is common knowledge that all its iterated expectations converge to the same value. In that paper, it is pointed out that 'the stochastic analysis of type spaces is finer than the convex analysis used for the nonagreement condition'. It turns out that this statement holds true even more strongly in the countable space context. Whereas, as pointed out above, the no betting criterion fails to extend as a characterisation even of common improper priors in countable type spaces, it is proved in Section 5 that the iterated expectations criterion extends, almost 'as is', to a full characterisation of the existence of a common prior. But the iterated expectations criterion *does* fail as a characterisation of common improper priors, and cannot be used for the goal of identifying common improper priors.

The study of common priors in the context of countable spaces involves a much richer set of concepts than the finite space context. Unlike the finite case, in the infinite one the non existence of a common prior can have different characteristics that have bearing on the question of consistency. Mapping the relationships between these concepts is an on-going effort, as of this writing. Research studies in this field complementary to this paper include Lehrer and Samet (2012), who provide a sufficient condition for type spaces that have a common improper prior to admit an agreeable bet.

Acknowledgements. Special thanks are due to Ehud Lehrer and Dov Samet for many helpful conversations as well as crucial assistance in formulating and proving Lemma 3 and Proposition 7. Thanks are also due to Aviad Heifetz for helpful correspondence on the subject of his conjecture.

2. Preliminaries

2.1. Knowledge and Belief.

A *knowledge space* for a nonempty set of *players* I, is a pair $(\Omega, \mathbf{\Pi})$. In this context, Ω is a nonempty set called a *state space*, and $\mathbf{\Pi} = (\Pi_i)_{i \in I}$ is a *partition profile*, where for each $i \in I$, Π_i is a partition of Ω into measurable sets with positive measure. We will assume throughout this paper that every state space Ω is either finite, or countably infinite, and that $|I| = m$, where

$m \geq 2$ is a finite integer. Denote by $\Delta(\Omega)$ the set of probability distributions over Ω.

When working with a knowledge space $(\Omega, \mathbf{\Pi})$, an element $\omega \in \Omega$ is typically termed a *state*. For each $\omega \in \Omega$, we denote by $\Pi_i(\omega)$ the element of Π_i containing ω. Π_i is interpreted as the information available to player i; $\Pi_i(\omega)$ is the set of all states that are indistinguishable to i when ω occurs. Player i is said to *know* an event E at ω if $\Pi_i(\omega) \subseteq E$. We define for each i a *knowledge operator* $K_i \colon 2^\Omega \to 2^\Omega$, by $K_i(E) = \{\omega \mid \Pi_i(\omega) \subseteq E\}$. Thus, $K_i(E)$ is the event that i knows E.

A partition Π' is a *refinement* of Π if every element of Π' is a subset of an element of Π. Refinement intuitively describes an increase of knowledge. The *meet* of $\mathbf{\Pi}$, denoted $\wedge\mathbf{\Pi}$, is the partition that is the finest among the partitions that are simultaneously coarser than all the partitions Π_i. $\mathbf{\Pi}$ is called *connected* when $\wedge\mathbf{\Pi} = \{\Omega\}$. (By abuse of notation, when $\mathbf{\Pi}$ is clear from context, we will sometimes say that Ω is connected when the intention is to say that $\mathbf{\Pi}$ is connected).

A *type function* for Π_i is a function $t_i : \Omega \to \Delta(\Omega)$ that associates with each state ω a distribution in $\Delta(\Omega)$, in which case the latter is termed the *type* of i at ω. Each type function t_i further satisfies the following two conditions:

(a) $t_i(\omega)(\Pi_i(\omega)) = 1$, for each $\omega \in \Omega$;
(b) t_i is constant over each element of Π_i.

A *type profile* for $\mathbf{\Pi}$ is an n-tuple of type functions, $\tau = (t_i)_{i \in I}$, where for each i, t_i is a type function for Π_i, which intuitively represents the player's beliefs. A type profile τ is *positive* if $t_i(\omega)(\omega) > 0$ for each i, and each state ω.

By definition of a type function, abusing notation, we may write $t_i(\omega)$ as short-hand for $t_i(\omega)(\omega)$, with the distinction between the intended interpretation of $t_i(\omega)$ as an element of $\Delta(\Omega)$ or as an element of $\mathbb{R}$ clear from context.

A *random variable* f over Ω is any element of $\mathbb{R}^\Omega$. Given a probability measure $\mu \in \Delta(\Omega)$ and a random variable f, the *expected value of f* with respect to μ, is

$$E_\mu f := \sum_{\omega \in \Omega} f(\omega)\mu(\omega). \tag{1}$$

For a random variable f, denote by $E_i f$ the element of $\mathbb{R}^\Omega$ defined by

$$E_i f(\omega) := \sum_{\omega' \in \Pi_i(\omega)} t_i(\omega') f(\omega'). \tag{2}$$

We will alternatively also sometimes write $E_i(f|\omega)$ in place of $E_i f(\omega)$, and call this the *interim expected value* player i ascribes to f at ω.

2.2. Priors.

If Ω is a countable state spaces, an *improper prior* for a type function t_i is a non-negative and non-zero function $p\colon \Omega \to \mathbb{R}$ such that for each $\pi \in \Pi_i$, $p(\pi) < \infty$ and $p(\pi)t_i(\omega) = p(\omega)$ for all $\omega \in \pi$. Note that although for any $\pi \in \Pi_i$, $p(\pi) < \infty$, the possibility that $p(\Omega) = \infty$ is not ruled out, so that p may not be normalisable.

A *prior* for a type function t_i is a probability distribution $p \in \Delta(\Omega)$ such that for each $\pi \in \Pi_i$, and $\omega \in \pi$, the equation $p(\pi)t_i(\omega) = p(\omega)$ is satisfied. Obviously, a prior is in particular an improper prior (normalising if necessary in the finite case), so that when we use the term 'improper prior' we will mean both concepts, but the term 'prior' alone will mean a normalisable prior.

An improper prior does not allow us to talk about the 'probabilities' that player i assigns to events at the *ex ante* stage, but it still allows us to discuss the relative likelihood that he ascribes to pairs of events; and the interim probability assessments of the player i.e., his types constitute a disintegration of the improper prior.

A *common improper prior* for a type profile τ is a $p \in D(\Omega)$ that is an improper prior for each player i.[1] A type profile τ is called *consistent* (the term is due to Harsányi) when it has a common prior.

Note also that if p is a common improper prior, then for any constant $\gamma > 0$, γp is also a common improper prior. In particular, if p is a common improper prior and $p(\Omega) < \infty$ then $[p(\Omega)]^{-1}p$ is a common prior. Thus, for a finite space, a type profile has a common prior if and only if it has a common improper prior. In light of this, we extend the definition of consistency to countable spaces: a type profile τ is *consistent* when it has a common improper prior and *inconsistent* otherwise.

2.3. Type Matrices and Common Priors.

Samet (1998a) introduced a matrix-based approach to the analysis of common priors that is convenient for the results of this paper. Although that paper restricts attention to finite state spaces, the matrix definitions extend readily to countable state spaces.

[1] Contrasting a prior for t_i with the types $t_i(\omega, \cdot)$, the latter are referred to as the posterior probabilities of i.

We will denote by **I** the square identity matrix

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix},$$

trusting that the size of this matrix in each case will be clear by context (including cases in which the matrix is countably infinite).

For a type space τ, over either finite or countable Ω, define for each player i a *type matrix* M_i in $\mathbb{R}^{\Omega^2}$, by $M_i(\omega, \omega') = t_i(\omega)(\{\omega'\})$. M_i is a Markov matrix representing the transition function t_i. Each type matrix M_i satisfies the following properties:

- M_i is idempotent, i.e., $M_i M_i = M_i$. This follows from the fact that the multiplication of the j-th row of M_i with the k-th column of M_i is equal to the multiplication of the j-th row by the constant $t_i(\omega_j)(\{\omega_k\})$. But the j-th row sums to 1.
- A p such that for each $\pi \in \Pi_i$, $p(\pi) < \infty$, is an improper prior for i if and only if $pM_i = p$. This follows from the fact that the condition that $p(\omega_j)$ equals p times the j-th column of M_i is equivalent to the condition $p(\pi) t_i(\omega_j)(\pi) = p(\omega_j)$.
- For each random variable f, $M_i f = E_i f$. We will therefore use the notation $E_i f$ and $M_i f$ interchangeably.

Furthermore, for any $q \in \mathbb{R}^\Omega$, if $p = qM_i$ is a well-defined element $p \in \mathbb{R}^\Omega$, then p is an improper prior for i. To see this, using the idempotence of M_i, note that $pM_i = qM_iM_i = qM_i = p$.

Given a permutation σ of I, denote

$$M_\sigma := M_{\sigma_1} \cdots M_{\sigma_n} = t_{\sigma_1} \cdots t_{\sigma_n}$$

and term this a *permutation matrix*.

For every $P \in \wedge\mathbf{\Pi}$, and every i, denote by M_i^P the matrix M_i restricted to P, and similarly denote by M_σ^P the matrix M_σ restricted to P.

Given the identification of $E_i f$ with $M_i f$, we can write, given a permutation σ of I,

$$E_\sigma := E_{\sigma_1} \cdots E_{\sigma_n} = M_\sigma.$$

The *iterated expectation of f with respect to σ* is the sequence $\{E_\sigma^k f\}_{k=1}^\infty$. Samet (1998a) proves that in the finite state space case, p is a common prior if and only if p is an invariant probability measure of the Markov matrix M_σ for each permutation σ. Furthermore, if the knowledge space is connected, there exists a common prior if and only if for each random variable f, the iterated expectations of f, with respect to all permutations σ, converge to the same limit, and if p is the common prior, then this limit is pf. The analogues of these results are explored in this paper in Section 5.

2.4. Common Knowledge.

An event $E \subseteq \Omega$ is *self-evident* if for all $\omega \in E$ and each $i \in I$

$$\Pi_i(\omega) \subseteq E. \tag{3}$$

In particular, every element of the meet, $M \in \wedge\mathbf{\Pi}$, is self-evident.

An event E is *common knowledge at* $\omega \in \Omega$ if and only if there exists a self-evident event $F \ni \omega$ such that for all $i \in I$

$$F \subseteq K_i(E). \tag{4}$$

In fact, the element of the meet containing ω is also known as the *common knowledge component* of ω, because it is the smallest self-evident set containing ω.

Working with a connected space is thus particularly convenient for theorems involving common knowledge, because if $\mathbf{\Pi}$ is connected, then an event E is common knowledge at ω if and only if $E = \Omega$.

2.5. Characterisation of the Existence of Common Priors.

We adopt the notation that for n-tuples $x_1, x_2 \in \mathbb{R}^\Omega$, $x_1 > x_2$ means that $x_1(\omega) > x_2(\omega)$ for all $\omega \in \Omega$, and $x_1 > 0$ means that $x_1(\omega) > 0$ for all ω. $x_1 \geq x_2$ will be interpreted to mean that $x_1(\omega) \geq x_2(\omega)$ for all $\omega \in \Omega$, and there is at least one ω^* such that $x_1(\omega^*) = x_2(\omega^*)$.

DEFINITION 1. An m-tuple of random variables $\{f_1, \ldots, f_m\}$ is a *bet* if $\sum_{i=1}^m f_i = 0$. Given an m-player type space τ, a bet f is an *agreeable bet* if $E_i f_i > 0$ for all i. ♦

In the two-player case, the condition that $\sum_{i=1}^m f_i = 0$ is the same as $f_1 = -f_2$. Hence in this case we will sometimes say that a bet is agreeable if there is a random variable f such that $E_1 f > 0 > E_2 f$.

We can also slightly tweak the definition, replacing strict inequality with weak inequality, to obtain:

DEFINITION 2. Given an m-player type space τ, an m-tuple of random variables $\{f_1, \ldots, f_m\}$ is a *weakly agreeable bet* if $\sum_{i=1}^m f_i = 0$, and $E_i f_i \geq 0$ for all i, with $E_j f_j > 0$ for at least one $j \in I$. ♦

We will say that an agreeable bet $\{f_1, \ldots, f_m\}$ is bounded if $|f_i|$ is bounded for all $i \in I$. In addition, given two sequences of r.v. $f = \{f_1, \ldots, f_m\}$ and $g = \{g_1, \ldots, g_m\}$, we define $f + g := \{f_1 + g_1, \ldots, g_1 + g_m\}$

The characterisation of the existence of common priors in finite spaces is accomplished by:

> A finite type space τ has a common prior if and only if there is no agreeable bet.

The functions f_i, which sum to zero, can be interpreted as a bet between the players. The condition $E_i(f_i|\omega) > 0$, for each state ω, amounts to saying that the positivity of $E_i f_i$ is always common knowledge amongst the players. Thus, a necessary and sufficient condition for the existence of a common prior is that there is no bet for which it is always common knowledge that all players expect a positive gain. This establishes a fundamental, and remarkable, two-way connection between posteriors and priors.

The most accessible proof of this result is in Samet (1998b). It was proved by Morris (1995) for finite type spaces and independently by Feinberg (2000) for compact type spaces. Bonanno and Nehring (1999) proved it for finite type spaces with two agents.

3. Main Results

3.1. Agreeable Betting.

THEOREM 1. *Let τ be a type space over $\{\Omega, \mathbf{\Pi}\}$, where Ω is countable.*

(a) *If τ has a common prior, then there is no bounded weakly agreeable bet relative to τ.*
(b) *If τ has no common improper prior, then there exists a bounded agreeable bet relative to τ.*

PROPOSITION 1. *Let τ be a type space over $\{\Omega, \mathbf{\Pi}\}$, where Ω is countable. If the set of priors of at least one player is compact, then τ has a common prior if and only if there is no bounded agreeable bet relative to τ.*

COROLLARY 1. *Let τ be a type space over $\{\Omega, \mathbf{\Pi}\}$, where Ω is countable. If the partition Π_j of at least one player j is finite, then τ has a common prior if and only if there is no bounded agreeable bet relative to τ.*

3.2. Counterexamples.

We show here by counter-examples that the converses to the statements in Theorem 1 do not obtain.

The first example shows that the converse to Theorem 1(b) does not obtain, by exhibiting a type space with both a common improper prior and a bounded agreeable bet.

EXAMPLE 1. The state space is $\Omega = \{\omega_0, \omega_1, \ldots\}$. There are two players, Anne and Ben. Anne's knowledge partition, Π_A, is given by

$$\{\{0\}, \{1,2\}, \{3,4\}, \{5,6\}, \ldots\}.$$

Ben's partition, Π_B, is given by

$$\{\{0,1\}, \{2,3\}, \{4,5\}, \ldots\}.$$

Anne's type function, t_A, is given by $t_A(\omega_0, \omega_0) = 1$, and $t_A(\omega_n, \omega_n) = 0.5$ for all $n \geq 1$. Ben's type function, t_B, is given by $t_B(\omega_n, \omega_n) = 0.5$ for all $n \geq 0$. See Figure 1.

Anne	1 |	1/2	1/2 |	1/2	1/2 |	1/2	1/2 |	...
Ben	1/2	1/2 |	1/2	1/2 |	1/2	1/2 |	1/2	...

FIGURE 1. The partition profile of Example 1.

Let $p(\omega) = 1$ for all $\omega \in \Omega$. Then p is a common improper prior.

Define a bounded random variable f as follows:

$$f(\omega_n) = \begin{cases} 1 & \text{if } n = 0 \\ 1 + \sum_{i=1}^{n} \frac{1}{2^i} & \text{if } n \text{ is even} \\ -(1 + \sum_{i=1}^{n} \frac{1}{2^i}) & \text{if } n \text{ is odd} \end{cases}$$

Then $E_A(f|\omega) > 0 > E_B(f|\omega)$ for all $\omega \in \Omega$, hence $\{f, -f\}$ is a bounded agreeable bet. ♦

Example 1 exhibits a type space with a common improper prior and no common prior, but it cannot serve as a counterexample to Theorem 1(a) because of the existence of a bounded agreeable bet. The next example however, exhibits a type space with neither a common prior nor a bounded agreeable bet. The example is taken from Feinberg (2000), and is included here for completeness.

EXAMPLE 2. The state space is $\Omega = \{\omega_1, \omega_2, \ldots\}$. There are two players, Anne and Ben. Anne's knowledge partition, Π_A, is given by

$$\{\{1\}, \{2,3\}, \{4,5\}, \{6,7\}, \ldots\}.$$

Ben's partition, Π_B, is given by

$$\{\{1,2\}, \{3,4\}, \{5,6\}, \ldots\}.$$

Ben's type function, t_B, is given by $t_B(\omega_n, \omega_n) = 0.5$ for all $n \geq 1$.

Anne's type function, t_A, is given by

$$t_A(\omega_n, \omega_n) = \begin{cases} 1 & \text{if } n = 1 \\ 2/3 & \text{if } n = 2^k - 2 \text{ for } k \geq 1 \\ 1/3 & \text{if } n = 2^k - 1 \text{ for } k \geq 1 \\ 1/2 & \text{otherwise} \end{cases}$$

See Figure 2.

Anne	1	2/3	1/3	1/2	1/2	2/3	1/3	1/2	...
Ben	1/2	1/2	1/2	1/2	1/2	1/2	1/2	1/2	...

FIGURE 2. The partition profile of Example 2.

Let p be a candidate for being a common improper prior; by the structure of the partition profile, the value of $p(\omega_1)$ determines the value of $p(\omega_n)$ for all n. It cannot be the case that $p(\omega_1) = 0$, because then $p(\omega_n) = 0$ for all n. There must therefore be a real number $\alpha > 0$ such that $p(\omega_1) = \alpha$. This then implies that $p(\omega_2) = \alpha$, $p(\omega_3) = p(\omega_4) = p(\omega_5) = p(\omega_6) = \alpha/2$, $p(\omega_7) = p(\omega_8) = \ldots = p(14) = \alpha/4$, and so on. It follows that for any positive α, setting $p(\omega_1) := \alpha$ determines a common improper prior p; it cannot, however, be a common prior, because $\sum_{n=1}^{\infty} p(\omega_n) = \infty$.

There is no agreeable bet in this example. Let f be a candidate for a random variable satisfying $E_A(f|\omega) > 0 > E_B(f|\omega)$ for all $\omega \in \Omega$. Then $f(\omega_1) = E_A(f|\omega_1) > 0$, $f(\omega_2) = 2E_B(f|\omega_1) - f(\omega_1) < -f(\omega_1)$, $f(\omega_3) = 3E_A(f|\omega_2) - 2f(\omega_2) > -2f(2)$. It emerges that $f(n)$ is an alternating sequence whose absolute value tends to infinity. ⧫

Examples 1 and 2 show that the no betting criterion is insufficiently subtle to be used as a tool for determining when a type space has a common improper prior but no common prior, as both examples satisfy that property, but one has a bounded agreeable bet and the other does not.

3.3. Unbounded Bets. The restriction to bounded bets in the statement of Theorem 1(a) is necessary. Example 3 shows that the statement in Theorem 1(a) does not hold for unbounded bets.

EXAMPLE 3. The state space is $\Omega = \{1, 2, \ldots\}$. There are two players, Anne and Ben. Anne's knowledge partition, Π_A, is given by

$$\{\{1\}, \{2, 3\}, \{4, 5\}, \{6, 7\}, \ldots\}.$$

Ben's partition, Π_B, is given by

$$\{\{1, 2\}, \{3, 4\}, \{5, 6\}, \ldots\}.$$

Anne's type function, t_A, is given by

$$t_A(n,n) = \begin{cases} 1 & \text{if } n = 1 \\ \frac{2}{3} & \text{if } n \text{ is even} \\ \frac{1}{3} & \text{if } n \text{ is odd, } n > 1 \end{cases}$$

Ben's type function, t_B, is given by

$$t_B(n,n) = \begin{cases} \frac{2}{3} & \text{if } n \text{ is odd} \\ \frac{1}{3} & \text{if } n \text{ is even} \end{cases}$$

See Figure 3.

Anne	1	2/3	1/3	2/3	1/3	2/3	1/3 ...
Ben	2/3	1/3	2/3	1/3	2/3	1/3	2/3 ...

FIGURE 3. The partition profile of Example 3.

Let $p(n) = 2^{-n}$ for all n. Then p is a common prior.

Fix $\varepsilon > 0$. Define an unbounded random variable f as follows:

$$f(n) = \begin{cases} 1 & \text{if } n = 1 \\ -(2f(n-1) + \varepsilon) & \text{if } n \text{ is even} \\ 2f(n-1) + \varepsilon & \text{if } n \text{ is odd, } n > 1 \end{cases}$$

Then $E_A f > 0 > E_B f$, hence $\{f, -f\}$ is an unbounded agreeable bet.

♦

We also have a sufficient condition for ensuring that a knowledge space admits an unbounded bet[2] for *any* positive type space defined over it.

THEOREM 2. *If a knowledge space $\{\Omega, \mathbf{\Pi}\}$ over a countably infinite space of states Ω satisfies the property that the number of partition elements of cardinality greater than* 1 *is infinite, then any positive type space $\{\Omega, \mathbf{\Pi}, \tau\}$ over $\{\Omega, \mathbf{\Pi}\}$ has an agreeable bet (bounded or unbounded).*

3.4. The Iterated Expectations Criterion.

In contrast to the no betting criterion, which fails to characterise either common priors or common improper priors in countable spaces, the iterated expectations criterion of Samet (1998a) does extend to a full characterisation of the existence of common priors.

[2] Lehrer and Samet (2012) provide another sufficient condition for the existence of unbounded bets: a partition is *locally finite* if each partition element is finite. Then Lehrer and Samet (2012) prove that every locally finite knowledge space admits an unbounded bet for any positive type space defined over it.

THEOREM 3. *Let τ be a positive type space, and suppose that $\wedge\mathbf{\Pi} = \{\Omega\}$. Then τ has a common prior if and only if for each bounded non-zero random variable $f \geq 0$, the iterated expectations of f, with respect to all permutations σ, converge to the same non-zero limit. Moreover, if p is the common prior, then this limit is pf.*

4. Acceptable Bets and Countable Spaces

First note that for the proof of Theorem 1, it suffices to prove the two statements of when the knowledge space $\mathbf{\Pi}$ is connected. For part (a), if there exists a common prior p over a non-connected $\mathbf{\Pi}$, we can decompose Ω into disjoint connected components, $\Omega = \sum_j S_j$, and then p restricted to each S_j is a common prior; if there is no acceptable bet over S_j, then there can be no acceptable bet over Ω. For part (b), again decompose Ω disjointly as $\Omega = \sum_j S_j$. If there is no common improper prior over Ω, then there can be no *cip* over each S_j (otherwise that *cip* could serve as a *cip* for all of Ω), and if we show the existence of a bounded agreeable bet f_{S_j} over each S_j, then the function $f(\omega) := f_{S_j}(\omega)$ for $\omega \in S_j$ is a bounded agreeable bet over Ω.

We will therefore assume that for all type spaces τ over $\{\Omega, \mathbf{\Pi}\}$ in this section, $\mathbf{\Pi}$ is connected.

The proof of Theorem 1(a) is a straightforward extension of the proof of the same statement in the finite state space case.

Proof of Theorem 1(a). Suppose by contradiction that there exists a bounded weakly agreeable bet f. First note that for any $j \in I$,

$$E_p(f_j) = \sum_{\omega' \in \Omega} f_j(\omega')p(\omega')$$

This quantity is well-defined given the assumptions that f is bounded, and that p is a (proper) common prior.

Next, as p is a prior for each $j \in I$, it follows that $E_p(f_j) = E_p(E_j f_j)$, where $E_j f_j$ is regarded as a function from Ω to $\mathbb{R}$.

As f is a bet, $\sum_{i \in I} f_i = 0$. Hence

$$0 = E_p\left(\sum_i f_i\right) = \sum_i E_p f_i = \sum_i E_p(E_i f_i). \tag{5}$$

But by the assumption that f is a weakly agreeable bet, there is at least one player i such that $E_i f_i > 0$, in which case $E_p E_i f_i > 0$, hence $\sum_i E_p(E_i f_i) > 0$, contradicting Equation 5. ∎

Proof of Proposition 1. For each $i \in I$, denote the set of all priors of player i by P_i. Further denote $P := \times_{i \in I} P_i$, and let D denote the diagonal, i.e., the set of all m-tuples $(p, \ldots, p)$ such that $p \in \Delta(\Omega)$. Clearly, τ has a common prior if and only if $P \cap D \neq \emptyset$. Suppose that P_i is compact, for some $i \in I$.

Suppose next that $P \cap D = \emptyset$, but that the distance between P and D (in any appropriate metric) is 0, i.e., P and D are not strictly separated. Then there exists a sequence $\overline{p}_1, \overline{p}_2, \ldots$, where for each t, $\overline{p}_t = (p_t^1, \ldots, p_t^m) \in P$, and there exists a sequence of probabilities $q_1, q_2, \ldots$ such that $\|\overline{p}_n - d_n\| \to 0$, where $d_n = (q_n, \ldots q_n) \in D$.

Then w.l.o.g. there exists a sequence $q_1^i, q_2^i \ldots$ of elements of P_i such that $\|q_n^i - q_n\| \to 0$. By the assumption that P_i is compact, there is a point $q \in P_i$ such that both q_n^i and q_n converge to q. We can also show that $q \in P_j$ for all j: for each player j, there is a sequence $q_1^j, q_2^j \ldots$ of elements of P_j such that $\|q_n^j - q_n\| \to 0$. Since $q_n \to q$, it follows that $q_n^j \to q$. As P_j is closed, $q \in P_j$.

We therefore deduce that there does not exist common prior if and only if P and D can be strictly separated. The rest of the proof then proceeds as in the analogous case in the finite state space setting, as presented in Samet (1998b). ∎

4.1. Chains.

Following a concept introduced in Hellman and Samet (2012), we have the following definition:

DEFINITION 3. A *chain* of length $n \geq 0$, for a partition profile $\mathbf{\Pi}$, from one state to another, is defined by induction on n. A state ω_0 is a chain of length 0 from ω_0 to ω_0. A chain of length $n+1$, from ω_0 to ω, is a sequence $c \xrightarrow{i} \omega$, where c is a chain of length n from ω_0 to ω', and $\omega \in \Pi_i(\omega')$. Thus, a chain of positive length n is a sequence $c = \omega_0 \xrightarrow{i_0} \omega_1 \xrightarrow{i_1} \cdots \xrightarrow{i_{n-1}} \omega_n$, such that $\omega_{s+1} \in \Pi_{i_s}(\omega_s)$ for $s = 0, \ldots, n-1$. ♦

We write $\omega \to \omega'$ when there is a chain from ω to ω', in which case we say that ω and ω' are connected by a chain. The binary relation $\to$ is the transitive closure of the union of the relations $\xrightarrow{i}$, and it is an equivalence relation.

DEFINITION 4. Given a chain $c = \omega_0 \xrightarrow{i_0} \omega_1 \xrightarrow{i_1} \cdots \xrightarrow{i_{n-1}} \omega_n$, its *reverse chain* c^{-1} is defined as

$$c^{-1} := \omega_n \xrightarrow{i_{n-1}} \omega_{n-1} \xrightarrow{i_{n-2}} \cdots \xrightarrow{i_0} \omega_0.$$

A chain c is *alternating* if no two consecutive states ω_s and ω_{s+1} in c are the same, and no two consecutive agents i_s and i_{s+1} in c are the same. ⧫

Hellman and Samet (2012) prove that a partition profile $\mathbf{\Pi}$ is connected if and only if every two states are connected by at least one chain.

4.2. Positive, Zero, and Singular States.

Towards the aim of proving Theorem 1(b), we introduce here a categorisation of states in Ω, relative to a type profile τ, which will be needed for the proofs.

DEFINITION 5. Given a type profile τ, a state $\omega \in \Omega$ is:

- *positive* if $t_i(\omega) > 0$ for all $i \in I$;
- *zero* if $t_i(\omega) = 0$ for all $i \in I$;
- *singular* if it is neither positive nor zero. ⧫

Based on the categorisation of states in Definition 5, define the following:

DEFINITION 6. Given a type profile τ,

- A subset $S \subseteq \Omega$ is *i-positive* if $t_i(\omega) > 0$ for all $\omega \in S$.
- A subset $S \subseteq \Omega$ is *positive* if it is i-positive for all i (equivalently, if every $\omega \in S$ is a positive state). A chain c satisfying the condition that every element $\omega \in c$ is a positive state is a *positive chain*.
- A subset $S \subseteq \Omega$ *i-non-singular* if $t_i(\omega) = 0$ for every singular $\omega \in S$.
- A subset $S \subseteq \Omega$ is *non-singularly positive* if it is positive, and every maximal chain c entirely contained in S satisfies the property that for every $\omega \in c$ and every $i \in I$, $\pi_i(\omega)$ is i-non-singular. ⧫

A subset S of Ω is thus non-singular positive if it is positive, and for every $\omega \in S$, and every $i \in I$, every $\omega' \in \Pi_i(\omega)$ satisfies the condition that either

- $\omega' \in S$, or
- ω' is a zero state, or
- ω' is a singular state such that $t_i(\omega') = 0$.

Note that it is immediate by definition that if Ω is a positive state space, then trivially the entire space Ω is non-singularly positive.

The distinctions made in Definition 6 are illustrated in the following examples.

EXAMPLE 4. Let $\Omega = \{\omega_1, \omega_2\}$, $\Pi_1 = \Pi_2 = \Omega$, and let τ_1 be $\{\{1, 0\}\}$ and τ_2 be $\{\{0, 1\}\}$.

Ω has no positive subset, because all the states in Ω are singular. Clearly, this partition profile also can have no common prior. ♦

EXAMPLE 5. $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, and let τ_1 be given by

$$\tau_1 = \{\{1\}, \{1/2, 1/2, \}, \{1/2, 1/2\}, \{1\}\},$$

and τ_2 be given by

$$\tau_2 = \{\{1/2, 1/2\}, \{1/2, 1/2\}, \{0, 1\}\}$$

(with the structures of Π_1 and Π_2 clear from the structures of τ_1 and τ_2).

τ_1	1	1/2	1/2	1/2	1/2	1
τ_2	1/2	1/2	1/2	1/2	0	1

FIGURE 4. The partition profile of Example 5.

The subset $\{\omega_1, \omega_2, \omega_3, \omega_4\}$ is positive, but it is not non-singularly positive, because $\omega_5 \in \Pi_1(\omega_4)$, but ω_5 is a singular state, and $\tau_1(\omega_5) = \frac{1}{2} > 0$. Ω, however, does have a non-singularly positive subset: the subset $\{\omega_6\}$ meets the conditions listed in Definition 6 for a positive subset. This partition profile has a common prior: $\{0, 0, 0, 0, 0, 1\}$ ♦

EXAMPLE 6. $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7\}$, and let τ_1 be given as

$$\tau_1 = \{\{0, 1/2, 1/2\}, \{1/2, 1/2\}, \{0, 1\}\},$$

and τ_2 be given as

$$\tau_2 = \{\{1, 0\}, \{1/2, 1/2\}, \{1/2, 1/2, 0\}\}$$

(with the structures of Π_1 and Π_2 clear from the structures of τ_1 and τ_2).

τ_1	0	1/2	1/2	1/2	1/2	0	1
τ_2	1	0	1/2	1/2	1/2	1/2	0

FIGURE 5. The partition profile of Example 6.

The subset $S = \{\omega_3, \omega_4, \omega_5\}$ is positive (as is any subset of it), but Ω has no non-singularly positive subset. For example, S is not non-singularly positive, because $\omega_2 \in \Pi_1(\omega_3)$, but ω_2 is a singular state, and $t_1(\omega_2) = \frac{1}{2} > 0$. A similar analysis can be conducted on each subset of Ω to show that it is not non-singularly positive. This partition profile has no common prior. ♦

LEMMA 1. *If S is a non-singularly positive subset of Ω, then for any $\omega \in S$, every $\omega' \in \Omega$ that is connected to ω via a positive chain is also an element of S. It follows that every non-singularly positive subset S can be decomposed as $S = \cup T_j$, where each T_j is a non-singularly positive subset such that all of the members of T_j are connected to each other by positive chains.*

The proof is in the appendix.

4.3. Type Ratios.

The proofs of the propositions in this subsection, which are mainly technical, are located in the appendix.

DEFINITION 7. Let τ be a type profile and (ω_1, ω_2) an ordered pair of positive states in $\pi \in \Pi_i$. The *type ratio* of (ω_1, ω_2) relative to i is[3] $\mathrm{tr}^i_\tau(\omega_1, \omega_2) = t_i(\pi, \omega_1)/t_i(\pi, \omega_2)$. If a chain $c = \omega_0 \overset{i_0}{\to} \omega_1 \overset{i_1}{\to} \cdots \overset{i_{n-1}}{\to} \omega_n$ of length $n > 0$ is a positive chain, the *type ratio* of c is $\mathrm{tr}_\tau(c) = \times_{k=0}^{n-1} \mathrm{tr}^{i_k}_\tau(\omega_k, \omega_{k+1})$. For a positive chain c of length 0, $\mathrm{tr}_\tau(c) = 1$. Thus, if $c = c' \overset{i}{\to} \omega$ where c' is a positive chain from ω_0 to ω' and ω' is a positive state, $\mathrm{tr}_\tau(c) = \mathrm{tr}_\tau(c')\mathrm{tr}^i_\tau(\omega', \omega)$.[4] ♦

We note here for later use two equalities involving type ratios that follow immediately from the definitions:

- For any chain c,

$$\mathrm{tr}(c^{-1}) = [\mathrm{tr}(c)]^{-1}, \tag{6}$$

- If $c = \omega_1 \overset{i}{\to} \omega_2 \overset{i}{\to} \omega_3$ (i.e., $\omega_2, \omega_3 \in \Pi_i(\omega_1)$), then

$$\mathrm{tr}(c) = \mathrm{tr}^i(\omega_1, \omega_3). \tag{7}$$

The following proposition extends the results of a proposition appearing in Hellman and Samet (2012), from positive type spaces to general type spaces.

PROPOSITION 2. *Let τ be a type profile over a connected knowledge space. Then there exists a common improper prior for τ if and only if Ω has a non-singularly positive subspace S with respect to τ, and for each ω_0 and ω in S, and chains c and c' entirely contained in S from ω_0 to ω, $\mathrm{tr}_\tau(c) = \mathrm{tr}_\tau(c')$.*

[3] The type ratio defined in Hellman and Samet (2012) is the inverse of the one defined here, i.e., there $\mathrm{tr}^i_t(\omega_1, \omega_2) = t_i(\pi, \omega_2)/t_i(\pi, \omega_1)$. Which definition is used is immaterial, as long as one keeps to it consistently in an exposition. The definition chosen here is more convenient for the equations developed in this paper.

[4] When we discuss only one type profile we omit the subscript τ in tr_τ.

DEFINITION 8. A chain $c = \omega_0 \xrightarrow{i_0} \omega_1 \xrightarrow{i_1} \cdots \xrightarrow{i_{n-1}} \omega_n$, where $\omega_{s+1} \in \Pi_{i_s}(\omega_s)$ for $s = 0, \ldots, n-1$, is a *cycle*[5] if $\omega_n = \omega_0$. If with respect to a cycle c of length n there is a pair $s, s' \in \{0, \ldots, n-1\}$ such that $s' > s+1$ and $\omega_{s'} \in \Pi_{i_s}(\omega_s)$, then we say that c has a *self-crossing point* at $\omega_{s'}$. A cycle c is a *non-crossing* cycle if it is alternating, and has no self-crossing points, i.e., for every pair $s, s' \in \{0, \ldots, n-1\}$ such that $s' > s+1$, $\omega_{s'} \notin \Pi_{i_s}(\omega_s)$. ♦

Definition 8 leads to an immediate corollary of Proposition 2:

COROLLARY 2. *Let τ be a type profile over a connected knowledge space. Then there exists a common improper prior for τ if and only if Ω has a non-singularly positive subspace S with respect to τ, and for each ω in S, every cycle $\overline{c} = \omega \to \omega$ that is entirely contained in S satisfies* $\mathrm{tr}(\overline{c}) = 1$.

In fact, we can do even better, and show that it suffices to check the type ratios only of non-crossing cycles, instead of all cycles:

PROPOSITION 3. *Let τ be a type profile over a connected knowledge space. Then there exists a common improper prior for τ if and only if Ω has a non-singularly positive subspace S with respect to τ, and every non-crossing cycle $\overline{c}$ that is entirely contained in S satisfies* $\mathrm{tr}(\overline{c}) = 1$.

4.4. No Common Improper Prior Implies Existence of a Bounded Agreeable Bet.

DEFINITION 9. Let τ be a type space over $(\Omega, \mathbf{\Pi})$, and let $X \subseteq \Omega$ be a subset of Ω. Define $\mathbf{\Pi}$ *restricted to* X, denoted $\mathbf{\Pi}^X$, to be the partition profile over X given by $\Pi_i^X(\omega) := \Pi_i(\omega) \cap X$ for any state ω. Furthermore, let τ^X, a type function τ *restricted to* X, to be any type function over $(X, \mathbf{\Pi}^X)$ that satisfies the property that for any $\omega \in \Omega$, $t_i(\omega)(\Pi_i^X)t_i^X(\omega) = t_i(\omega)$.

In the special case in which X is a positive subset of Ω, τ^X is explicitly given by:

$$t_i^X(\omega) := \frac{t_i(\omega)}{t_i(\Pi_i^X(\omega))}$$

for any $\omega \in X$ and any $i \in I$. ♦

Intuitively, Π_i^X is the partition of X derived from the partition Π_i of Ω by 'ignoring all states outside of X'. It then follows intuitively that $t_i^X(\omega)$, for each state $\omega \in X$, is $t_i(\omega)$ scaled relative to the other states in $\Pi_i^X(\omega)$ in such a way that $\sum_{\omega \in X} t_i^X(\omega) = 1$.

[5] Cf. a similar definition in Rodrigues-Neto (2009).

For a random variable f, denote

$$E_i^X(f \mid \omega) := \sum_{\omega' \in \Pi_i^X(\omega)} t_i^X(\omega')f(\omega').$$

A set of random variables $f = \{f_1, \ldots, f_m\}$ is an agreeable bet relative to τ^X if for all $\omega \in X$, $\sum_i f_i(\omega) = 0$, and $E_i^X(f|\omega) > 0$ for all $i \in I$.

Note that it follows from the definitions that if (ω_1, ω_2) is an ordered pair of positive states in $\pi \in \Pi_i^X$ that

$$\text{(8)} \qquad \mathrm{tr}^i_{\tau^X}(\omega_1, \omega_2) = \frac{t_i^X(\omega_1)}{t_i^X(\omega_2)} = \frac{t_i(\omega_1)}{\tau_i(\pi)} \frac{\tau_i(\pi)}{t_i(\omega_2)} = \frac{t_i(\omega_1)}{t_i(\omega_2)} = \mathrm{tr}^i_{\tau}(\omega_1, \omega_2),$$

from which it further immediately follows that for any chain c of τ whose elements are entirely contained in X,

$$\text{(9)} \qquad \mathrm{tr}_{\tau^X}(c) = \mathrm{tr}_{\tau}(c).$$

We need one more definition.

DEFINITION 10. Let τ be a type space over $(\Omega, \mathbf{\Pi})$, let $X \subseteq \Omega$ be a positive subset of Ω, and let f be a random variable. A state $\omega \in X$ is a *surplus state for player* i relative to f and τ^X if $E_i^X(f|\omega) > 0$. In the context of a sequence $f = \{f_1, \ldots, f_m\}$ of r.v., we will say that ω is an i*-surplus state* if f_i is a surplus state for player i.

PROPOSITION 4. *Let τ be a type space over $(\Omega, \mathbf{\Pi})$, let S be a finite connected subset of positive states in Ω, and let $X \subseteq S$. Suppose that there exists an agreeable bet relative to τ^X. Then there exists an agreeable bet relative to τ^S.*

Proof. Let f be an agreeable bet relative to τ^X. If $X = S$, there is nothing to prove. If $X \subset S$, then by the assumption of the connectedness of S, we can find at least one player i and a point $\omega' \notin X$ such that $\Pi_i(\omega') \cap X \neq \emptyset$. By the assumption of positivity, $t_i(\omega') > 0$, and by the assumption that f is an agreeable bet, every state $\omega \in \Pi_i(\omega')$ is an i-surplus state relative to f.

Denote $Y := X \cup \omega'$, and let ε be the (by the i-surplus state assumption) positive value

$$\text{(10)} \qquad \varepsilon := \sum_{\omega'' \in \Pi_i(\omega') \cap X} f_i(\omega'') t_i^X(\omega'').$$

Next, let $\overline{f}_i(\omega')$ be a negative real number satisfying

$$\text{(11)} \qquad 0 > \overline{f}_i(\omega') > \frac{-(1 - t_i^Y(\omega'))}{t_i^Y(\omega')} \varepsilon \,,$$

and for $j \neq i$, set $\overline{f}_j(\omega') := -\overline{f}_i(\omega')/(m-1) > 0$, where $m = |I|$. Clearly, by construction, $\sum_{j \in I} \overline{f}_j(\omega') = 0$. Complete the definition of $\overline{f}$ by letting $\overline{f}(\omega'') := f(\omega'')$ for all $\omega'' \in X$.

It is straightforward to check that $\overline{f}$ is an agreeable bet relative to τ^Y. Now simply repeat this procedure as often as necessary to extend the agreeable bet to every state in the finite set S. ∎

We already have enough to conclude that for countable positive type spaces if there is no common prior then there is an agreeable bet.

PROPOSITION 5. *Let τ be a type space over $(\Omega, \mathbf{\Pi})$, where Ω is countable and every state in Ω is positive with respect to τ. If τ has no common prior then there exists an agreeable bet relative to τ.*

Proof. It suffices to suppose that τ is connected (if not, we can focus on each connected component separately).

Since τ has no common prior, by results in Hellman and Samet (2012) there is a finite subset $S_0 \subset \Omega$ such that there is a cycle c entirely contained in S_0 satisfying $\mathrm{tr}(c) \neq 1$. It follows that τ^{S_0} has no common prior. Therefore there is an agreeable bet f_0 (with $|f_0| \leq 1$) relative to τ^{S_0}.

Next, let $S_0 \subset S_1 \subset S_2 \ldots$ be a sequence of finite subsets, $S_i \subset \Omega$ for all i, such that $\bigcup_{i=0}^{\infty} S_i = \Omega$. By Proposition 4, for each $i \geq 0$ there exists an agreeable bet f_i relative to τ^{S_i} (with $|f_i| \leq 1$). For each i, define $\overline{f}_i : \Omega \to \mathbb{R}$ by

$$\overline{f}_i(\omega) = \begin{cases} f_i(\omega) & \text{if } \omega \in S_i \\ 0 & \text{otherwise} \end{cases}$$

Defining $f : \Omega \to \mathbb{R}$ by $f := \sum_{i=0}^{\infty} f_i/2^i$ yields an agreeable bet relative to τ. ∎

LEMMA 2. *Let τ be a type space over $(\Omega, \mathbf{\Pi})$, and let X be a non-crossing cycle such that $\mathrm{tr}(X) \neq 1$. Then there exists a random variable f that is an agreeable bet relative to τ^X.*

Proof. Write the non-crossing cycle as $X = \omega_1 \overset{i_1}{\to} \omega_2 \overset{i_2}{\to} \ldots \omega_n \overset{i_n}{\to} \omega_{n+1} = \omega_0$, where $\omega_{s+1} \in \Pi_{i_s}(\omega_s)$ for $s = 1, \ldots, n$. Assume without loss of generality that $\mathrm{tr}(X) < 1$ (otherwise simply reverse the ordering of states in the cycle). To cut down on notational clutter, write $r_s := \mathrm{tr}^{i_s}(\omega_s, \omega_{s+1})$, hence $\mathrm{tr}(X) = r_1 r_2 \cdots r_n$. Furthermore, denote by P the set such that $i \in P$ if and only if i is one of $i_1, i_2, \ldots i_n$ used in the presentation above of the cycle X.

We will several times make use of the following simple technical observation: suppose that i is a player, π is a partition element of Π_1, and

$\omega', \omega'' \in \pi$. Furthermore, suppose that g is a random variable satisfying the property that $g(\omega) = 0$ for all $\omega \in \pi$ such that $\omega \neq \omega', \omega''$. Then:

$$(12) \quad \left\{ \begin{array}{ll} E_i(g|\omega') = 0 & \text{if } g(\omega'') = -\mathrm{tr}^i(\omega', \omega'')g(\omega'), \\ E_i(g|\omega') > 0 & \text{if } g(\omega'') > -\mathrm{tr}^i(\omega', \omega'')g(\omega'), \\ E_i(g|\omega') < 0 & \text{if } g(\omega'') < -\mathrm{tr}^i(\omega', \omega'')g(\omega'). \end{array} \right\}$$

We now proceed to the construction of an agreeable bet, in stages.

First note that by the assumption that $r_1 r_2 \cdots r_n < 1$, we may choose a $\delta_n > 1$ such that $r_1 r_2 \cdots r_n \delta_n < 1$. We may then further define δ_{n-1} such that $\delta_n > \delta_{n-1} > 1$ and so on, to yield a sequence $\delta_n > \delta_{n-1} > \ldots > \delta_2 > 1$. Use this to define $\overline{f}$ by:

$$\begin{array}{ll} \overline{f}_{i_1}(\omega_1) = -1 & \\ \overline{f}_{i_n}(\omega_1 = \omega_{n+1}) = 1 & \\ \overline{f}_{i_{s-1}}(\omega_s) = \delta_s r_1 r_2 \cdots r_{s-1} & \text{for } 2 \leq s \leq n \\ \overline{f}_{i_s}(\omega_s) = -\delta_s r_1 r_2 \cdots r_{s-1} & \text{for } 2 \leq s \leq n \\ \overline{f}_j(\omega) = 0 & \text{for all other } \omega \text{ and } j. \end{array}$$

Using Equation (12) repeatedly (here is also where we use the assumption that X is non-crossing, which ensures that in each partition element of every player i there are at most two states at which $\overline{f}_i$ takes on non-zero values) we deduce that $E_{i_s}(\overline{f}_{i_s}|\omega_s) > 0$ for $1 \leq s \leq n+1$.

We still need to ensure that the players who are not in P have positive expectations at the states participating in X. To do so, note the following: since $r_1 r_2 \cdots r_n \delta_n < 1$, we can choose an ε_{n+1} such that $1 - \varepsilon_{n+1} > r_1 r_2 \cdots r_n \delta_n$. Similarly, since for any $2 \leq s \leq n$, $\delta_s r_1 r_2 \ldots r_{s-1} > \delta_{s-1} r_1 r_2 \ldots r_{s-1}$, we can choose ε_s such that

$$\delta_s r_1 r_2 \ldots r_{s-1} - \varepsilon_s > \delta_{s-1} r_1 r_2 \ldots r_{s-1}.$$

At each state in $\omega_s \in X$, therefore, we intuitively can take away a positive part of the positive value of $\overline{f}_{i_{s-1}}(\omega_s)$ and 'redistribute' it among the other players. This enables the following construction:

$$\begin{array}{ll} f_{i_s}(\omega_s) = \overline{f}_{i_s}(\omega_s) & \text{for } 1 \leq s \leq n \\ f_{i_{s-1}}(\omega_s) = \overline{f}_{i_{s-1}}(\omega_s) - \varepsilon_s & \text{for } 1 \leq s \leq n \\ f_j(\omega_s) = \varepsilon_s/(m-2) & \text{for } 1 \leq s \leq n, \text{ for } j \in I, j \neq i_s, i_{s-1} \\ f_j(\omega) = 0 & \text{for all other } \omega \text{ and } j. \end{array}$$

By construction, $f = \{f_1, \ldots, f_m\}$ is an agreeable bet relative to τ^X, which is what we needed to show. ∎

The following proposition, one half of the no betting characterisation for finite type spaces, has several proofs in the literature, all of which ultimately rely on convex separation theorem, or equivalents thereof. The proof presented here (building on the previous lemmas) is, in contrast, constructive and entirely combinatorial.

PROPOSITION 6. *Let τ be a type space over $(\Omega, \mathbf{\Pi})$, where Ω is a finite state space. If τ has no common prior then there exists an agreeable bet relative to τ.*

Proof. Recalling that in a finite state space every improper prior can be normalised, and is therefore a proper prior, we may refer to all the previous lemmas and apply them restricting attention to the special case of common priors, rather than the more general common improper prior.

Let $S \subseteq \Omega$ be the set of non-singularly positive states in Ω, and (using Lemma 1) decompose S disjointly as $S = \cup_{j=1}^{k} T_j$, where each T_j is non-singularly positive and connected. By Proposition 3, for each $1 \leq j \leq k$, there is a non-crossing cycle c_{T_j} contained in T_j such that $\mathrm{tr}(c_{T_j}) \neq 1$. By Proposition 4, there exists $f^{T_j} = \{f_1^{T_j}, \ldots, f_k^{T_j}\}$ such that $f_i^{T_j}(\omega) = 0$ for all $\omega \notin T_j$ and $i \in I$, and f^{T_j} is an agreeable bet relative to τ^{T_j}.

Let Z be the set of zero states, and denote by Q the complement of $S \cup Z$. Define f^Q in stages as follows.

In stage 0, let W_0 be the set of singular states in Q. For every $\omega \in W_0$ decompose I as $I = J \cup K$, where $\tau_j(\omega) > 0$ for all $j \in J$ and $\tau_i(\omega) = 0$ for $i \in K$, and choose an arbitrary $i' \in K$. Then define $g_j^0(\omega) = 1/|J|$ for every $j \in J$, with $g_{i'}^0(\omega) = -1$, and $g_i^0(\omega) = 0$ for all other $i \in K$, $i \neq i'$. Set $g_k^0(\omega) = 0$ for all states ω not in W_0 and all players $k \in I$.

Note, for the rest of this proof, that by definition every state in $Q \setminus W_0$ is a positive state.

In stage 1, let W_1 be the set of all states $\omega \in Q \setminus W_0$ satisfying the property that there is at least one player i, and a state $\omega' \in W_0$, such that $\omega \in \pi_i(\omega')$. For each $\omega \in W_1$, choose ω' as just described. If $t_i(\omega') > 0$, then by construction ω' is an i-surplus state relative to g^0. Hence we can apply a process similar to that described in the paragraph preceding and following Equation 11 to extend g^0 to a function g^1 such that $g_i^1(\omega) < 0$, but for all $j \neq i$, $g_j^1(\omega) > 0$, $\sum_{j \in I} g_j^1(\omega) = 0$, and is a surplus state for all players.

If $t_i(\omega') = 0$, then note that since ω is positive but not contained in any non-singularly positive set, there must be a chain $c = \omega \stackrel{i_0}{\rightarrow} \omega_1 \stackrel{i_1}{\rightarrow} \cdots \stackrel{i_{n-1}}{\rightarrow} \omega_n$ entirely contained in $Q \setminus W_0$ such that there is a player i and a state $\omega' \in X_0$ satisfying $\omega_n \in \pi_i(\omega')$ and $t_i(\omega') > 0$. But then we can apply the same argument as in the previous paragraph to define extend g^0 to g^1 by induction over all the states $\omega_n, \omega_{n-1}, \ldots, \omega$, yielding a function such that $\sum_{j\in I} g_j^1(\omega) = 0$, and is a surplus state for all players.

In all stages $l > 1$, in stage $l-1$, a function g^{l-1} has been defined such that each state $\omega \in W_{l-1}$ is an i-surplus state for every player i relative to g^{l-1}. Denote by W_l the set containing every state $\omega \in Q \setminus W_{l-1}$ satisfying the property that there is at least one $i \in I$ and $\omega \in W_{l-1}$ such that $\omega' \in \pi_i(\omega)$. Since ω is an i-surplus state relative to g^{l-1}, we can again apply the same technique as in the previous paragraphs to extend g^{l-1} to g^l.

By the finiteness of Ω, this iterative process ends after a finite number of stages r. Finally, set $f^Q := g^r$, and define

$$f := f^{T_1} + f^{T_2} + \ldots f^{T_k} + f^Q.$$

It is straightforward to check that f, by construction, is an agreeable bet. ∎

LEMMA 3. *Let $(d_n)_{n=1}^{\infty}$ be an increasing sequence of integers. Let $(C_n)_{n=1}^{\infty}$ be a sequence satisfying the properties that each C_n is a non-trivial cone in $\mathbb{R}^{d_n}$ and $\varphi_n(C_{n+1}) \subseteq C_n$, where φ_n is the projection from $\mathbb{R}^{d_{n+1}}$ to $\mathbb{R}^{d_n}$. Assume in addition that for each $m > n$ the projection of C_m to C_n is a non-trivial cone. Then there exists a sequence of non-trivial cones $(C'_n)_{n=1}^{\infty}$ such that $C'_n \subseteq C_n$ and $\varphi_n(C'_{n+1}) = C'_n$ for all n.*

Proof For each n define a sequence $(C_n^k)_{n=1}^{\infty}$. To begin with, set $C_n^1 = C_n$ and then set $C_n^{k+1} = \varphi(C_{n+1}^k)$. The sets C_n^k are closed cones. Moreover, $C_n^{k+1} = \varphi_n\varphi_{n+1}\cdots\varphi_{n+k-1}(C_{n+k}^1)$, hence C_n^{k+1} is the projection of C_{n+k} and therefore by the assumption in the statement of the lemma it is a non-trivial cone. We show by induction that $C_n^{k+1} \subseteq C_n^k$ for all $k \geq 1$. For $k=1$ this holds by the properties satisfied by the sequence $(C_n)_{n=1}^{\infty}$. If it holds for k then, since projection functions preserve inclusion, $C_n^{k+1} = \varphi_n(C_{n+1}^k) \subseteq \varphi_n(C_{n+1}^{k-1}) = C_n^k$. Define $C'_n = \cap_{k\geq 1} C_n^k$. As an intersection of a decreasing sequence of non-trivial closed cones, C'_n is a non-trivial cone. Moreover, $\varphi_n(C'_{n+1}) = \varphi(\cap_{k\geq 1} C_{n+1}^k) = \cap_{k\geq 1}\varphi_n(C_{n+1}^k) = \cap_{k\geq 1} C_n^{k+1} = C'_n$. ∎

PROPOSITION 7. *Let τ be a type space over $\{\Omega, \mathbf{\Pi}\}$. There exists a common improper prior for τ if and only if there exists an increasing sequence $(\Omega_n)_{n=1}^{\infty}$ of finite subsets of Ω such that $\cup_n \Omega_n = \Omega$ and for each n there is a common prior for τ^{Ω_n}.*

Proof. Suppose that μ is a common improper prior for τ. Choose a finite subset Ω_1 of Ω such that $\mu(\Omega_1) > 0$ and then let $(\Omega_n)_{n=1}^{\infty}$ be any sequence of finite subsets of Ω satisfying $\Omega_{n-1} \subseteq \Omega_n$ such that $\cup_n \Omega_n = \Omega$. We have that $\mu(\Omega_n) > 0$ for all n. Let μ^{Ω_n} be the restriction of μ to τ^{Ω_n} and let $\overline{\mu}^{\Omega_n}(\omega) := \mu^{\Omega_n}(\omega)/\mu(\Omega_n)$ for each $\omega \in \Omega_n$. $\overline{\mu}^{\Omega_n}$ is clearly a common prior for τ^{Ω_n}.

In the other direction, suppose there exists a sequence $(\Omega_n)_{n=1}^{\infty}$ satisfying the condition in the statement of the proposition. Let $d_n = |\Omega_n|$ and let C_n be the non-trivial cone of common priors for τ^{Ω_n} in $\mathbb{R}^{d_n}$. Clearly $\varphi_n(C_{n+1}) \subseteq C_n$ for each n. We show that for $m > n$ the projection of C_m on C_n is a non-trivial cone. To see this, let $\mu \in C_m$. If $\mu(\Omega_n) > 0$ then the projection is straightforwardly non-trivial. If $\mu(\Omega_n) = 0$ and $\mu' \in C_n$ then $\mu + \mu' \in C_m$ and again we have that the projection is non-trivial. Consider the sequence of non-trivial cones $(C'_n)_{n=1}^{\infty}$ constructed in Lemma 3. Choose $\overline{\mu} \in C'_1$. The pointwise limit of the vectors $\overline{\mu}$, $\varphi_1^{-1}(\overline{\mu})$, $\varphi_2^{-1}(\varphi_1^{-1}(\overline{\mu})), \ldots$ is a common improper prior. ∎

Proof of Theorem 1(b). Let τ be a type space over $\{\Omega, \mathbf{\Pi}\}$. If Ω is finite, then Proposition 6 suffices. Suppose therefore that Ω is countably infinite. Applying Proposition 7, let $(\Omega_n)_{n=1}^{\infty}$ be an arbitrary increasing sequence of finite subsets such that $\cup_n \Omega_n = \Omega$. Since we are assuming that τ has no *cip*, there must be infinitely many n's such that there is no common prior for Ω_n. Thus we may assume without loss of generality that for each n there is no common prior for τ^{Ω_n}. By Proposition 6, for each n there is an agreeable bet f^n on Ω_n bounded by 1. Let $\hat{f}^n$ be the extension of f^n to all of Ω defined to be 0 at all points not in f^n. Then $f = \sum_n 2^{-n} \hat{f}^n$ is a bounded agreeable bet over Ω. ∎

As mentioned in the introduction, putting together the elements of the proofs in this section yields an algorithm that can be applied in finite type spaces. The algorithm determines whether the space has a common prior, by listing the connected non-singularly positive subspaces and checking whether there is such a subspace such that every non-crossing cycle contained in it satisfies $\mathrm{tr}(c) = 1$ (see Proposition 3). If it does have a common prior, the algorithm then constructs the common prior, by the method in the proof of Proposition 2. If the space does not have a common prior, the algorithm constructs an agreeable bet by the method in the proof of Proposition 6, thus finding a sequence of random variables about whose expected values the players 'agree to disagree'.

EXAMPLE 7. This simple example illustrates how to compute an acceptable bet given a knowledge space without a common prior. The space of states is $\Omega = \{w, x, y, z\}$. Player 1's partition of Ω is $\{\{w, x\}, \{y, z\}\}$, with corresponding posteriors $\{\{1/8, 7/8\}, \{4/5, 1/5\}\}$ and Player 2's partition is $\{\{w, y\}, \{x, z\}\}$, with corresponding posteriors $\{\{1/2, 1/2\}, \{1/4, 3/4\}\}$.

The sequence w, x, y, z forms a cycle, with the corresponding type ratios $\frac{1}{7}$, $\frac{1}{3}$, $\frac{1}{4}$ and 1. Multiplying them together gives $\frac{1}{84}$, hence the type ratio of the cycle is not equal to one, and therefore this type space has no common prior, hence there must exist an agreeable bet.

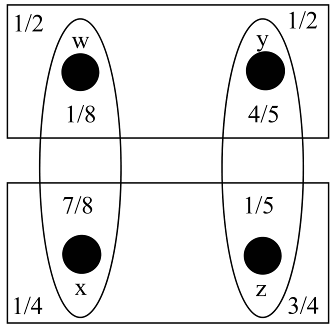

FIGURE 6. The knowledge space in Example 7.

To compute an agreeable bet, choose real numbers $1 < \delta_2 < \delta_3 < \delta_4 < 84$. For example, $let\delta_2 = 2$, $\delta_3 = 6$ and $\delta_4 = 8$. Use these to define a function $f(w) := -1$, $f(x) := \delta_2 \frac{1}{7} = \frac{2}{7}$, $f(y) := -\delta_3 \frac{1}{7}\frac{1}{3} = \frac{6}{21}$ and $f(z) := \delta_4 \frac{1}{7}\frac{1}{3}\frac{1}{4} = \frac{8}{84} = \frac{2}{21}$. $\{f, -f\}$ is an agreeable bet. ♦

4.5. Unbounded Bets.

DEFINITION 11. Define the *graph of knowledge* $G_{\mathbf{\Pi}}$ over $\mathbf{\Pi}$ to be the following undirected graph (V, E). The set of vertices V is the set of states Ω, i.e., $V = \Omega$. Two vertices ω_0 and ω_1 are connected by an edge in E if there is a chain of length 1 from ω_0 to ω_1. ♦

DEFINITION 12. For any state ω, define a *subtree of knowledge* T_ω to be a subtree of the graph of knowledge $G_{\mathbf{\Pi}}$ that is rooted at ω and satisfies the property that the path in the tree between ω and any of its descendants ω' is an alternating chain. For every $\omega' \in T_\omega$, denote the set of its children by $ch(\omega)$. ♦

A knowledge subtree T of $G_{\mathbf{\Pi}}$ that satisfies the property that every branch of T (which corresponds to a maximal alternating chain) has infinite height is an *infinite tree of knowledge*.

PROPOSITION 8. *If the graph of knowledge $G_{\mathbf{\Pi}}$ corresponding to $\mathbf{\Pi}$ contains an infinite subtree of knowledge T_{ω_0}, then any type space $\{\Omega, \mathbf{\Pi}, \tau\}$ over $\{\Omega, \mathbf{\Pi}\}$ has an agreeable bet.*

Proof. By assumption, there is a rooted subtree $T \subseteq G_{\mathbf{\Pi}}$ satisfying the property that every maximal alternating chain in T emanating from the root ω_0 is infinite.

Define a bet f in a recursive manner as follows. At step 1, start with the root, ω_0, of the subtree of knowledge. All the elements in $ch(\omega_0)$ (the set of children of ω_0) are members of the same partition of some player i_1.

Let f_{i_1} be defined over all $\omega \in ch(\omega_0)$ in such a way that $E_{i_1} f_{i_1}(\omega)$ is well-defined and positive for all $\omega \in ch(\omega_0) \cup \omega_0$. For every $j \in I$, $j \neq i_1$ and $\omega \in ch(\omega_0)$, let $f_j(\omega) = -f_{i_1}(\omega)/(n-1)$.

At step k, let Ψ_{k-1} be the set of all elements in T for which values of f were defined in stage $k-1$. For each $\omega \in \Psi_{k-1}$, the set of children $ch(\omega)$ is non-empty (by the assumption that T is an infinite tree), and all the elements in that set are members of the same partition of some player i_k. Let f_{i_k} be defined over all $\omega \in ch(\omega_0)$ in such a way that $E_{i_k} f_{i_k}(\omega)$ is well-defined and positive for all $\omega \in ch(\omega) \cup \omega$. For every $j \in I$, $j \neq i_k$ and $\omega' \in ch(\omega)$, let $f_j(\omega') = -f_{i_k}(\omega')/(n-1)$.

After a countable number of such steps, $f = \{f_1, \ldots, f_n\}$ is an agreeable bet relative to τ^T. An application of Proposition 4 completes the proof. ∎

COROLLARY 3. *If a knowledge space* $\{\Omega, \mathbf{\Pi}\}$ *contains an infinite alternating chain, then any type space* $\{\Omega, \mathbf{\Pi}, \tau\}$ *over* $\{\Omega, \mathbf{\Pi}\}$ *has an agreeable bet.*

Proof. An infinite alternating chain is an infinite tree of knowledge. ∎

Proof of Theorem 2. An infinite alternating chain can always be defined within the subset of states contained in partition elements that are of cardinality greater than 1, if that subset is infinite. ∎

Finally, the following example shows that if the condition in Theorem 2 does not hold, an agreeable bet may not exist.

EXAMPLE 8. The state space Ω is the set $\mathbb{Z}$ of all the integers, i.e., $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

There are two players, with partitions

$$\Pi_1(n) = \begin{cases} \{n\} & \text{if } n < 0 \\ \{0, 1, 2, \ldots\} & \text{if } n \geq 0 \end{cases}$$

and

$$\Pi_2(n) = \begin{cases} \{n\} & \text{if } n > 0 \\ \{0, -1, -2, \ldots\} & \text{if } n \leq 0 \end{cases}$$

Player 1	...	-3	-2	-1	0	1	2	3	...
Player 2	...	-3	-2	-1	0	1	2	3	...

FIGURE 7. The partition profile of Example 8.

Let p be any arbitrary common prior from which a type space over this partition profile is derived.

There is no bounded or unbounded agreeable bet. For suppose that f is an agreeable bet. Then:

(1) since $E_1 f(n) > 0$ for $n < 0$, it must be the case that $f(n) > 0$ for $n < 0$
(2) since $E_2 f(n) < 0$ for $n > 0$, it must be the case that $f(n) < 0$ for $n > 0$

It remains to determine the sign of $f(0)$. But if $f(0) \geq 0$, then $E_2 f(0) > 0$, and if $f(0) \leq 0$, then $E_1 f(0) < 0$. This is a contradiction. ♦

5. Iterated Expectations, Countable Spaces, and Common Priors

5.1. Markov Theory Preliminaries.

We recall some basic concepts and results from Markov theory. Let M be a Markov matrix over a state space Ω that is either finite or countable. Fix an ordering of the elements of Ω as $\{\omega_1, \omega_2, \ldots\}$.

States ω_i and ω_j *communicate* if there are $k, m \geq 0$ such that $M^k(\omega_i, \omega_j) > 0$ and $M^m(\omega_j, \omega_i) > 0$. This relation partitions the state space into equivalence classes. M is *irreducible* if all the states communicate. A state ω_i is said to have *period* d if $M^k(\omega_i, \omega_i) = 0$ whenever k is not divisible by d, and d is the largest integer with this property. A state with period 1 is *aperiodic*. Every pair of communicating states have the same period, hence an irreducible Markov matrix has a well-defined period.

A state ω_i is called *recurrent* if

$$\sum_{k=0}^{\infty} M^k(\omega_i, \omega_i) = \infty.$$

Otherwise the state is called transient. Every pair of communicating states are either both recurrent or both transient, have we can categorise irreducible Markov matrices into recurrent and transient matrices.

An irreducible aperiodic Markov matrix is called *null recurrent* if it is recurrent and for every pair of states ω_i, ω_j

$$\lim_{k \to \infty} M^k(\omega_i, \omega_j) = 0.$$

If M is recurrent but not null recurrent then it is called *positive recurrent*. Denoting

$$\gamma_i := \lim_{k \to \infty} M^k(\omega_i, \omega_i), \tag{13}$$

M is positive recurrent if and only if $\gamma_i > 0$ for all i. If M is irreducible, aperiodic, and positive recurrent it is termed *ergodic*.

A measure $\mu \in \Omega^{\mathbb{R}}$ satisfying $\mu \geq 0$ is an *invariant measure* of M if it is a left eigenvector of M, i.e. $\mu M = \mu$, or equivalently for each state ω':

$$\mu(\omega') = \sum_{\omega \in \Omega} \mu(\omega) M(\omega, \omega').$$

If μ is an invariant measure and also satisfies the property that it is normalisable, i.e. $\sum_{\omega \in \Omega} \mu(\omega) < \infty$, then it is termed an *invariant probability measure*.

The following are well-known results in Markov theory:

(1) If M is recurrent then it has an invariant measure that is unique up to scalar multiplication.
(2) M has a unique invariant probability measure if and only if M is ergodic.
(3) If M is ergodic, then the measure π given by $\pi(i) = \gamma_i > 0$ (where γ_i is defined in Equation 13) is the unique invariant probability measure.
(4) If M is irreducible, aperiodic, and recurrent, then $\lim_{k \to \infty} M^k$ exists. If M is ergodic, then

$$\lim_{k \to \infty} M^k = \mathbf{I}\pi.$$

If M is null recurrent, then $\lim_{k \to \infty} M^k = 0$.

5.2. Iterated Expectations and Common Priors over Countable State Spaces.

This section clearly owes a large debt to Samet (1998a), and most of the proof ideas are taken from that paper. The need to distinguish between null recurrent and positive recurrent classes, however, adds subtleties that are not present in the finite state space.

PROPOSITION 9. *Let σ be a permutation of I, and let $P \in \wedge\mathbf{\Pi}$ be an element of the meet. Then P is an irreducible, aperiodic, and recurrent class of M_σ. Thus, M_σ^P has a unique (up to scalar multiplication) invariant measure μ_σ^P on P.*

Proof. Let $i \in I$, and let $\omega, \omega' \in P \in \Pi_{\sigma(i)}$. Then

$$M_\sigma(\omega, \omega') \geq M_{\sigma(1)}(\omega, \omega) \cdots M_{\sigma(i-1)}(\omega, \omega) M_{\sigma(i)}(\omega, \omega') \times M_{\sigma(i+1)}(\omega', \omega') \cdots M_{\sigma(n)}(\omega', \omega').$$

Therefore, any two states in the same element of a partition of an arbitrary player communicate. Hence, if ω is in an equivalence class of states, then $\Pi_i(\omega)$, for each i, is a subset of this class. This means that each class is a union of elements of $\wedge\mathbf{\Pi}$. Also, for each $P \in \wedge\mathbf{\Pi}$, the probability of $\omega \in P$

staying in P under M_σ is 1, and therefore P is an irreducible equivalence class. That the Markov matrix M_σ is aperiodic and recurrent follows from the fact that $M_\sigma(\omega, \omega) > 0$ for every state. ∎

COROLLARY 4. *Let σ be a permutation of I, and let $P \in \wedge\mathbf{\Pi}$ be an element of the meet. If P is a positive recurrent class of M_σ, then M_σ^P has a unique invariant probability measure μ_σ^P.* ∎

PROPOSITION 10. *The following conditions are equivalent for each $P \in \wedge\mathbf{\Pi}$:*

(1) *p is the common prior on P.*
(2) *p is the invariant probability measure of M_i^P for every $i \in I$.*
(3) *The permutation matrix M_σ^P is ergodic for every permutation σ, and p is the invariant probability measure of M_σ^P.*

Proof. With slight abuse of notation, write M_i in place of M_i^P throughout this proof, for ease of exposition.

That (1) and (2) are equivalent is straightforward. To see that (2) implies (3), note that if p is an invariant probability measure of M_i for every i, then it must also be an invariant probability measure of M_σ for every permutation σ. But M_σ has an invariant probability measure if and only if it is ergodic.

Next, suppose that (3) is true and let p be the invariant probability measure of $M_{\sigma_1} := M_1 M_2 \cdots M_n$ that exists by (3). Then

$$pM_1 M_2 \cdots M_n = p.$$

Multiplying M_{σ_1} from the right by M_1 yields

$$pM_1 M_2 \cdots M_n M_1 = pM_1.$$

It follows that pM_1 is an invariant probability measure of

$$M_{\sigma_2} := M_2 \cdots M_n M_1.$$

However, by (3), p is also an invariant probability measure of M_{σ_2}, and by Corollary 4, each permutation matrix over restricted to P has a unique invariant probability measure. Hence $pM_1 = p$, and similarly, $pM_i = p$ for each $i \in I$. ∎

Proof of Theorem 3. Since $E_\sigma^k f = M_\sigma^k f$ for each f and k,

$$\lim_{k\to\infty} E_\sigma^k f = \lim_{k\to\infty} M_\sigma^k f.$$

As M_σ is recurrent for each σ by Proposition 9, $\lim_{k\to\infty} M_\sigma^k$ exists.

If τ has a common prior, then by Proposition 10, M_σ is ergodic for each σ, and therefore $\lim_{k\to\infty} M_\sigma^k = \mathbf{I}p_\sigma$, where p_σ is the unique invariant probability measure of M_σ. But by the common prior assumption, $p_\sigma = \varphi$

for all σ. Furthermore, $\varphi > 0$, hence the iterated expectations of f with respect to all permutations converge to the same non-zero limit pf.

In the other direction, suppose that for each bounded non-zero random variable $f \geq 0$, the iterated expectations of f, with respect to all permutations σ, converge to the same non-zero limit. Then it cannot be the case that $\lim_{k\to\infty} M_\sigma^k = 0$, hence M_σ is ergodic. It follows that $\lim_{k\to\infty} M_\sigma^k = \mathbf{I}p_\sigma$, where p_σ is the unique invariant probability measure of M_σ, and therefore the iterated expectation of every f with respect to σ converges to $p_\sigma f$. By the assumption that $p_\sigma f$ is the same for all permutations σ, there is a single p, $p = p_\sigma$ for all σ, such that p is the unique invariant probability measure of M_σ. We conclude from Proposition 10 that p is the common prior. ∎

Theorem 3 provides a characterisation of common priors in countable spaces using the iterated expectations criterion. That criterion cannot be used for identifying common improper priors in cases in which a common prior does not exist; by Proposition 10, if there is no common prior, then for at least one permutation σ, the permutation matrix M_σ^P is not ergodic. This in turn means that M_σ^P must be null recurrent, hence $\lim_{k\to\infty} M_\sigma^k = 0$.

6. Appendix

Proof of Lemma 1. Suppose that $S \subseteq \Omega$ is non-singularly positive, and let $\omega_0 \in S$ be chosen arbitrarily. Suppose that ω_0 is connected to ω_n by a positive chain $c = \omega_0 \stackrel{i_0}{\to} \omega_1 \stackrel{i_1}{\to} \cdots \stackrel{i_{n-1}}{\to} \omega_n$. Since ω_0 is in S and ω_1 is in the same partition element of player i_0 as ω_0, by Definition 6, ω_1 must also be in S, and continuing the same argument by induction, we conclude that all the elements in c are members of S. ∎

Proof of Proposition 2. Suppose that there exists a common improper prior p for τ. Let $S = \{\omega \in \Omega \mid p(\omega) > 0\}$. S is guaranteed to be positive, because $p \neq 0$. We next show that S is non-singularly positive: Suppose that for arbitrary $i \in I$ and $\omega \in S$, $\omega' \in \Pi_i(\omega)$. Furthermore, suppose that $\omega' \notin S$. Then $p(\omega') = 0$, while $p(\omega) > 0$. Hence $p(\Pi_i(\omega)) > 0$, and by the definition of an improper prior,

$$t_i(\omega') = \frac{p(\omega')}{p(\Pi_i(\omega))} = 0.$$

It follows from Definition 6 that S is non-singularly positive.

To complete this part of the proof, note that for any pair of states $\omega_1, \omega_2 \in S$ such that ω_1 and ω_2 are in the same element of Π_i for some $i \in I$, $\mathrm{tr}_t^i(\omega_1, \omega_2) = p(\omega_1)/p(\omega_2)$. It then easily follows from the definition of the type ratio of a chain that for any chain c entirely contained in S and connecting ω_0 and ω, one has $\mathrm{tr}_\tau(c) = p(\omega_0)/p(\omega)$.

Conversely, suppose that Ω has a non-singularly positive subspace S with respect to τ, and that for each ω_0 and ω in S, any pair of chains c and c' entirely contained in S connecting ω_0 to ω satisfy $\mathrm{tr}_\tau(c) = \mathrm{tr}_\tau(c')$. Using Lemma 1, we may assume that S is connected (replacing S by a connected subset of itself if necessary).

We will construct a *cip* p. For $\omega \notin S$, set $p(\omega) = 0$. Otherwise, fix $\omega_0 \in S$ and for each $\omega \in S$, let $p(\omega) = \mathrm{tr}(c)$ for some chain c from ω_0 to ω contained in S.

To see that p is a *cip* consider $\pi \in \Pi_i$. Suppose first that $\pi \cap S = \emptyset$. Then for all $\omega \in \pi$, $p(\omega) = 0$, hence $p(\pi) = 0$, and $p(\pi)t_i(\omega) = p(\omega)$ is satisfied.

Suppose instead that $\pi \cap S \neq \emptyset$, and that $\omega \in \pi \cap S$. Let c be a chain from ω_0 to ω entirely contained in S. For $\omega' \in \pi \cap S$, consider the chain $c' = c \xrightarrow{i} \omega'$. Then, by the definitions of tr and p, $p(\omega') = \mathrm{tr}(c') = \mathrm{tr}(c)\mathrm{tr}^i(\omega, \omega') = p(\omega)t_i(\pi, \omega')/t_i(\pi, \omega)$. Thus,

$$p(\pi) = \sum_{\omega' \in \pi \cap S} p(\omega') = [p(\omega)/t_i(\pi, \omega)] \sum_{\omega' \in \pi \cap S} t_i(\pi, \omega') = p(\omega)/t_i(\pi, \omega) < \infty$$

and

$$p(\omega) = p(\pi)t_i(\pi, \omega).$$

Finally, suppose that $\pi \cap S \neq \emptyset$, and that $\omega \in \pi$ is such that $\omega \notin S$. By construction, $p(\omega) = 0$, and by the assumption that S is non-singularly positive, it must be the case that $t_i(\omega) = 0$. We have already shown that $p(\pi) < \infty$, hence $p(\pi)t_i(\omega) = p(\omega)$ is satisfied. ∎

Proof of Corollary 2. It suffices to note the following: suppose that c_1 and c_2 are two distinct chains entirely contained in S connecting a pair of states ω and ω'. Then $\overline{c} := c_1 c_2^{-1}$ is a cycle connecting ω to itself. By Equation (6), $\mathrm{tr}(c_1) = \mathrm{tr}(c_2)$ if and only if $\mathrm{tr}(\overline{c}) = 1$. ∎

Proof of Proposition 3. If there exists a *cip*, then by Corollary 2, there is a non-singularly positive $S \subseteq \Omega$ such that every cycle contained in S has type ratio equal to 1, hence in particular every non-crossing cycle satisfies the same property.

In the other direction, if there does not exist a *cip*, then for every non-singularly positive $S \subseteq \Omega$, there is at least one cycle $\overline{c}$ entirely contained in S such that $\mathrm{tr}(\overline{c}) \neq 1$. Suppose that $\overline{c}$ is not non-crossing.

If $\overline{c}$ fails to be non-crossing because it is not alternating, this 'flaw' can easily be corrected: if two consecutive states ω_s and ω_{s+1} in $\overline{c}$ are identical, since that implies that $\mathrm{tr}^i(\omega_s, \omega_{s+1}) = 1$, the state ω_{s+1} is redundant and can be removed from $\overline{c}$ without changing the type ratio. Similarly, if ω_s, ω_{s+1}, ω_{s+2} and ω_{s+3} are consecutive states that are all members of the same partition element of player i, then $\mathrm{tr}^i(\omega_s, \omega_{s+1})\mathrm{tr}^i(\omega_{s+2}, \omega_{s+3}) = \mathrm{tr}^i(\omega_s, \omega_{s+3})$, hence we may remove ω_{s+1} and ω_{s+2} from $\overline{c}$ without changing the type ratio.

We will therefore assume that $\overline{c}$ is alternating but not non-crossing, and that we can then write $\overline{c} = \omega_0 \xrightarrow{i_0} \omega_1 \xrightarrow{i_1} \cdots \xrightarrow{i_{n-1}} \omega_n = \omega_0$, where $\omega_{s+1} \in \Pi_{i_s}(\omega_s)$ for

$s = 0, \ldots, n-1$, where there exists at least one pair r, k such that $k > r+1$, and $\omega_k \in \Pi_{i_r}(i_r)$.

We can 'shorten' $\overline{c}$ into another cycle:

$$\widehat{c} = \omega_0 \xrightarrow{i_0} \omega_1 \xrightarrow{i_1} \ldots \omega_r \xrightarrow{i_r} \omega_k \xrightarrow{i_k} \cdots \xrightarrow{i_{n-1}} \omega_n = \omega_0.$$

If $\mathrm{tr}(\widehat{c}) \neq 1$, then we have a cycle of type ratio not equal to 1, with a number of self-crossing points that is strictly less than the number of self-crossing points in $\overline{c}$, and we can continue by induction to apply the same process to $\mathrm{tr}(\widehat{c})$.

Suppose, therefore, that $\mathrm{tr}(\widehat{c}) = 1$. Denote:

$$c_0 = \omega_0 \xrightarrow{i_0} \omega_1 \xrightarrow{i_1} \cdots \xrightarrow{i_{r-1}} \omega_r,$$

$$c_k = \omega_k \xrightarrow{i_k} \omega_{k+1} \xrightarrow{i_{k+1}} \cdots \xrightarrow{i_{n-1}} \omega_n,$$

and

$$c_l = \omega_r \xrightarrow{i_r} \omega_{r+1} \xrightarrow{i_{r+1}} \cdots \xrightarrow{i_{k-1}} \omega_k.$$

Then $\overline{c} = c_0 c_l c_k$, and $\widehat{c} = c_0(\omega_r, \omega_k)c_k$. By assumption,

$$1 = \mathrm{tr}(\widehat{c}) = \mathrm{tr}(c_0)\mathrm{tr}^{i_r}(\omega_r, \omega_k)\mathrm{tr}(c_k).$$

It follows that $[\mathrm{tr}^{i_r}(\omega_r, \omega_k)]^{-1} = \mathrm{tr}^{i_r}(\omega_k, \omega_r) = \mathrm{tr}(c_0)\mathrm{tr}(c_k)$. We also assumed that $\mathrm{tr}(\overline{c}) \neq 1$, so $1 \neq \mathrm{tr}(c_0 c_l c_k) = \mathrm{tr}(c_0)\mathrm{tr}(c_k)\mathrm{tr}(c_l) = \mathrm{tr}^{i_r}(\omega_k, \omega_r)\mathrm{tr}(c_l)$.

Writing out the last inequality in full yields

$$\mathrm{tr}(\omega_k \xrightarrow{i_r} \omega_r \xrightarrow{i_r} \omega_{r+1} \xrightarrow{i_{r+1}} \cdots \xrightarrow{i_{k-1}} \omega_k) \neq 1.$$

But by Equation 7, $\mathrm{tr}(\omega_k \xrightarrow{i_r} \omega_r \xrightarrow{i_r} \omega_{r+1}) = \mathrm{tr}(\omega_k \xrightarrow{i_r} \omega_{r+1})$, hence

$$\mathrm{tr}(\omega_k \xrightarrow{i_r} \omega_{r+1} \xrightarrow{i_{r+1}} \cdots \xrightarrow{i_{k-1}} \omega_k) \neq 1.$$

We deduce then that the cycle $\widetilde{c} = \omega_k \xrightarrow{i_r} \omega_{r+1} \xrightarrow{i_{r+1}} \cdots \xrightarrow{i_{k-1}} \omega_k$ satisfies both that $\mathrm{tr}(\widetilde{c}) \neq 1$, and that it has a number of self-crossing points that is strictly less than the number of self-crossing points in $\overline{c}$. We can continue by induction to apply the same process to $\mathrm{tr}(\widetilde{c})$.

After applying this reasoning as often as necessary, we arrive at the existence of a cycle entirely contained in S with no self-crossing points, i.e., a non-crossing cycle, whose type ratio is not equal to 1, which is what we needed to show. ∎

Bibliography

Aumann, R. J. (1976), Agreeing to disagree, *Ann. Statist.*, 4(6), 1236–1239.

Aumann, R. J. (1998), Common Priors: A Reply to Gul, *Econometrica*, 66, 929–938.

Bonanno, G., and Nehring K. (1999), How to make sense of the common prior assumption under incomplete information, *International Journal of Game Theory*, 28, 409–434.

Feinberg, Y. (2000), Characterizing Common Priors in the Form of Posteriors, *Journal of Economic Theory* 91, 127–179.

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

Harsányi, J. C. (1967-1968), Games with Incomplete Information Played by Bayesian Players, *Management Science*, 14, 159–182, 320–334, 486–502.

Heifetz, A. (2006), The Positive Foundation of the Common Prior Assumption, *Games and Economic Behavior*, 56, 105–120.

Hellman, Z. (2011), Iterated Expectations, Compact Spaces, and Common Priors, *Games and Economic Behavior*, 72, 163–171.

Hellman, Z., and Samet, D. (2012), How Common are Common Priors, *Games and Economic Behavior*, 74 , 517–525.

Lehrer, E., and Samet, D. (2011), Agreeing to Agree, *Theoretical Economics*, 6, 269287.

Lehrer, E., and Samet, D. (2012), Belief Consistency and Trade Consistency, *mimeo*, Tel Aviv University.

Milgrom, P. and Stokey, N. (1982), Information, Trade and Common Knowledge, *Journal of Economic Theory*, 26, 17-27.

Morris, S. (1995), Trade with Heterogenous Prior Beliefs and Asymmetric Information, *Econometrica*, (62), 1327–1347.

Morris, S. (2002), Notes on Iterated Expectations, *Working Paper*.

Rodrigues-Neto, J. A. (2009), From Posteriors to Priors via Cycles, *J. of Economic Theory*, 144, 876–883.

Rubinstein, A., and Wolinsky, A., (1990), On the Logic of "Agreeing to Disagree" Type Results, *J. of Economic Theory*, 51, 184-193.

Samet, D. (1998a), Iterated Expectations and Common Priors, *Games and Economic Behavior*, (24), 131–141.

Samet, D. (1998b), Common Priors and Separation of Convex Sets, *Games and Economic Behavior*, (24), 172–174.

CHAPTER 5

A Game with no Bayesian Approximate Equilibrium

Abstract

R. Simon (2003) presented an example of a 3-player Bayesian games with no Bayesian equilibria but has been an open question whether or not there are games with no Bayesian approximate equilibria. We present an example of a Bayesian game with two players, two actions and a continuum of states that possesses no Bayesian approximate equilibria, thus resolving the question. As a side benefit we also have for the first time an an example of a 2-player Bayesian game with no Bayesian equilibria and an example of a strategic-form game with no approximate Nash equilibria. The construction makes use of techniques developed in an example by Y. Levy of a discounted stochastic game with no stationary equilibria.

A Game with no Bayesian Approximate Equilibria

1. Introduction

One of the seminal contributions of Harsányi (1967) was the analysis of Bayesian games for studying games of incomplete information, which included showing that every finite Bayesian game (finite number of players, finite actions, finite states of the world) has a Bayes-Nash, or Bayesian, equilibrium. The fact that modellers could safely assume the existence of at least one equilibrium was undoubtedly an element in the widespread acceptance of Bayesian games in modelling a wide range of economic situations. Indeed, at this point it is impossible to imagine modern game theory and economic modelling without the theory of Bayesian games.

The question of the existence or non-existence[1] of Bayesian equilibria in games with uncountably many states, however, remained open for many years, until Simon (2003) gave a negative answer by presenting an example of a three-player Bayesian game with no Bayesian equilibrium.

That important result left in its wake (at least) two open questions: (1) are there examples of games that have no Bayesian ε-equilibria?; (2) are there examples of two-player games that have no Bayesian equilibria? In particular, a negative answer to the first question would imply that modellers could always assume that Bayesian equilibria can be approximated as closely as desired in games with uncountably many states, thus significantly weakening Simon (2003)'s result,

We show here, however, that the answer to both questions is yes by constructing a two-player Bayesian[2] game with no Bayesian ε-equilibria. As a side-benefit, the example also shows that there there exist strategic-form games with a continuum of players and no Nash ε-equilibria[3] and that there

[1] By the existence of an equilibrium we mean the existence of a measurable equilibrium. There are several reasons for restricting attention to measurable strategies (and hence measurable equilibria); to consider just two reasons, if a strategy is not measurable it cannot be constructed explicitly, and the payoffs of non-measurable strategies haven't got well-defined expected values. Measurability has in fact been included as a basic requirement in the definition of an equilibrium over uncountable spaces since the earliest literature on the subject (see Schmeidler (1973) for one such example). We therefore throughout this paper use the term 'existence of an equilibrium' as synonymous with 'existence of a measurable equilibrium' without further qualification.

[2] The game constructed here is not only a Bayesian game, it is an *ergodic game* as defined in Simon (2003).

[3] This result does not contradict the result in Schmeidler (1973), which assumes that no deviation from equilibrium undertaken by a finite number (or even a measure zero set) of players can affect payoffs; we do not assume that here. Sion and Wolfe (1957) presents an example of a finite-player game with no equilibrium, but the example there assumes each

exist two-player Bayesian games with no Harsányi equilibria (meaning *ex ante* Nash equilibria over the common prior of a Bayesian game), which had also been open questions.

We make extensive use of techniques developed in Levy (2012) in his paper on stochastic games without stationary Nash equilibria. The fact that these techniques have now been shown to be useful for generating counter-examples in separate subject fields (stochastic games and Bayesian games) may indicate that they have potential application in many other fields of interest.

The significance of counter-examples to the existence of equilibria and approximate equilibria such as the example here (and those in Simon (2003) and Levy (2012)) is that they serve as a sharp warning signal to modellers: although you routinely assume the existence of equilibria when you work with finite games, you cannot automatically do so in games with an uncountable number of states. A large percentage of economic models rely on the use of uncountably many states to represent quantities such as prices (as in models of auctions or bargaining, such as that of Chatterjee and Samuelson (1983) for example), profits and outputs in market models (for example Radner (1980)), time, accumulated wealth or stocks, population percentages, share percentages, and so forth.

In addition, an extensively-used approach to dealing with a Bayesian game with a finite but large number of states is to analyse instead a similar game with a continuum of states. Myerson (1997), for example, informs readers of Chapter 2 of his textbook on game theory, when referring to Bayesian games, that "it is often easier to analyze examples with infinite type sets than those with large finite type sets." Given this, it is important for modellers working with Bayesian games with uncountably many states to keep in mind that they cannot blindly rely on the well-known results in finite games guaranteeing the existence of equilibria and approximate equilibria.

2. Preliminaries and Notation

2.1. Information Structures and Knowledge.

A *space of states* is a pair $(\Omega, \mathbf{B})$ composed of a *set of states* Ω and a σ-field $\mathbf{B}$ of measurable subsets (*events*) of Ω.

We will work throughout this paper with a two-element set of players I. The two players will be denoted by Player Red and Player $Green$ (with capitalised initial letters). An *information structure* over the state space $(\Omega, \mathbf{B})$

player has a continuum of actions while we assume that each player has a finite action space.

is then given by a pair of partitions of Ω labelled Π_R and Π_G, respectively, of Player *Red* and Player *Green*. For $i \in I$ and for each state $\omega \in \Omega$ we denote by $\Pi_i(\omega)$ the element in Π_i that contains ω. Furthermore, denote by Γ_i the sub-σ-algebra of $\mathbf{B}$ generated by Π_i.

2.2. Types and Priors.

A *type function* t_i of player i for $(\Omega, \mathbf{B}, (\Pi_i)_{i \in I})$ is a function $t_i : \Omega \to \Delta(\Omega)$ from states to probability measures over $(\Omega, \mathbf{B})$ such that the mapping $t_i(\cdot)$ satisfies:

(1) $t_i(\omega)(E)$ is measurable for any fixed event E,
(2) $t_i(\omega)(\Pi_i(\omega)) = 1$,
(3) $t_i(\omega) = t_i(\omega')$ for all $\omega' \in \Pi_i(\omega)$.

For each ω, $t_i(\omega)$ is called player i's type at ω. Therefore, a quintuple $(I, \Omega, \mathbf{B}, (\Pi_i)_{i \in I}, (t_i)_{i \in I})$, where each t_i is a type function, is a *type space*.

A probability measure μ_i over $(\Omega, \mathbf{B})$ is a *prior* for a type function t_i if for each event A

$$\mu_i(A) = \int_\Omega t_i(\omega)(A) \, d\mu_i(\omega). \tag{14}$$

A probability measure μ that is a prior for each of the players' type function in a type space is a *common prior*.

2.3. Agents Games.

Recall the definition of the agents game $\mathbf{K}$ associated with a Bayesian game $\mathbf{B}$: $\mathbf{K}$ is a strategic-form game whose set of players, which is a measurable space, has a (measurable) bijection η with the the set of all types of all the players in $\mathbf{B}$. The action set of each player θ in $\mathbf{K}$ is equal to the action set of the player j in $\mathbf{B}$ associated with $\eta(\theta)$, and the payoff to player θ for an action profile is the corresponding expected payoff of j at $\eta(\theta)$. Every strategy $\hat{\psi}$ of $\mathbf{B}$ is naturally associated in this way with a strategy ψ in $\mathbf{K}$.

The analysis of the equilibria of a Bayesian game $\mathbf{B}$ can be accomplished by analysing the associated strategic-form game $\mathbf{K}$ in the sense that, for any $\varepsilon \geq 0$, the strategy ψ in $\mathbf{K}$ is a (measurable) Nash ε-equilibria if and only if $\hat{\psi}$ is a (measurable) Bayesian ε-equilibria in $\mathbf{B}$.

3. Constructing the Bayesian Game

We define in this section an ergodic game $\mathbf{B}$ that we will eventually show possesses no Bayesian ε-equilibrium.

3.1. The State Space.

Let X be the set of infinite sequences of 1 and -1, i.e. $X := \{-1, 1\}^{\mathbb{Z}_{\geq 0}}$. A generic element $x \in X$ is a sequence $x_0, x_1, x_2, \ldots$, where we denote the i-th coordinate of x by x_i.

Next, define the following two sets:

$$red = \{r\} \times X$$

and

$$green = \{g\} \times X,$$

that is, a generic element of red is $(r, x_0, x_1, x_2, \ldots)$ and a generic element of green is $(g, x_0, x_1, x_2, \ldots)$.

Our state space is $\Omega := red \cup green$. The measure μ we will work with over Ω gives each of red and $green$ the Lebesgue measure over X and independently gives equal probability to red and to $green$.

NOTATION 1.

- Denote by ι the operator on Ω defined by

$$\iota(r, x_0, x_1, x_2, \ldots) = (r, -1 \cdot x_0, x_1, x_2, \ldots)$$

$$\iota(g, x_0, x_1, x_2, \ldots) = (g, -1 \cdot x_0, x_1, x_2, \ldots).$$

  Note that ι is colour-preserving.
- Denote by S the operator on Ω defined by

$$S(r, x_0, x_1, x_2, \ldots) = (g, x_1, x_2, x_3 \ldots)$$

$$S(g, x_0, x_1, x_2, \ldots) = (r, x_1, x_2, x_3 \ldots).$$

  S is the product of the measure preserving involution $r \leftrightarrow g$ and the Bernoulli shift, which is well-known to be measure preserving,[4] hence S is measure preserving. Denote

$$S_+^{-1}(r, x_0, x_1, x_2, \ldots) = \{(g, 1, x_0, x_1, x_2, \ldots)\}$$

  and

$$S_-^{-1}(r, x_0, x_1, x_2, \ldots) = \{(g, -1, x_0, x_1, x_2, \ldots)\}$$

  with the same operator defined for $S_+^{-1}(g, x_0, \ldots)$ and $S_-^{-1}(g, x_0, \ldots)$ with the colours reversed . Then $S^{-1}(\omega) = S_+^{-1}(\omega) \cup S_-^{-1}(\omega)$, hence it maps each point to *two* points,

[4] See, for example, Halmos (1956).

3.2. The Knowledge Partitions and Type Functions.

We next define the partitions of Player *Red* and Player *Green*.

A generic element of Π_R is of the form

$$\{(r, x_1, x_2, x_3, \ldots),\ (g, 1, x_1, x_2, \ldots),\ (g, -1, x_1, x_2, \ldots)\}$$

A generic element of Π_G is of the form

$$\{(g, x_1, x_2, x_3, \ldots),\ (r, 1, x_1, x_2, \ldots),\ (r, -1, x_1, x_2, \ldots)\}$$

Another way of expressing this, using the above-defined operators, is: If ω is a *red* state then

$$\Pi_G(\omega) = \{\omega, \iota(\omega), S(\omega)\}$$
$$\Pi_R(\omega) = \{\omega\} \cup S^{-1}(\omega)$$

If ω is a *green* state then the partition elements are exactly the same with colours reversed

$$\Pi_R(\omega) = \{\omega, \iota(\omega), S(\omega)\}$$
$$\Pi_G(\omega) = \{\omega\} \cup S^{-1}(\omega)$$

To complete the description of the types, all that's left is to give the probabilities, which for all partition elements are:

$$\frac{1}{4}, \quad \frac{1}{4}, \quad \frac{1}{2}$$
$$\{\omega, \iota(\omega), S(\omega)\}$$

In words, the type function of Player *Green*, t_G, ascribes 1/2 to the *green* state in each of Player *Green*'s partition elements and 1/4 to each of the *red* states in his partition elements. The type function of Player *Red*, t_R, ascribes 1/2 to the *red* state in each of Player *Red*'s partition elements and 1/4 to each of the *green* states in her partition elements.

LEMMA 4. *The functions t_R and t_G satisfy the conditions for being type functions with μ as their common prior.*

The proof of Lemma 4 is in the appendix.

3.3. The Game Forms.

The action set we will work with is identical at every state for both players: $A_R^\omega = A_G^\omega = \{U, D\}$, for all $\omega \in \Omega$.

The payoff functions are more complicated. Let ρ_G be Player *Green*'s payoff function and ρ_R be Player *Red*'s payoff function, i.e.,

$$\rho_G, \rho_R : \Omega \times \{U, D\} \times \{U, D\} \to \mathbb{R}.$$

- For all $\omega \in red$, $\rho_G(\omega, \cdot, \cdot) = 0$, and similarly for all $\omega \in green$, $\rho_R(\omega, \cdot, \cdot) = 0$.

  In words, at red (respectively, $green$) states, Player $Green$ (respectively, Player Red) gets payoff 0 *no matter what* actions are played by him or the other player; strategically, Player $Green$ (respectively, Player Red) can ignore the red (respectively, $green$) states.
- For $\omega \in green$, $\rho_G(\omega, \cdot, \cdot)$ is given in Table 1, where Player $Green$ is row and Player Red is column.

  Note that this means that Player $Green$ cannot strategically ignore the $green$ states. Since each of his partition elements contains both $green$ and red states, he will focus his attention at each partition element on what might occur at the $green$ state.

If $x_0 = 1$:

	U	D
U	1	0
D	0.3	0.3

If $x_0 = -1$:

	U	D
U	0.7	0.7
D	1	0

TABLE 1. The payoff matrix.

- For $\omega \in red$, $\rho_R(\omega, \cdot, \cdot)$ is also given by Table 1, but now Player Red is row and Player $Green$ is column.

  Note that this means that Player Red cannot strategically ignore the red states. Since each of her partition elements contains both $green$ and red states, she will focus her attention at each partition element on what might occur at the red state.

This completes the construction of the Bayesian game $\mathbf{B}$.

3.4. Reduction to an Agents Game.

In the agents game, each type becomes an agent. A type of Player Red has the form

$$\{(r, x_0, x_1, x_2 \ldots),\ (g, 1, x_0, x_1, \ldots),\ (g, -1, x_0, x_1, \ldots)\}$$

We can therefore *uniquely* identify this type of Player Red by the state $(r, x_0, x_1, x_2 \ldots)$, and in this way we identify $\{\Pi_R(\omega)\}_{\omega \in \Omega}$ with $r \times X$. Hence we can consider Player Red's agents to be the set of red states and by the same reasoning, we can consider Player $Green$'s agents to be the set of $green$ states.

Formally, we have a bijection η between Ω and the collection of types $\{\Pi_R(\omega)\}_{\omega\in\Omega} \cup \{\Pi_G(\omega)\}_{\omega\in\Omega}$ as follows:

(15) $$\eta(\omega) = \begin{cases} \Pi_R(\omega) & \text{if } \omega \in \mathit{red} \\ \Pi_G(\omega) & \text{if } \omega \in \mathit{green}. \end{cases}$$

To form the agents game $\mathbf{K}$ corresponding to $\mathbf{B}$, we use this bijection. Abusing notation, we will use ω to identify both the state and the agent associated with that state. We will use the same measure μ over $(\Omega, \mathcal{B})$ in both $\mathbf{B}$ and $\mathbf{K}$. The agents in $\mathbf{K}$ all share the same action set, $A_\omega = \{U, D\}$, hence each strategy profile ψ of $\mathbf{K}$ is given by $\psi : \Omega \to \Delta(\{U, D\})$.

The definition of the mixed strategy profile $\hat{\psi}$ in $\mathbf{B}$ associated with a mixed strategy profile ψ in $\mathbf{K}$ explicitly gives

$$\psi(\omega) = \begin{cases} \hat{\psi}_1(\omega) & \text{if } \omega \text{ is odd} \\ \hat{\psi}_2(\omega) & \text{if } \omega \text{ is even.} \end{cases}$$

In moving from the Bayesian game $\mathbf{B}$ to the agents game $\mathbf{K}$, the payoff functions ρ_G and ρ_R of $\mathbf{B}$ induce payoff functions

$$u_\omega : \{U, D\} \times \{U, D\} \to \mathbb{R}$$

of $\mathbf{K}$ (one for each agent ω). The payoff function u_ω of each agent ω depends only on the action $a(\omega)$ chosen by agent ω and the action $a(S(\omega))$ of agent $S(\omega)$; in effect, each agent ω perceives himself or herself as playing a game against $S(\omega)$.

To see this, consider for example Player *Red*'s perspective when the true state of the world is a *green* state ω. The relevant partition element for Player *Red* is $\Pi_R(\omega) = \{\omega, \iota(\omega), S(\omega)\}$. By construction, $S(\omega)$ is a *red* state and is the only *red* state in $S(\omega)$. Since Player *Red*'s action choice can lead to a non-zero payoff for him only at $S(\omega)$, for calculating his expected payoff at $\Pi_R(\omega)$ he needs only to take into account the state of nature associated with $S(\omega)$ and accordingly to plan a best reply to Player *Green*'s actions at $S(\omega)$. Similar considerations then imply that at $S(\omega)$ Player *Green* need only plan a best reply to Player *Red*'s action at $S^2(\omega)$, and so forth.

The payoff function in the agents game is then given by the matrices of Table 2 at each ω, where the row player is ω and the column player is $S(\omega)$.

3.5. Observations on Equilibria in K.

Since there is an equivalence between the Bayesian ε-equilibria of a Bayesian game and the Nash ε-equilibria of the associated agents game, and in addition every strategy profile of the Bayesian game is measurable if and only if its associated profile of mixed strategies in the agents game is also

If $x_0 = 1$:

	U	D
U	1	0
D	0.3	0.3

If $x_0 = -1$:

	U	D
U	0.7	0.7
D	1	0

TABLE 2. The payoff matrices in the agents game, where the row player is ω and the column player is $S(\omega)$

measurable, for showing that there is no measurable Bayesian ε-equilibrium in $\mathbf{B}$ it suffices to show that there is no measurable Nash ε-equilibrium in $\mathbf{K}$.

From here to the end of the paper, fix the following values

$$0 < \varepsilon < \frac{1}{50};\ \delta = 10\varepsilon < \frac{1}{5}.$$

Assume, by way of contradiction, that there exists a measurable Nash ε-equilibrium in $\mathbf{K}$ denoted by ψ.

DEFINITION 13. An agent $\omega \in \Omega$ is *U-quasi-pure* (respectively, *D-quasi-pure*) if under ψ he plays U with probability greater than $1 - \delta$ (respectively, less than δ).

Denote

$$\Xi_U = \{\omega \mid \omega \text{ is } U\text{-quasi-pure}\},$$

$$\Xi_D = \{\omega \mid \omega \text{ is } D\text{-quasi-pure}\}$$

and

$$\Xi_M = \Omega \setminus (\Xi_U \cup \Xi_D).$$

LEMMA 5. *If $S(\omega)$ is quasi-pure under ψ then so is ω (i.e., $S(\omega) \in (\Xi_U \cup \Xi_D)$ implies $\omega \in (\Xi_U \cup \Xi_D)$). If the former is a-quasi-pure (for $a \in \{U, D\}$) then the latter is as well if and only if $x_0 = 1$ (i.e., if $S(\omega) \in \Xi_U$ then $\omega \in \Xi_U$ only if $x_0 = 1$ and similarly if $S(\omega) \in \Xi_D$, then $\omega \in \Xi_D$ only if $x_0 = 1$).*

Proof. The proof is identical to the proof of Lemma 3.3.4 in Levy (2012). The intuition behind the proof is as follows: a careful analysis of the matrices in Table 2 shows that under an ε-equilibrium if $S(\omega)$ is quasi-pure then ω will want to 'quasi-match' iff $x_0 = 1$ and to 'quasi-mismatch' iff $x_0 = -1$. ∎

LEMMA 6. *For all $\omega \in \Omega$, at least one of the two states in $S^{-1}(\omega)$ is in Ξ_U or Ξ_D (even if ω itself is not quasi-pure).*

Proof. The proof is identical to the proof of Lemma 3.3.5 in Levy (2012). The intuition behind the proof is as follows: if $S(\omega)$ plays U with probability greater than 2/5 then if $x_0 = 1$ agent ω will 'quasi-match' and play U with probability greater than $1 - \delta$. If agent $S(\omega)$ plays U with probability less than 3/5 (i.e. plays D with probability greater than 2/5) then if $x_0 = -1$ agent ω will 'quasi-mismatch' and play U with probability greater than $1 - \delta$. But agent $S(\omega)$ must play U with probability greater than 2/5 or less than 3/5 (or both). ∎

3.6. Nonexistence of ε-Equilibria.

THE MAIN THEOREM. *The game $\mathbf{B}$ has no Bayesian ε-equilibria.*

Proof. Although the following statement does not appear as an explicit theorem in Levy (2012), it is in effect what is proved by putting together the results Lemma 4.0.8, Lemma 4.0.9 and Proposition 4.0.10 of that paper:

Let F be a finite set and τ be a permutation of F. Denote $\Theta := X \times F$ and let S denote the measure-preserving operator defined by the cross product of the Bernoulli shift operator and τ over Θ. Then there cannot exist a decomposition of Θ into three disjoint measurable subsets Ξ_D, Ξ_U and Ξ_M satisfying

(1) $\Theta = \Xi_D \cup \Xi_U \cup \Xi_M$ up to a null set;
(2) Lemma 5;
(3) Lemma 6

But this is exactly the situation we have developed for $\mathbf{K}$. This is a contradiction. If $\mathbf{K}$ has no Nash ε-equilibrium then $\mathbf{B}$ has no Bayesian ε-equilibrium. ∎

3.7. Robustness to Perturbations.

For $\varepsilon > 0$, an ε-perturbation of a Bayesian game $\mathbf{B}$ is a Bayesian game $\mathbf{B}'$ over the same type space and action sets, with a set of payoff functions v_i^ω satisfying $\|v_i^\omega - u_i^\omega\|_\infty < \varepsilon$ for all $i \in I$.

Fixing a game $\mathbf{B}'$ that is an ε-perturbation of the ergodic game $\mathbf{B}$ defined above, it is straightforward to show that the inequalities that are essential for the proofs of Lemmas 5 and 6 continue to hold with respect to the agents game of $\mathbf{B}'$ (we need to restrict to ε because if we perturb a payoff from 0 to $-\varepsilon$ agent ω can still concentrate only on the actions of $S(\omega)$ and ignore the other states in calculating an ε-best reply). We therefore immediately have the following corollary.

COROLLARY 5. *For sufficiently small ε, an ε-perturbation of the game* B *has no Bayesian ε-equilibria.*

A similar result holds for sufficiently small perturbations of the posterior probabilities defining the types t_R and t_G.

4. End Remarks

An Harsányi ε-equilibrium of a Bayesian game with a common prior μ is a profile of mixed strategies $\Psi = (\Psi_i)_{i \in I}$ such that for each player i and any unilateral deviation strategy $\widehat{\Psi}_i$,

$$\int_\Omega u_i^\omega(\Psi(\omega))\, d\mu(\omega) \geq \int_\Omega u_i^\omega(\widehat{\Psi}_i(\omega), \Psi_{-i}(\omega))\, d\mu(\omega) - \varepsilon.$$

Simon (2003) shows that the existence of a measurable 0-Harsányi equilibrium implies the existence of a measurable 0-Bayesian equilibrium. However, this result is known not to hold for $\varepsilon > 0$. The example in this paper does not, therefore, imply that there is no Harsányi ε-equilibrium. We leave the open question of whether or not there are examples of games that have no measurable Harsányi ε-equilibria for future research.

Furthermore, in the example here there is incomplete information on both sides: if we define the colour and value of x_0 to comprise the state of nature at a state of the world ω then neither player knows the true state of nature. Because of this ignorance of the state of nature and the way the payoff matrices are defined in Table 1, neither player ever knows the true payoffs.

This contrasts with the example in Simon (2003), where there is incomplete information on one and a half sides (that is, one player always knows the true payoff relevant data but not always what the other players might know) and by construction players know their own payoffs at each state. It is presently unknown whether an example can be constructed of a game in which players know their payoffs at each state but the game has no Bayesian ε-equilibrium.

5. Appendix

Proof of Lemma 4. By construction, $t_i(\omega)(\Pi_i(\omega)) = 1$ for all ω and $t_i(\omega) = t_i(\omega')$ for $\omega' \in \Pi_i(\omega)$. Two more items need to be checked: that for each event A, $t_i(\omega)(A)$ is measurable and that $\mu(A) = \int_\Omega t_i(\omega)(A)\, d\mu(\omega)$. We will prove these for $i = R$, with the proof for $i = G$ conducted similarly.

For the rest of this proof, denote by $1_A(\omega)$ the indicator function that returns 1 if $\omega \in A$ and 0 if $\omega \notin A$. Fix an event A. Then:

$$t_R(\omega)(A) = \begin{cases} \frac{1}{4}1_A(\omega) + \frac{1}{4}1_A(i(\omega)) + \frac{1}{2}1_A(S(\omega)) & \text{if } \omega \in green \\ \frac{1}{4}1_A(S_+^{-1}(\omega)) + \frac{1}{4}1_A(S_-^{-1}(\omega)) + \frac{1}{2}1_A(\omega)) & \text{if } \omega \in red \end{cases}$$

from which we conclude that $t_R(\omega)(A)$ is measurable.

Next, we divide up A as follows:

$$\begin{aligned} A_1 :=& \{\omega \in A \mid \omega \in green \text{ and } \iota(\omega), S(\omega) \notin A\}, \\ A_2 :=& \{\omega \in A \mid \omega \in green \text{ and } \iota(\omega) \in A, S(\omega) \notin A\}, \\ A_3 :=& \{\omega \in A \mid \omega \in green \text{ and } \iota(\omega), S(\omega) \in A\}, \\ A_4 :=& \{\omega \in A \mid \omega \in red \text{ and } S^{-1}(\omega) \subset A^c\}, \\ A_5 :=& \{\omega \in A \mid \omega \in red \text{ and } S^{-1}(\omega) \cap A \neq \emptyset, S^{-1}(\omega) \not\subset A\} \\ & \cup \{\omega \in A \mid \omega \in green \text{ and } S(\omega) \in A, \iota(\omega) \notin A\}, \\ A_6 :=& \{\omega \in A \mid \omega \in red \text{ and } S^{-1}(\omega) \subset A\}. \end{aligned}$$

The sets (A_j) are all disjoint from each other. The proof proceeds by showing that $\mu(A_j) = \int t_R(\omega)(A_j)\, d\mu(\omega)$ for each $1 \leq j \leq 6$, which is straight-forward but tedious. We show how it is accomplished in two of the cases, trusting that the technique for the rest of the cases will be clear enough.

Case A_1. Define $B := A_1 \cup \iota(A_1) \cup S(A_1 \cup \iota(A_1))$. By the measure-preserving properties of ι and S, one has $\mu(\iota(A_1)) = \mu(A_1)$ and

$$\mu(S(A_1 \cup \iota(A_1))) = \mu(A_1 \cup \iota(A_1)),$$

hence $\mu(B) = 4\mu(A_1)$. On the other hand,

$$\int_\Omega t_R(\omega)(A_1)\, d\mu(\omega) = \int_B t_R(\omega)(A_1)\, d\mu(\omega) = \int_B \frac{1}{4}\, d\mu(\omega) = \mu(B)/4,$$

leading to the conclusion that $\int_\Omega t_R(\omega)(A_1)\, d\mu(\omega) = \mu(A_1)$.

Case A_5. Define $C := A_5 \cup (\iota(A_5 \cap green))$. Using similar reasoning as in the previous case, relying on measure-preserving properties, one deduces that $\mu(C) = \frac{4}{3}A_5$. On the other hand,

$$\int_\Omega t_R(\omega)(A_5)\, d\mu(\omega) = \int_C t_R(\omega)(A_5)\, d\mu(\omega) = \int_C \frac{3}{4}\, d\mu(\omega) = \frac{3}{4}\mu(C),$$

hence $\int_\Omega t_R(\omega)(A_5)\, d\mu(\omega) = \mu(A_5)$. ∎

Bibliography

Chatterjee, K., and W. Samuelson (1973), Bargaining Under Incomplete Information, *Operations Research*, 5, 835–851.

Halmos, P. (1956), *Lectures on Ergodic Theory*, New York, Chapman and Hall/CRC.

Harsányi, J. C. (1967), Games with Incomplete Information Played by Bayesian Players, *Management Science*, 14, 159–182.

Levy, Y., (2012), A Discounted Stochastic Game with No Stationary Nash Equilibrium, *Discussion Paper 596*, The Centre for the Study of Rationality, Hebrew University of Jerusalem.

Myerson, R. (1997), *Game Theory: Analysis of Conflict*, Cambridge, Harvard University Press.

Radner, R. (1980), Collusive Behavior in Noncooperative Epsilon-Equilibria of Oligopolies with Long but Finite Lives, *Journal of Economic Theory*, 22, 136–154.

Schmeidler, D. (1973), Equilibrium Points of Nonatomic Games, *Journal of Statistical Physics*, 7 (4), 295–300.

Simon, R. S., (2003), Games of Incomplete Information, Ergodic Theory and the Measurability of Equilibria, *Israel Journal of Mathematics*, (138), 73–92.

Sion, M., and P. Wolfe (1957), On a Game with No Value, *Annals of Mathematical Studies*, (39), 299–306.

פרק שלישי

ישנם שני אפיונים עיקריים של ה-CPA. האפיון המוכר יותר ("אפיון ההסכמה") מבוסס על הסכמה ביחס להתערבויות בין השחקנים (כבפרק הקודם). ב-1998 הציג פרופ' דב סמט אפיון המבוסס על שרשראות מרקוב, כאשר מטריצות המרקוב נגזרות מאמונות השחקנים.

פרק זה מרחיב את אפיונו של סמט למרחבים אינסופיים קומפקטים.

פרק רביעי

אפיון ההסכמה לא תקף במרחבים בני מניה. יוסי פיינברג, בעבודת הדוקטור שלו באוניברסיטה העברית בשנת 1996, הציג דוגמה של מרחב אמונות בר מניה בו קיום תנאי ההסכמה לא גורר קיום התפלגות הסתברות אפריורית משותפת.

ב-2006 הציע אביעד חפץ השערה שטענה שבמרחבים בני מניה אפיון ההסכמה מאפיין לא התפלגויות הסתברות אפריוריות משותפות אלא התפלגויות הסתברות משותפות שלא ניתנות לנירמול. בפרק זה אנו מראים שהשערה זו אינה נכונה: נכון הדבר שתנאי ההסכמה גורר קיום התפלגות הסתברות משותפת שלא בהכרח ניתנת לנירמול (הוכחה לכך מופיעה בפרק), אבל הכיוון השני לא תקף – יש דוגמה של מרחב אמונות בר מניה בו קיום התפלגות הסתברות משותפת שלא ניתנת לנירמול לא גורר את תנאי ההסכמה.

פרק חמישי

ב-2003 הציג רוברט סיימון דוגמה של משחק בייסאני עם שלושה שחקנים ואינסוף מצבי עולם שאין לו שווי-משקל בייסיאני. תוצאה זו הותירה שתי שאלות: א) האם יש דוגמה למשחק בייסיאני שאין לו שווי-משקל בייסיאני אפילו בקירוב? ב) האם יש דוגמה של משחק בייסיאני עם שני שחקנים ללא שווי-משקל בייסיאני?

התשובה היא חיובית לשתי השאלות יחד: בפרק זה אנו מציגים דוגמה של משחק בייסיאני עם שני שחקנים ללא שווי-משקל בייסיאני בקירוב (ולכן גם במדויק).

תקציר

הנחת התפלגות הסתברות אפריורית משותפת הינה הנחה רווחת מאוד בספרות המקצועית בתורת המשחקים. הנחה זו מקובלת, כמעט כמובנת מאליה, במודלים רבים החוקרים ציפיות רציונליות, סחר בניירות ערך, מיקוח, מכרזים, סיכון מוסרי, בעיות פשיטת רגל, משחקים חוזרים ומשחקים בייסיאנים – וזו רשימה חלקית בלבד. למרות השימוש הנרחב שנעשה בהנחה זו, היא שנויה במחלוקת ולא מעט חוקרים העלו סימני שאלה לגבי הצדקתה.

עבודה זו כוללת חמישה פרקים. כל פרק בוחן את הנחת התפלגות הסתברות אפריורית משותפת (CPA) ממבט שונה. להלן סיכום של כל אחד מהפרקים.

פרק ראשון

בפרק זה נשאלת השאלה מהי השכיחות של מרחבי טיפוסים עם התפלגות הסתברות אפריורית משותפת? הדבר נבחן ממבט טופולוגי. אם נקבע מרחב ידיעה וכנגד נשנה את אמונות השחקנים, האם נגלה שקבוצת המרחבים עם התפלגויות הסתברות אפריוריות משותפות היא קבוצה 'גדולה' או 'קטנה' במונחים טופולוגים?

התשובה תלויה במבנה החלוקות של השחקנים במרחב הידיעה. אם מבנה זה הוא 'הדוק' (כפי שבחרנו לכנות תנאי זה) אזי קיימת תמיד התפלגות הסתברות אפריורית משותפת, ללא תלות באמונות השחקנים. מאידך, בניגוד מוחלט, אם המבנה לא הדוק אזי קבוצת המרחבים עם התפלגויות הסתברות אפריוריות משותפות היא קבוצה דלילה טופולוגית; מרחב גנרי לא הדוק לא יקיים את הנחת ההתפלגות האפריורית המשותפת.

פרק שני

בפרק זה נשאלת השאלה מה קורה כשהתפלגויות ההסתברות האפריוריות *אינן* משותפות? האם במצב זה ניתן להסיק מסקנות לא טריוויאליות על התנהגות השחקנים?

התשובה היא כן. בהינתן מרחב אמונה ללא התפלגות הסתברות אפריורית משותפת, אנו מציגים מידה המודדת כמה רחוקים השחקנים ממצב בו יש להם התפלגות הסתברות אפריורית משותפת. את מרחק זה אנו מכנים "המרחק האפריורי". בכל מרחב אמונה עם מרחק אפריורי δ, לכל התערבות f, לא ייתכן שתהיה ידיעה משותפת שכל השחקנים מצפים לרווח חיובי של δ כפול sup-norm(f) או יותר. כאשר $\delta=0$, למרחב האמונות יש התפלגות הסתברות אפריורית משותפת, והתוצאה הנ"ל מתלכדת עם משפטי ההסכמה המוכרים בספרות.

בנוסף, ככל שמעדנים את חלוקות השחקנים הולך וקטן המרחק δ. בפרק אנו מחשבים חסם עליון למספר העידונים הנדרשים בכדי ש- $\delta=0$, המבטיח קיום התפלגות הסתברות אפריורית משותפת.

עבודה זו נעשתה בהדרכתם של:

פרופסור דב סמט

פרופסור סרג'יו הרט

התפלגויות הסתברות אפריוריות משותפות ולא משותפות

חיבור לשם קבלת תואר דוקטור לפילוסופיה

מאת

זיו הלמן

הוגש לסנט האוניברסיטה העברית בירושלים
ספטמבר 2012
אלול תשע״ב