ALMOST COMMON PRIORS

ZIV HELLMAN

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.

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 '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 (1967-8) 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

---

Ziv Hellman, Department of Mathematics and The Centre for the Study of Rationality, The Hebrew University of Jerusalem, *email:* ziv.hellman@mail.huji.ac.il.
The author acknowledges funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013/ERC grant agreement No. 249159).

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 δ, 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

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

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

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

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 $(\Omega, \mathbf{\Pi})$. In this context, Ω is a nonempty set called a *state space* (and each $\omega \in \Omega$ is called a state), 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. 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 $\mathbf{\Pi} = (\Pi_1, \dots, \Pi_{|I|})$, the total number of partition elements, i.e. $\sum_{i=1}^{|I|} |\Pi_i|$, is denoted $|\mathbf{\Pi}|$ (the *size* of $\mathbf{\Pi}$).

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 $\mathbf{\Pi}$, 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., $\tau = \{\Omega, \mathbf{\Pi}, (t_i)_{i\in I}\}$, where the t_i are defined relative to $\mathbf{\Pi}$. 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 $\mathbf{\Pi}'$ is a *(proper) refinement* of $\mathbf{\Pi}$ if for at least one player i, Π'_i is a (proper) refinement of Π_i. That $\mathbf{\Pi}'$ is a proper refinement of $\mathbf{\Pi}$ is denoted $\mathbf{\Pi} \prec \mathbf{\Pi}'$.

If τ is a type profile over Ω and $\mathbf{\Pi}$, then a refinement $\mathbf{\Pi}'$ of $\mathbf{\Pi}$ 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 $\mathbf{\Pi}'$ is a proper refinement of $\mathbf{\Pi}$. 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 $|\mathbf{\Pi}|$ of the partition profile $\mathbf{\Pi}$ over which τ is defined.

If $\mathbf{\Pi}'$ is a refinement of $\mathbf{\Pi}$, we also say that $\mathbf{\Pi}'$ is *finer* than $\mathbf{\Pi}$, and that $\mathbf{\Pi}$ is *coarser* than $\mathbf{\Pi}'$. The *meet* of $\mathbf{\Pi}$ is the finest common coarsening of the players' partitions. Each element of the meet of $\mathbf{\Pi}$ is called a *common knowledge component* of $\mathbf{\Pi}$. Denote by $C(\mathbf{\Pi})$ the number of common knowledge components in the meet of $\mathbf{\Pi}$.

A type profile $\mathbf{\Pi}$ 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]

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:

---

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

> 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 Section 3.1 to situations in which there is no common prior. For some motivating intuition, consider the following scenario.

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 naive 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 $\mathbb{R}^{mn}$.

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

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 the minimal L^1 distance between points in X and points in D, i.e. $\min_{x \in X, p \in D} \|x - p\|_1$. 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 straightfoward 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 $(f_1, \ldots, f_n)$ satisfies $\max_\omega f_i(\omega) - \min_\omega f_i(\omega) \leq 2$ 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 $\mathbb{R}^m$,

$$\min_{x_1 \in C_1, x_2 \in C_2} \|x_1 - x_2\|_p = \max_{\{f \in \mathbb{R}^m | \|f\|_q \leq 1\}} \inf_{x_1 \in C_1} f \cdot x_1 - \sup_{x_2 \in C_2} f \cdot x_2, \tag{4}$$

where $\|\cdot\|_p$ and $\|\cdot\|_q$ are dual norms, with the maximum in the right-hand side of Equation (4) attained by some $f \in \mathbb{R}^m$ such that $\|f\|_q \leq 1$. 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 $\max_\omega f_i(\omega) - \min_\omega f_i(\omega) \leq 2$ for each i. It follows that there is a real number c_i such that, setting $g_i := f_i + c_i e$, $\|g_i\|_\infty \leq 1$.

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 $f = (f_1, \dots, f_n) \in \mathbb{R}^{nm}$ 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 $x = (x_1, \dots, x_n)$ in X and $y = (p, \dots, p) \in D$.

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 $fx \geq b > d \geq fy$ and $b - d > n\delta$. Rearranging terms, we have $fx - d \geq b - d > n\delta$ and $fy - d < 0$.

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

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

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

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

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. 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 $\tau_0 \prec \tau_1 \prec \ldots \prec \tau_k$ 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 $\delta_0 \geq \delta_1 \geq \ldots \geq \delta_k$. 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 $\tau_0 \prec \tau_1 \prec \ldots$ be a sequence of successive proper refinements, with $\delta_0 \geq \delta_1 \geq \ldots$ 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 $\mathbf{\Pi}$, if

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

then any type profile over $\mathbf{\Pi}$ 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 $\tau_0 \prec \tau_1 \prec \ldots$ 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.

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

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.

REFERENCES

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

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

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

Dax, A. (2006), The distance between two convex sets, *Linear Algebra and its Applications*, 416, 184–213.

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

Geanakoplos, J. (1994), Common Knowledge, *Handbook of Game Theory* Volume 2, 1439-1496, Elsevier Science B.V., the Netherlands.

Harsányi, J. C. (1967-8), Games with Incomplete Information Played by Bayesian Players, *Management Science*, 14, 159–182, 320–334, 486–502.

Hellman, Z., and D. Samet (2012), How Common are Common Priors, *Games and Economic Behavior*, 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, *Economics and Philosophy*, 11, 227–253.

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

Samuelson, L. (2004), Modeling Knowledge in Economic Analysis, *Journal of Economic Literature*, 42, 367–403.