האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

DELUDEDLY AGREEING TO AGREE

By

ZIV HELLMAN

Discussion Paper # 605 April 2012

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

DELUDEDLY AGREEING TO AGREE

ZIV HELLMAN

Abstract. We study conditions relating to the impossibility of agreeing to disagree in models of interactive KD45 belief (in contrast to models of S5 knowledge, which are used in nearly all the agreements literature). Agreement and disagreement are studied under models of belief in three broad settings: non-probabilistic decision models, probabilistic belief revision of priors, and dynamic communication among players. We show that even when the truth axiom is not assumed it turns out that players will find it impossible to agree to disagree under fairly broad conditions.

1. Introduction

One of the strongest assumptions underpinning the standard model of knowledge, known as S5, is the *truth axiom*, which essentially states that 'everything that a player knows is true'. This is equivalent, from one perspective, to asserting that no mistakes are ever made in the processing of signals.

Mistakes, of course, abound around us, and sometimes such mistakes can have significant consequences. Consider, for example the following scenario[1] (a variation of an example appearing in Hart and Tauman (2004)): There are two traders. They trade on a daily basis, and since a trade involves one trader selling and the other buying, they can at least observe each others' willingness to trade. We may imagine that these two traders are the 'market leaders', in the sense that their actions are followed by others in the market and copied.

Let Ω be the set of all states of the world, with Ω containing nine states; $\Omega = \{1, 2, \ldots, 9\}$. For simplicity we will assume that there is a common prior p over Ω, with $p(\omega) = 1/9$ for all states ω. The private information of the two traders, Anne and Bob are summarized by partitions Π_A and Π_B respectively, with

$$\Pi_A = 1234|5678|9$$

and

$$\Pi_B = 123|456|789.$$

One standard interpretation of the structure of such partitional knowledge is that Anne and Bob receive signals. If the true state is 2, for example, Anne receives a signal that enables her to rule out the states $5, 6, 7, 8, 9$, and she therefore knows that the true state is one of $1, 2, 3, 4$. Bob, at the true state 3, receives a signal that enables him to rule out the states $4, 5, 6, 7, 8, 9$, and he therefore knows that the

The Department of Mathematics and the Centre for the Study of Rationality, The Hebrew University of Jerusalem, *email:* `ziv.hellman@mail.huji.ac.il`. This research was supported in part by the European Research Council under the European Commission's Seventh Framework Programme (FP7/20072013)/ERC grant agreement no. 249159.

[1] The scenario was suggested to the author by Uri Weiss.

true state is one of $1, 2, 3$. Specifically, suppose that Bob may receive any one of three signals, $\sigma_1, \sigma_2, \sigma_3$, where σ_1 informs Bob that the true state is one of $1, 2, 3$, σ_2 informs Bob that the true state is one of $4, 5, 6$, and σ_3 informs Bob that the true state is one of $7, 8, 9$ (we will be less interested in this example with specifying Anne's possible signals).

$$\begin{array}{ccc} \text{Signal} & & \text{States} \\ \sigma_1 & \rightarrow & \{1,2,3\} \\ \sigma_2 & \rightarrow & \{4,5,6\} \\ \sigma_3 & \rightarrow & \{7,8,9\} \end{array}$$

Figure 1: Bob's signals and their interpretation when there are no processing errors.

So far, so standard. Now consider the possibility of a mistake in signals processing on the part of Bob. Suppose that Bob inputs the signals he receives into a black box that he has been assured outputs $1, 2, 3$, $4, 5, 6$, or $7, 8, 9$ if the input is σ_1,σ_2, or σ_3 respectively. Unbeknownst to Bob (and to Anne), however, Bob's black box is defective; when either σ_1 or σ_2 are given as input, the box outputs $4, 5, 6$.

$$\begin{array}{ccc} \text{Signal} & & \text{States} \\ \sigma_1 & \rightarrow & \{4,5,6\} \\ \sigma_2 & \rightarrow & \{4,5,6\} \\ \sigma_3 & \rightarrow & \{7,8,9\} \end{array}$$

Figure 2: Bob's signal processing error.

Consider next the event $E = \{4, 9\}$. This event will be interpreted as a 'good' outcome (e.g., company earnings are about to rise), with the complement representing a 'bad' event that ought to trigger the sale of shares. Suppose that the true state is 2, and that each one of the two traders behaves each day according to the following rule:

$$\begin{cases} \text{Buy} & \text{if the probability of } E \text{ is } 0.3 \text{ or more;} \\ \text{Sell} & \text{if the probability of } E \text{ is less than } 0.3. \end{cases}$$

Given these assumptions, the following sequence of actions transpires. On Day 1, Anne, who processes signals correctly, supposes that the true state is one of $1, 2, 3, 4$, judges the probability of E to be $1/4$ and seeks to sell shares. Bob erroneously supposes that the true state is one of $4, 5, 6$, judges the probability of E to be $1/3$, and therefore buys shares from Anne.

Since Bob was willing to buy on Day 1, Anne 'learns' that the true state is not in $1, 2, 3$. She therefore erroneously supposes on Day 2 that the true state is 4 and offers to buy on Day 2. Bob does the same. By Day 3, it is 'common knowledge' that 4 is the 'true state' – Bob's error has now become Anne's error. Both traders seek to buy as many shares as they can, to their detriment, and a bubble has developed.

Geanakoplos (1989) and Morris (1996) show that in knowledge models that satisfy the truth axiom (but are not necessarily S5) more information is always beneficial for a player, in the sense that with more information a rational player will never choose an action that gives him less in expectation than an action that he chooses when he has less information. Without the truth axiom, that no longer holds true. Indeed, as the example here shows, without the truth axiom, not only is the 'mistaken' player in danger of choosing detrimental actions, his errors can cascade and 'infect' other players to their detriment: in Day 1 above, Anne makes the right decision in seeking to sell shares, but on Day 2, due to Bob's mistake, she is buying shares. Arguably, Anne has been mistaken all along, in accepting Bob's reports at face value, without considering the possibility that Bob might be mistaken.

The above story motivates the study of agreement and disagreement in models of *belief* as opposed to models of *knowledge*, which is the standard setting of most of the agreement literature. In this paper, we consider theorems relating to impossibility of agreeing to disagree in models of belief in three broad settings: in the non-probabilistic case, in which players make abstract decisions based on their beliefs (in the spirit of literature on the subject initiated by Bachrach (1985)); in the probabilistic case (as in most of the agreements literature) where agreement is defined in terms of the expectations of players revising prior probabilities by conditioning on partition refinements; and in the dynamic setting consisting of communications of reports made by players (in the spirit of literature on the subject initiated by Geanakoplos and Polemarchakis (1982)).

Considering belief as opposed to knowledge as arising from players making mistakes in belief revision from priors leads to concepts of *deluded belief revision* and the *deluded conditional expected value* of a random variable, which are introduced here. It is hoped that these may make a contribution to the literature on bounded rationality. Somewhat surprisingly, even when the truth axiom is not assumed it turns out players will find it impossible to agree to disagree under fairly broad conditions, as detailed in the claims in this paper. In fact, as Theorem 2 and the corollary and example following it show, situations in which there are states at which all players uniformly make mistakes lead to impossibility of agreeing to disagree results, although the players might agree to disagree if they did not make any mistakes at all!

The paper also includes a study of the structures several types of belief models, and introduces a new concept of belief revision from priors that is appropriate for belief models, which we term delusional revision.

1.1. **Review of Literature.** Battigalli and Bonanno (1999) and Bonanno and Nehring (1999) contain a wealth of ground-breaking results relating to belief models in interactive settings, as does Samet (2011). Many results of this paper build on ideas appearing in those papers.

Samet (2010b) and Samet (2008) are two papers devoted to studying conditions for the impossibility of agreeing to disagree in the non-probabilistic case in knowledge models and non-partitional models; a large debt is owned to ideas appearing in those papers in the non-probabilistic section of this paper. Tarbush (2011) extends some of D. Samet's results to KD45 models (in a mainly modal setting) in a spirit similar to that of this paper.

Tallon, Vergnaud and Zamir (2004) study novel concepts of revision in KD45 models when players directly communicate their beliefs to each other. In this paper, in contrast, each player communicates only his decisions based on his beliefs, leading to indirect assessments of their beliefs by the other players.

1.2. **Acknowledgements.** Thanks are due to Uri Weiss for suggesting the topic and to Ilan Nehama, Yehuda Levy and Dov Samet for helpful conversations.

2. Preliminaries

2.1. Belief Structures.

Fix a finite set of players I and a finite set of *states of the world*[2] denoted by Ω. Subsets of Ω are called *events*. The set of probability distributions over Ω is denoted by $\Delta(\Omega)$.

A function $B_i : 2^\Omega \to 2^\Omega$ associated with player i is called a *belief operator* for i. A *type function* t_i over Ω for player i is defined by assigning, for each ω, a probability distribution $t_i(\omega) \in \Delta(\Omega)$ representing player i's beliefs at ω. We associate with each type function t_i a partition Π_i of Ω defined[3] by $\Pi_i(\omega) = \{\omega' \mid t_i(\omega') = t_i(\omega)\}$. If we impose on a type function the property that $t_i(\omega)(\Pi_i(\omega)) = 1$, then the type functions is *partitional*. A *probabilistic belief structure* over Ω is then a set of partitional type functions $(t_i)_{i\in I}$ over Ω.

For each i, define the belief operator B_i^1 by ; $B_i^i E$ is termed the event that i is *certain* of, given E.[4]

The belief operator B_i^1 is a special case of a p-belief operator B_i^p, where $B_i^p E = \{\omega \mid t_i(\omega)(E) \geq p\}$ (see Monderer and Samet (1989)), and the full probabilistic belief structure defined above is appropriate for the general study of p-belief operators. Since we want to concentrate solely on 1-belief here, a simpler structure, termed a belief structure, will suffice for our purposes.

A function $b_i : \Omega \to 2^\Omega \setminus \emptyset$ is a *possibility function.* The event $b_i(\omega)$ is interpreted as the set of states that are considered possible for i at ω, while all other states are excluded by i at ω. We will call a possibility function $b_i : \Omega \to 2^\Omega \setminus \{\emptyset\}$ that is measurable with respect to a partition Π_i and satisfies $b_i(\omega) \subseteq \Pi_i(\omega)$ for each $\omega \in \Omega$ a *KD45 possibility function.*

The following four axioms on a belief operator $B_i : 2^\Omega \to 2^\Omega$ are standard in the literature:

(K) $B_i(\neg E \cup F) \cap B_i E \subseteq B_i F$
(D) $B_i E \subseteq \neg B_i \neg E$
(4) $B_i E \subseteq B_i B_i E$
(5) $\neg B_i E \subseteq B_i \neg B_i E$

Given possibility function b_i, define a belief operator $B_i : 2^\Omega \to 2^\Omega$ by

$$B_i E := \{\omega \mid b_i(\omega) \subseteq E\}. \tag{1}$$

Samet (2011) shows that a belief operator B_i satisfies the above axioms K,D,4 and 5 if and only if there exists a KD45 possibility function b_i such that the belief operator derivable from b_i is the operator B_i.[5]

---

[2] In the basic definitions of elements of belief structures we largely follow Samet (2011).

[3] The presentation here reverses most presentations of belief structures, in which partitions are given and used to define type functions; here we are starting with type functions and using them to define the partitions.

[4] The property that $t_i(\omega)(\Pi_i(\omega)) = 1$ at each state ω therefore states that player i is always certain of his type.

[5] Hence the name KD45.

A *belief structure* over Ω is a set of pairs $\mathbf{\Pi} = (\Pi_i, b_i)_{i \in I}$, where each b_i is a KD45 possibility function with respect to the partition Π_i of Ω. We will sometimes also call such a structure a KD45 belief structure. Since one may define a belief structure equivalently by specifying either the possibility functions or the belief operators as primitives, we will allow ourselves to specify either of them as is convenient.

A probabilistic belief structure $(t_i)_{i \in I}$ over Ω induces a belief structure $(\Pi_i, b_i)_{i \in I}$ over Ω, where Π_i is the partition of Ω into the types of player i and $b_i(\omega)$ is the set of states in $\Pi_i(\omega)$ that have positive $t_i(\omega)$ probability. Conversely, every belief structure over Ω is induced by a probabilistic belief structure over Ω. We will sometimes make use of this by choosing, for a given belief structure $\mathbf{\Pi} = (\Pi_i, b_i)_{i \in I}$, an arbitrary probabilistic belief structure $(t_i^b)_{i \in I}$ that induces $\mathbf{\Pi}$.

2.2. **Intrapersonal Belief Consistency.**

Belief is 'one axiom short of knowledge', in the following sense. An operator K_i on Ω is a *knowledge operator* derived from a partition of Ω if and only if it satisfies the KD45 axioms and additionally the truth axiom, defined as:

(T) $K_i E \subseteq E$.

A belief structure that satisfies the axioms KD45+T is called an S5 (or knowledge) structure or model.

The truth axiom states that knowledge is correct; if E is known it must true. This can equivalently be written as $\neg K_i E \cup E = \Omega$. This is where the distinction between knowledge and belief lies: for a belief operator, we do *not* assume that $B_i E \subseteq E$.

Equivalently, we do not assume that $\neg B_i E \cup E = \Omega$ always holds (i.e., it may not be true at all states). However, as shown in Samet (2011), a player always *believes* that it holds, that is,

Intrapersonal Belief Consistency: KD45 $\Rightarrow$ for each event E

$$B_i(\neg B_i E \cup E) = \Omega.$$

The equivalent concepts in the context of possibility functions are defined quite simply in terms of set containment. Let $\mathbf{\Pi} = (\Pi_i, b_i)_{i \in I}$ be a belief structure. If $\omega \in b_i(\omega)$ then b_i is *non-deluded* at ω. If $\omega \notin b_i(\omega)$ then b_i is *deluded* at ω; in this case we will also sometimes say that ω is a deluded state for player i.

If there is at least one state at which b_i is deluded, then b_i is *delusional*, and we will similarly say that the corresponding belief operator B_i is delusional if this is the case. It is straight-forward to show that a belief structure $\mathbf{\Pi}$ is non-delusional for all players if and only if it is an S5 structure, and it is similarly straight-forward to show that a state ω is non-deluded for player i if and only if $t_i^b(\omega) = 0$ for any probabilistic belief structure $(t_i^b)_{i \in I}$ that induces $\mathbf{\Pi}$.

2.3. **Partitional Aspect of KD45 Structures.**

It is well-known that in a belief structure $\mathbf{\Pi} = (\Pi_i, b_i)_{i \in I}$, the axioms of transitivity (i.e., for all i and all $\omega, \omega' \in \Omega$, if $\omega' \in b_i(\omega)$ then $b_i(\omega') \subseteq b_i(\omega)$) and euclideaness (i.e., for all i and all $\omega, \omega' \in \Omega$, if $\omega' \in b_i(\omega)$ then $b_i(\omega) \subseteq b_i(\omega')$). It follows that if $\omega' \in b_i(\omega)$ then $b_i(\omega') = b_i(\omega)$.

Denote

$$f_i(\omega) := \{\omega' \mid \omega' \notin b_i(\omega) \text{ and } b_i(\omega') = b_i(\omega)\}.$$

By definition, $b_i(\omega) \cap f_i(\omega) = \emptyset$. Furthermore, it is easy to show that $b_i(\omega) \cup f_i(\omega) = \Pi_i(\omega)$ for all $\omega \in \Omega$.

The general structure of a model of KD45 belief of a player i is therefore of an over-arching partition Π_i, with each partition element $\pi \in \Pi_i$ furthermore partitioned into $b_i(\omega)$ and $f_i(\omega)$ (using an arbitrary $\omega \in \pi$). Every element $\omega' \in f_i(\omega)$ is mapped by b_i into $bi(\omega)$, where it is 'trapped', in the sense that $b_i(b_i(\omega')) = b_i(\omega') = b_i(\omega)$.

In examples, we will compactly express KD45 belief structures by separating states in different partition elements of Π_i by the square boxes. Within each partition element we will denote states that are in the same component of $b_i(\omega)$ by an oval box.

For example, if we write

$\boxed{1\ 2}$ 3 4	5 $\boxed{6\ 7}$	8 $\boxed{9}$

then the intention is, for example, that 5, 6 and 7 are all in the same partition element, i.e., $\Pi_i(5) = \{5, 6, 7\}$, but 5 is a delusional state such that $b_i(5) = \{6, 7\}$.

2.4. **Standard belief revision and priors.**

Let μ be a probability distribution over Ω, and let Π_i be a partition of of Ω. The *(standard) revision* of μ at ω according to Π_i is the probability distribution $\widehat{\mu}(\omega)$ such that

$$\widehat{\mu}(\omega)(\omega') = \begin{cases} \frac{\mu(\omega')}{\mu(\Pi_i(\omega))} & \text{if } \omega' \in \Pi_i(\omega) \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

if $\mu(\Pi_i(\omega)) > 0$; otherwise it is undefined.

Let f be a random variable over Ω, μ be a probability distribution over Ω, and Π_i a partition of Ω. Then the *conditional expected value of f at ω* is

$$E_i^{\mu}(f \mid \Pi_i(\omega)) := \frac{1}{\mu(\Pi_i(\omega))} \sum_{\omega' \in \Pi_i(\omega)} f(\omega')\mu(\omega)(\omega'), \tag{3}$$

if $\mu(\Pi_i(\omega)) \neq 0$ (otherwise it is not defined).

Let $(t_i)_{i \in I}$ be a probabilistic belief structure over Ω, with $(\Pi_i)_{i \in I}$ the corresponding partition. A *(standard) prior* for t_i is a probability distribution $\mu \in \Delta(\Omega)$, such that $\widehat{\mu}(\omega) = t_i(\omega)$ at each ω, where $\widehat{\mu}(\omega)$ is the standard revision of μ at ω according to Π_i as defined in Equation (2). A *(standard) common prior* for $(t_i)_{i \in I}$ is a probability distribution $\mu \in \Delta(\Omega)$ that is a prior for each t_i.

Given a probabilistic belief structure $(t_i)_{i \in I}$ with corresponding partition $(P_i)_{i \in I}$, player i's *posterior expected value of f at ω* is

$$E_i^{t_i}(f \mid \Pi_i(\omega)) := \sum_{\omega' \in \Pi_i(\omega)} t_i(\omega')f(\omega'). \tag{4}$$

If there is a common prior μ, then for any random variable f the posterior expected value of each player equals the conditional expected value of f relative to μ and Π_i, i.e., $E_i^{t_i}(f \mid \Pi_i(\omega)) = E_i^{\mu}(f \mid \Pi_i(\omega))$.

3. Interpretations of KD45 Belief Structures

There are several ways of attaching interpretations to KD45 belief structures. One way is to consider a deluded state ω for player i simply as a state to which i attaches zero probability. This expresses the player's 'certainty' that ω cannot be the true state 'under any circumstances'; the player is therefore wrong if ω does occur, because there is no way to persuade him that ω occurs.

That interpretation is a static one. From a dynamic perspective, standard belief revision, as expressed in Equation (2), leaves no room for error or delusion; a player can ascribe zero probability to a state after revision if and only if he ascribes zero probability to that state prior to the revision. Since we wish to model situations in which mistakes are made in the revision process, we need to revisit revision. This is done in Section 3.1.

Another possible interpretation was hinted at in the introduction: misinterpreted signals. A formal development of that interpretation is in Section 3.2.

3.1. **Delusional belief revision.** Let μ be a probability distribution over Ω, and let b_i be a belief structure over Ω with corresponding partition Π_i. We introduce here the *delusional revision* of μ at ω according to b_i, defining it as the probability distribution $\widehat{\mu}(\omega)$ such that

$$\widehat{\mu}(\omega)(\omega') = \begin{cases} \frac{\mu(\omega')}{\mu(b_i(\omega))} & \text{if } \omega' \in b_i(\omega) \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

if $\mu(b_i(\omega)) > 0$; otherwise it is undefined.

Let f be a random variable over Ω, let μ be a probability distribution over Ω, and let b_i be a belief structure over Ω with corresponding partition Π_i. Then the *delusional conditional expected value of f at ω* according to b_i is

$$E_i^{\mu}(f \mid b_i(\omega)) := \frac{1}{\mu(b_i(\omega))} \sum_{\omega' \in b_i(\omega)} f(\omega')\mu(\omega)(\omega'), \tag{6}$$

if $\mu(\Pi_i(\omega)) \neq 0$ (otherwise it is not defined).

Comparing Equation (5) with Equation (2), and similarly Equation (6) with Equation (3), one sees that the distinction lies in the question of whether the states in $b_i(\omega)$ or $\Pi_i(\omega)$ are to be taken into account. If there are no deluded states at all for player i, the distinction disappears. However, if ω is a deluded state for player i, then we argue that Equation (5) is more appropriate, because at ω player i believes that the true state is in $b_i(\omega)$, hence should condition only on the probability $\mu(b_i(\omega))$, not on the larger set $\Pi_i(\omega)$.

Let $(t_i)_{i\in I}$ be a probabilistic belief structure over Ω, with $(\Pi_i)_{i\in I}$ the corresponding partition. Let b_i be the belief structure induced by t_i. A *delusional prior* for t_i is a probability distribution $\mu \in \Delta(\Omega)$, such that $\widehat{\mu}(\omega) = t_i(\omega)$ at each ω, where $\widehat{\mu}(\omega)$ is the delusional revision of μ at ω according to b_i as defined in Equation (5). A *common delusional prior* for $(t_i)_{i\in I}$ is a probability distribution $\mu \in \Delta(\Omega)$ that is a prior for each t_i.

Let ϕ_i be a standard prior for t_i, and suppose that for a state ω, $t_i(\omega)(\omega) = 0$, and therefore that $\omega \in \Pi_i(\omega)$ but $\omega \notin b_i(\omega)$. Then by Equation (2) it must be the case that $\phi_i(\omega) = 0$. The same reasoning does not hold for a delusional prior; a standard prior is a delusional prior, but the converse is not necessarily true.

Example 1. Consider a one-player probabilistic belief structure over a state space $\Omega = \{\omega_1, \omega_2, \omega_3\}$ defined by

$$t(\omega_k)(\omega_1) = 0;\ t(\omega_k)(\omega_2) = \frac{1}{2};\ t(\omega_k)(\omega_3) = \frac{1}{2}$$

for $k \in \{1, 2, 3\}$:

$$t = \boxed{\overbrace{\omega_1}^{0} \quad \overbrace{\omega_2}^{\frac{1}{2}} \quad \overbrace{\omega_3}^{\frac{1}{2}}}.$$

This induces a belief structure

$$b(\omega_1) = b(\omega_2) = b(\omega_3) = \{\omega_2, \omega_3\},$$

with ω_1 a deluded state, visualised as

$$\boxed{\omega_1 \quad \boxed{\omega_2 \quad \omega_3}}$$

The probability structure has only one (standard) prior, $\mu = (0, 1/2, 1/2)$, but it has an infinite number of delusional priors. The set of delusional priors includes, for example, $(0, 1/2, 1/2)$ and $(1/3, 1/3, 1/3)$. ♦

The fact that if ϕ_i is a standard prior for t_i, and $t_i(\omega)(\omega) = 0$ then $\phi(\omega) = 0$, has implications for common priors. It is immediate that if $t_i(\omega)(\omega) = 0$ and $t_j(\omega)(\omega) > 0$ for some pair of players i and j and any state ω, then the probabilistic belief structure $(t_i)_{i \in I}$ has no standard common prior; it may, however, have a common delusional prior. If such non-singularities are avoided, the following lemma holds.

Lemma 1. *Suppose that a probabilistic belief structure $(t_i)_{i \in I}$ satisfies the property that for each $\omega \in \Omega$ either $t_i(\omega)(\omega) = 0$ for all i or $t_i(\omega)(\omega) > 0$ for all i. Then $(t_i)_{i \in I}$ has a common delusional prior if and only if it has a common standard prior.*

We note here that there is no need to define a delusional version of the posterior expected value of a random variable f that is essentially different from the standard posterior expected value. One can define

$$E_i^{t_i}(f \mid b_i(\omega)) := \sum_{\omega' \in b_i(\omega)} t_i(\omega') f(\omega'). \tag{7}$$

But then

$$E_i^{t_i}(f \mid b_i(\omega)) = E_i^{t_i}(f \mid \Pi_i(\omega)) = \sum_{\omega' \in \Pi_i(\omega)} t_i(\omega') f(\omega'). \tag{8}$$

This holds because for any state ω', $t_i(\omega') \neq 0$ only if $\omega' \in b_i(\omega')$. If i is deluded at ω', the right-hand side of Equation (8) does not involve $f(\omega')$, and the value of $f(\omega')$ is therefore immaterial for calculating conditional expectation.

3.2. **KD45 from Misinterpreting Signals.**

For each player i, let S_i be a set of distinct *signals*. A *signalling function* σ_i is a function $\sigma_i : \Omega \to S_i$, which naturally induces a partition $\Pi_i^{\sigma_i}$ of Ω by setting

$$\Pi_i^{\sigma_i}(\omega) = \{\omega' \mid \sigma_i(\omega') = \sigma_i(\omega)\}. \tag{9}$$

It is straight-forward to show that any such partition $\Pi_i^{\sigma_i}$ derived from a signalling function is an S5 knowledge structure.

Next, let $M_i \subsetneq S_i$ be a set intended to represent the signals that i misreads. Let $\eta_i : M_i \to S_i \setminus M_i$ be an arbitrary function, and define $\chi_i : \Omega \to S_i$ by

$$\chi_i(\omega) = \begin{cases} \eta_i(\sigma_i(\omega)) & \text{if } \sigma_i(\omega) \in M_i \\ \sigma_i(\omega) & \text{otherwise} \end{cases} \tag{10}$$

This represents the action of misreading signals on the part of i. When ω is a state such that $\sigma_i(\omega) \in M_i \setminus S_i$, then $\chi_i(\omega) = \sigma_i(\omega)$, and the signal is read correctly by player i; however, if $\sigma_i(\omega) \in M_i$, then $\chi_i(\omega) = \eta(\sigma_i(\omega)) \in S_i \setminus M_i$, and i is misled into thinking that the signal differs from the true signal $s_i(\omega)$.

Let

$$\Theta_i^{\chi_i}(\omega) = \{\omega' \mid \chi_i(\omega') = \chi_i(\omega)\}, \tag{11}$$

and let

$$b_i^{\chi_i}(\omega) = \{\omega' \mid \omega' \in \Theta_i^{\chi_i}(\omega) \text{ and } \sigma_i(\omega') \in S_i \setminus M_i\}. \tag{12}$$

It is easy to check that $(b_i^{\chi_i})_{i \in I}$ constructed in such a way satisfies the KD45 axioms of a belief structure, with $(\Theta_i^{\chi_i})_{i \in I}$ the corresponding partition: for any pair $\omega, \omega' \in \Omega$, if $\omega' \in b_i(\omega)$ then $b_i(\omega') = b_i(\omega)$, and furthermore, the states in $\Pi_i^{\chi_i}(\omega) \setminus b_i^{\chi_i}(\omega)$ are all feed-in states to $b_i^{\chi_i}(\omega)$.

Furthermore, for each ω, $\Theta_i^{\chi_i}(\omega)$ (as defined in Equation (11)) is a union of elements of $\Pi_i^{\sigma_i}$ (as defined in Equation (9)):

$$\Theta_i^{\chi_i}(\omega) = \bigcup_{\{\omega' \in \Theta_i^{\chi_i}(\omega)\}} \Pi_i^{\sigma_i}(\omega');$$

letting $\hat{\omega}$ be any state $\hat{\omega} \in b_i^{\chi_i}(\omega)$, we also have

$$b_i^{\chi_i}(\omega) = \Pi_i^{\sigma_i}(\hat{\omega}). \tag{13}$$

4. Interactive Belief

Axiomatic structures of knowledge and belief, such as S5 and KD45, have been intensely studied for decades (especially in the context of modal logic). These are by construction single-agent models. In the literature on interactive decision theory, such models are usually extended to multi-player contexts simply by postulating that each player individually holds belief or knowledge operators that satisfy the appropriate axioms (as in the section above in this paper).

Interactive situations, however, open up scope for considering interactions between the belief operators of different players, enabling one to consider interesting new doxastic models.[6]

---

[6] Battigalli and Bonanno (1999) contains a survey of concepts related to interactive KD45 belief.

4.1. **Interpersonal Belief Consistency.**

Intrapersonal consistency expresses the property that every player always believes that his own beliefs are correct. A stronger condition is interpersonal consistency, in which each player believes that not only he, but all other players have correct beliefs. Formally, a belief structure satisfies *interpersonal belief consistency* if for each player i and j and event $E \subseteq \Omega$,

$$B_i(\neg B_j E \cup E) = \Omega. \tag{14}$$

In S5 knowledge models, interpersonal knowledge consistency always holds, i.e., $K_i(\neg K_j E \cup E) = \Omega$ is true for all i, j and E. Adding Equation (14) as an axiom to KD45 belief structures therefore results in models that are intermediate between pure KD45 structures and S5 structures.[7]

Another way of expressing the same concept is in terms of perceived worlds. The *world perceived by player* i is the minimal event F that satisfies $B_i(F) = \Omega$. Samet (2011) proves that the world perceived by a player always exists and is unique, and that players' beliefs are interpersonally consistent if and only if all players perceive the same world.

Denote by $D_i \subset \Omega$ the set of player i's deluded states, and by $ND_i = \Omega \setminus D_i$.

Lemma 2. *The world perceived by* i *is* ND_i.

We also introduce the following. Let $ND = \bigcap_{i \in I} ND_i$ be the set of states at which all players are uniformly non-deluded and $D = \bigcap_{i \in I} D_i$ be the set of states at which all players are uniformly deluded. A state $\omega \in \Omega$ is *non-singular* if $\omega \in ND \cup D$, otherwise it is *singular*. A singular state is one in which there exist players i and j such that i is deluded at ω and j is non-deluded at ω.

A belief structure is non-singular if all the states $\omega \in \Omega$ are non-singular with respect to it. One way to think of this is to consider a non-singular belief structure to be a structure such that at each state 'either all players are right, or all players are wrong (although they made be wrong in different ways, meaning that we allow the possibility that $b_i(\omega) \neq b_j(\omega)$ at a state ω at which the players are deluded)'.

4.2. **Convictions.**

For a knowledge operator K_i, the set of all events E that player i knows at ω is called player i's *ken* at ω; formally,

$$\mathrm{ken}_i(\omega) := \{E \mid \omega \in K_i(E)\}.$$

We introduce here concepts broadly analogous to those defined for kens in Samet (2008) for belief structures.

Given a belief operator B_i, call the set of all events E that player i believes at ω player i's *conviction*,[8] and denote it by

$$\mathrm{con}_i(\omega) := \{E \mid \omega \in B_i(E)\}.$$

---

[7] Bonanno and Nehring (1999) introduces a similar concept to interpersonal belief consistency that is called 'common belief of no error' in that paper.

[8] The word 'ken' is usually defined as 'range of knowledge' in dictionaries. The word 'conviction' may be defined as 'fixed or strong belief'.

Con_i will denote the family of all of i's convictions, i.e.,

$$\text{Con}_i := \{\text{con}_i(\omega) \mid \omega \in \Omega\}.$$

When we want to denote a particular conviction in Con_i without specifying the state ω, we will frequently use the notation $\mathbb{C}_i \in \text{Con}_i$. For each player i, $\text{con}_i(\omega)$ consists of all the supersets of $b_i(\omega)$, whether or not b_i is deluded at ω.

We have defined Con_i using the belief operator B_i, but we can also go in the other direction: given the family Con_i, one has that

$$B_i(E) = \{\omega \mid E \in \text{con}_i(\omega) \text{ and } \omega \in E\}. \tag{15}$$

We introduce here the following notation: let $\mathcal{C}_i \subseteq \text{Con}_i$ be a family of convictions of a player i. Then $\Omega\,|_{\mathcal{C}_i}$ will denote the set of all states in Ω such that i's convictions at ω are in $\mathcal{C}_i$, i.e.

$$\Omega\,|_{\mathcal{C}_i} := \{\omega \mid \text{con}_i(\omega) \in \mathcal{C}_i\} \tag{16}$$

From the definitions it easily follows that

$$\Omega\,|_{\text{Con}_i} = \Omega. \tag{17}$$

Furthermore, denote by

$$ND(\mathcal{C}_i) := \{\omega \in \Omega\,|_{\mathcal{C}_i} \mid \omega \text{ is not deluded for player } i\} \tag{18}$$

the set of non-deluded states whose convictions are in $\mathcal{C}_i$.

4.3. **Permission.**

For a given conviction of a player i, $\mathbb{C}_i$, and conviction of a player j, $\mathbb{C}_j$ we want a way to identify those convictions of another player j for which i's beliefs as represented by $\mathbb{C}_i$ are not contradicted by j's beliefs as represented by $\mathbb{C}_j$.

Formally, a conviction $\mathbb{C}_i \in \text{Con}_i$ *permits* a conviction $\mathbb{C}_j \in \text{Con}_j$ if for each $E \in \mathbb{C}_j$, $\neg B_j(E) \notin \mathbb{C}_i$, and for each $E \notin \mathbb{C}_j$, $B_j(E) \notin \mathbb{C}_i$. The set of all convictions in Con_j that are permitted by $\mathbb{C}_i$ is denoted by $\text{Permit}_j(\mathbb{C}_i)$.

Lemma 3. *Let $\mathbf{\Pi}$ be a $KD45$ belief structure. Then for any pair of players i and j and for each ω, for all $\omega' \in b_i(\omega)$*

$$con_j(\omega') \in Permit_j(con_i(\omega)),$$

4.4. **Interpersonal Belief Credibility.**

S5 knowledge structures, by dint of satisfying the truth axiom, satisfy the following property: $\bigcap_{i \in I} \text{ken}_i(\omega) \neq \emptyset$ for all states $\omega \in \Omega$. The corresponding property for belief structures,

$$\bigcap_{i \in I} \text{con}_i(\omega) \neq \emptyset \tag{19}$$

for all states ω, clearly does not hold in all belief structures. We can equivalently ask whether or not $\bigcap_{i \in I} b_i(\omega) \neq \emptyset$ for all $\omega \in \Omega$.

When Equation (19) does hold, we will say that the belief structure satisfies *interpersonal belief credibility*; the motivation for this name is that if at state ω player i reports to the other players that he believes $b_i(\omega)$, his report is 'credible' because all the other players believe that some state in $b_i(\omega)$ is true.

Since interpersonal credibility always holds in S5 knowledge structures, adding Equation (19) as an axiom to KD45 belief structures results in models that are intermediate between pure KD45 structures and S5 structures.

It is easy to construct examples of structures that satisfy interpersonal belief consistency but not interpersonal belief credibility, examples of structures that satisfy interpersonal belief credibility but not interpersonal belief consistency, and examples of structures that satisfy both properties but are not knowledge structures.

4.5. **Decisions.**

It is reasonable to expect that a rational player will base his or her decisions at a state of the world on the basis of her or her beliefs. This finds formal expression as follows.

Expanding on ideas in Samet (2008) and Samet (2010b), let D be a non-empty set of *decisions*, and let $(\Omega, (b_i)_{i \in I}$ be a belief structure. A *decision function* $\mathbf{d}_i$ for player i associates each of i's convictions (as determined by b_i) with a decision, i.e., $\mathbf{d}_i : \mathrm{Con}_i \to D$. A vector of decision functions $\mathbf{d} = (\mathbf{d}_i)_{i \in I}$ over is a *decision function profile*. With tolerable abuse of notation, we will write $\mathbf{d}_i(\omega)$ for $\mathbf{d}_i(\mathrm{con}_i(\omega))$. Similarly, when working with a possibility function b_i, since b_i defines a conviction at each ω, we will sometimes write $\mathbf{d}_i(b(\omega))$, with the intention clear from the context.

Denote by $[\mathbf{d}_i = d]$ the event that i's decision is d, i.e., $[\mathbf{d}_i = d] = \{\omega \mid \mathbf{d}_i(\omega) = d\}$.

5. KD45 Structures and Substructures

5.1. **Common Belief Operator.** Let $B(E) = \cap_i B_i(E)$ and let B^m denote the m-th power of the operator B. Define the *common belief operator* Q by $Q(E) := \cap_{m=1}^{\infty} B^m(E)$. When we wish to speak about common belief that exists solely between two players i and j, we will write $B_{ij} := B_i(E) \cap B_j(E)$, and $Q_{ij}(E) := \cap_{m=1}^{\infty} B_{ij}^m(E)$.

5.2. **Common Belief Set.** The operator Q can also be expressed in terms of possibility functions b_i. Let $\mathbf{\Pi} = \{\Omega, (b_i)_{i \in I}\}$ be a belief structure. Define a function $b : \Omega \to 2^{\Omega}$ by $b(\omega) := \bigcup_{i \in I} b_i(\omega)$. For any $m \geq 1$, let b^m be the composition of the function b repeated m times. Define, for each ω,

$$b^Q(\omega) := \bigcup_{m \geq 1} b^m(\omega) \tag{20}$$

Then for each event E,

$$Q(E) = \{\omega \mid b^Q(\omega) \subseteq E\}. \tag{21}$$

For any state ω, call $b^Q(\omega)$ the *common belief set* of ω in Ω. From the fact that KD45 belief structures satisfy transitivity, it follows easily that for any event E, if $\omega \in Q(E)$ then $\omega' \in Q(E)$ for all $\omega' \in b^Q(\omega)$. Using the same reasoning, it is easy to show that for any state ω,

$$b^Q(\omega) = \bigcup_{\{\omega' \in \bigcup_{i \in I} b_i(\omega)\}} b^Q(\omega'). \tag{22}$$

Denote, for each player i,

$$S^i(b^Q(\omega)) := \bigcup_{\omega'' \in b^Q(\omega)} b_i(\omega''), \tag{23}$$

and

$$T^i(b^Q(\omega)) := \bigcup_{\omega'' \in b^Q(\omega)} \Pi_i(\omega''), \tag{24}$$

and

$$\mathcal{C}_i(b^Q(\omega)) := \{\mathrm{con}_i(\omega') \mid \omega' \in \omega \cup b^Q(\omega)\}. \tag{25}$$

Lemma 4 is a technical result that will be needed later on. It states that for ascertaining the common belief convictions, one needn't confine attention solely to the states ω' in $b^Q(\omega)$; delusional states associated with states in $b^Q(\omega)$ are just as good for that purpose.

Lemma 5 describes the relationship between the information sink states in common belief sets, the common belief sets themselves, and associated delusional states; a similar result (in a modal setting) appears in Tarbush (2011).

Lemma 4. *Let* $\mathbf{\Pi} = (\Omega, (b_i)_{i \in I})$ *be a belief structure, with* (Π_i) *the corresponding partition profile. For each player* i *and every* $\omega \in \Omega$,

$$\mathcal{C}_i(b^Q(\omega)) = \{con_i(\omega'') \mid \omega'' \in T^i(b^Q(\omega))\}.$$

Lemma 5. *Let* $\mathbf{\Pi} = (\Omega, (b_i)_{i \in I})$ *be a belief structure, with* (Π_i) *the corresponding partition profile. For each player* i *and every* $\omega \in \Omega$,

$$S^i(b^Q(\omega)) \subseteq b^Q(\omega) \subseteq T^i(b^Q(\omega)).$$

6. Permission Consistency and Truth Equivalence

Knowledge structures are naturally partitioned into common knowledge components. Let $\{\Omega, (k_i)_{i \in I})\}$ be a knowledge structure. The *meet* is the finest common coarsening of the players' partitions. Each element of the meet of $\mathbf{\Pi}$ is called a *common knowledge component* of $\mathbf{\Pi}$.

Let $T \subseteq \Omega$ be a common knowledge component. T can be characterised in several ways. One way is by knowledge chains. Defining $k : \Omega \to 2^\Omega$ by $k(\omega) := \bigcup_{i \in I} k_i(\omega)$ and for $m \geq 0$ letting k^m be the composition of the function k repeated m times, it is well known that $T = \bigcup_{m \geq 1} k^m(\omega)$ for any $\omega \in T$.

Another property that T satisfies is that

$$\bigcup_{\omega \in T} k_i(\omega) = \bigcup_{\omega \in T} k_j(\omega) \tag{26}$$

for all $i, j \in I$.

Furthermore, let $T \subseteq \Omega$ and denote by $\mathcal{K}_i$ the set of all kens of player i at states in T, i.e., $\mathcal{K}_i := \{\mathrm{ken}_i(\omega) \mid \omega \in T\}$. Denote similarly $\mathcal{K}_j := \{\mathrm{ken}_j(\omega) \mid \omega \in T\}$. We will say that $\mathcal{K}_i$ and $\mathcal{K}_i$ are permission consistent if

$$\text{for each } \mathbb{K}_i \in \mathcal{K}_i, \mathrm{Permit}_j(\mathbb{K}_i) \subseteq \mathcal{K}_j,$$

and

$$\text{for each } \mathbb{K}_j \in \mathcal{K}_j, \mathrm{Permit}_i(\mathbb{K}_j) \subseteq \mathcal{K}_i.$$

Samet (2008) shows that in knowledge structures a pair of families of kens $\mathcal{K}_i$ and $\mathcal{K}_j$ of players i and j respectively, as defined above for $T \subseteq \Omega$, is permission consistent if and only if T is a common knowledge component, if and only if

$$\bigcap_{\mathbb{K}_i \in \mathcal{K}_i} \mathbb{K}_i = \bigcap_{\mathbb{K}_j \in \mathcal{K}_j} \mathbb{K}_j, \tag{27}$$

if and only if $\bigcup_{\omega \in T} k_i(\omega) = \bigcup_{\omega \in T} k_j(\omega)$.

6.1. **Permission Consistency.**

Definition 1. In a belief structure, a pair of families of convictions $\mathcal{C}_i \subseteq \text{Con}_i$ and $\mathcal{C}_j \subseteq \text{Con}_j$ of players i and j respectively is *permission consistent* if

$$\text{for each } \mathbb{C}_i \in \mathcal{C}_i, \text{Permit}_j(\mathbb{C}_i) \subseteq \mathcal{C}_j,$$

and

$$\text{for each } \mathbb{C}_j \in \mathcal{C}_j, \text{Permit}_i(\mathbb{C}_j) \subseteq \mathcal{C}_i.$$

By extension, a set of families of convictions $\mathcal{C} = (\mathcal{C}_i)_{i \in I}$ is permission consistent if the elements of the set are all pair-wise permission consistent. ♦

Proposition 1 shows that permission consistency roughly generalises the construction of common knowledge components by way of knowledge chains to belief spaces .

Proposition 1. *Recall the definition in Equation (25), of* $\mathcal{C}_i(b^Q(\omega)) := \{con_i(\omega') \mid \omega' \in \omega \cup b^Q(\omega)\}$.

(a) For any $\omega \in \Omega$ *and players* i *and* j*, the pair of families* $\mathcal{C}_i(b^Q(\omega))$ *and* $\mathcal{C}_j(b^Q(\omega))$ *are permission consistent.*

(b) Let $(\mathcal{K}_i)_{i \in I}$ *be a set of permission consistent families. Then for any state* ω *such that* $con_i(\omega) \in \mathcal{K}_i$ *for some player* i, $\mathcal{C}_j(b^Q(\omega)) \subseteq \mathcal{K}_j$ *for all* $j \in I$.

6.2. **Truth Equivalence.**

Definition 2. In a belief structure, a pair of families of convictions $\mathcal{C}_i \subseteq \text{Con}_i$ and $\mathcal{C}_j \subseteq \text{Con}_j$ of players i and j respectively is *truth equivalent* if

$$\bigcap_{\mathbb{C}_i \in \mathcal{C}_i} \mathbb{C}_i = \bigcap_{\mathbb{C}_j \in \mathcal{C}_j} \mathbb{C}_j.$$

By extension, a set of families of convictions $\mathcal{C} = (\mathcal{C}_i)_{i \in I}$ is truth equivalent if the elements of the set are all pair-wise truth equivalent. ♦

We show in Proposition 4 that permission consistency implies truth equivalence.

The definition of truth equivalence is inspired by the property expressed in Equation (27) for knowledge structures. The reason the name 'truth equivalence' was chosen for this concept in belief structures will be made clearer in the next series of results. In particular, Corollary 1 shows that truth equivalence implies that $\bigcup_{\omega \in T} b_i(\omega) = \bigcup_{\omega \in T} b_j(\omega)$ holds for some $T \subseteq \Omega$; compare this to Equation (26) above in knowledge structures.

Lemma 6. *Let* $\mathcal{C}_i \subseteq Con_i$ *be a family of convictions of player* i*. Then*

$$\bigcap_{\mathbb{C}_i \in \mathcal{C}_i} \mathbb{C}_i = \{E \mid \bigcup_{\omega \in \Omega|_{\mathcal{C}_i}} b_i(\omega) \subseteq E\}, \tag{28}$$

(where $\Omega\,|_{\mathcal{C}_i}$ is defined by Equation (16)).

Corollary 1. *Let $\mathcal{C}_i \subseteq Con_i$ and $\mathcal{C}_j \subseteq Con_j$ be a pair of truth equivalent families of convictions. Then*

$$\bigcup_{\omega \in \Omega|_{\mathcal{C}_i}} b_i(\omega) = \bigcup_{\omega' \in \Omega|_{\mathcal{C}_j}} b_j(\omega') \tag{29}$$

(where $\Omega\,|_{\mathcal{C}_i}$ and $\Omega\,|_{\mathcal{C}_j}$ are defined by Equation (16)).

If $\mathcal{C} = (\mathcal{C}_i)_{i\in I}$ is a set of truth equivalent families, then by the result of Corollary 1, $ND(\mathcal{C}_i) = ND(\mathcal{C}_j)$ for all $i, j \in I$ (where the operator ND is defined by Equation (18)). We can therefore define

$$ND(\mathcal{C}) := ND(\mathcal{C}_i) \tag{30}$$

for any $i \in I$.

Definition 3. Let $\mathcal{C} = (\mathcal{C}_i)_{i\in I}$ be a set of truth equivalent families. There is *strong common belief in truth* at a state ω if $b_i(\omega) \subseteq ND(\mathcal{C})$ for all $i \in I$.

We can now formalise the way that truth equivalence yields a form of 'local knowledge', in the sense that the set of non-delusional states $ND(\mathcal{C})$ corresponding to truth equivalent families of convictions 'looks like' a common knowledge component in a knowledge structure. Define the belief structure $\mathbf{\Pi}^{\mathcal{C}} = \{\Omega, b_1^{\mathcal{C}}, \ldots, b_n^{\mathcal{C}}\}$ to be the belief structure whose space of states is $ND(\mathcal{C})$, and whose possibility function $b_i^{\mathcal{C}}$, for each player i, is defined by $b_i^{\mathcal{C}}(\omega) = b_i(\omega)$. Then:

Proposition 2. *The belief structure $\mathbf{\Pi}^{\mathcal{C}}$ is a knowledge structure, i.e., it satisfies the S5 axioms.*

Lemma 7. *Let $\mathcal{C} = (\mathcal{C}_i)_{i\in I}$ be a set of truth equivalent families, and let ω be a state at which there is strong common belief of truth (i.e., $b_i(\omega) \in ND(\mathcal{C})$ for all $i \in I$). Then the common belief set $b^Q(\omega)$ satisfies $b^Q(\omega) \subseteq ND(\mathcal{C})$. In particular, if $\omega \in ND(\mathcal{C})$ then $b^Q(\omega) \subseteq ND(\mathcal{C})$.*

Lemma 8. *There is strong common belief in truth at a state ω if and only if there exists $\Omega_0 \subseteq \Omega$ such that $b^Q(\omega) = \bigcup_{\omega' \in \Omega_0} b_i(\omega')$ for all $i \in I$.*

Lemma 7 and Proposition 2 justify the use of the term 'common belief in truth': at a state in which there is strong common belief in truth, all the players believe that all the states of common belief are effectively elements of a knowledge structure in which the truth axiom holds.

Belief structures in which the set of all of player i's convictions is truth equivalent to the set of all the convictions of any other player j are precisely the non-singular structures:

Proposition 3. *For a belief structure $\mathbf{\Pi} = \{\Omega, (b_i)_{i\in I}\}$, the following are equivalent:*

1. *All players perceive the same world.*
2. *$\mathbf{\Pi}$ is non-singular (hence it satisfies interpersonal belief consistency).*
3. *For all players i and j, $\bigcap_{\mathbb{C}_i \in Con_i} \mathbb{C}_i = \bigcap_{\mathbb{C}_j \in Con_j} \mathbb{C}_j$.*
4. *For all players i and j, $\bigcup_{\omega\in\Omega} b_i(\omega) = \bigcup_{\omega\in\Omega} b_j(\omega)$.*
5. *There is strong common belief in truth at every state.*

We add here one more result: suppose that $\mathbf{\Pi} = \{\Omega, (b_i)_{i \in I}\}$ is a non-singular belief structure. For each ω and $i \in I$, define $ND^i(\omega)$ to be the smallest set of non-deluded states in Ω such that (a) $b_i(\omega) \subseteq ND^i(\omega)$ and (b) $\bigcup_{\omega' \in ND^i(\omega)} b_i(\omega') = \bigcup_{\omega' \in ND^i(\omega)} b_j(\omega')$, for all $j \in I$.

Lemma 9. *Let $\mathbf{\Pi} = \{\Omega, (b_i)_{i \in I}\}$ satisfy interpersonal belief consistency.*

(1) *If ω is a non-deluded state for all the players, then $ND^i(\omega) = ND^j(\omega)$ for all $i, j \in I$.*
(2) *For all $\omega \in \Omega$, the common belief set of ω, $b^Q(\omega)$, satisfies*

$$b^Q(\omega) = \bigcup_{i \in I} ND^i(\omega).$$

7. Agreement in Belief Structures: The Non-Probabilistic Case

7.1. Independence of Irrelevant Belief and Permission Consistent Consensus.

Definition 4 (**Independence of Irrelevant Belief**). A decision function profile $\mathbf{d} = (\mathbf{d}_i)_{i \in I}$ satisfies *independence of irrelevant belief* (IIB) if, for any pair of players $i, j \in I$, decisions d_i and d_j, and conviction families $\mathcal{C}_i \subseteq \mathrm{Con}_i$ and $\mathcal{C}_j \subseteq \mathrm{Con}_j$ if

1. $\mathcal{C}_i$ and $\mathcal{C}_j$ are truth equivalent and
2. for each $\mathbb{C}_i \in \mathcal{C}_i$, $\mathbf{d}_i(\mathbb{C}_i) = d_i$, and for each $\mathbb{C}_j \in \mathcal{C}_j$, $\mathbf{d}_j(\mathbb{C}_j) = d_j$,

then $d_i = d_j$. ♦

Definition 5 (**Permission Consistent Consensus**). A decision function profile $\mathbf{d} = (\mathbf{d}_i)_{i \in I}$ satisfies *permission consistent consensus* (PCC) if, for every pair of players $i, j \in I$, decisions d_i and d_j, and conviction families $\mathcal{C}_i \subseteq \mathrm{Con}_i$ and $\mathcal{C}_j \subseteq \mathrm{Con}_j$ if

1. $\mathcal{C}_i$ and $\mathcal{C}_j$ are permission consistent and
2. for each $\mathbb{C}_i \in \mathcal{C}_i$, $\mathbf{d}_i(\mathbb{C}_i) = d_i$, and for each $\mathbb{C}_j \in \mathcal{C}_j$, $\mathbf{d}_j(\mathbb{C}_j) = d_j$,

then $d_i = d_j$. ♦

The next definition captures the idea of impossibility of agreeing to disagree in our context: two players cannot agree to disagree on decisions if the set of states at which they have common belief that they choose different decisions is empty.

Definition 6 (**Impossibility of agreeing to disagree**). A decision function profile $\mathbf{d} = (\mathbf{d}_i)_{i \in I}$ satisfies impossibility of agreeing to disagree (IAD) if for all players i and j and decisions $d_i \neq d_j$,

$$Q_{ij}([\mathbf{d}_i = d_i] \cap [\mathbf{d}_j = d_j]) = \emptyset.$$

♦

Proposition 4. *Let $\mathbf{\Pi}$ be a belief structure, and let $\mathcal{C}_i \subseteq Con_i$ and $\mathcal{C}_j \subseteq Con_j$ be truth equivalent. Then $\mathcal{C}_i$ and $\mathcal{C}_j$ are permission consistent.*

Proposition 4 leads immediately to Proposition 5 as a corollary:

Proposition 5. *In belief structures, PCC implies IIB.*

We show by example that the converse does not hold.

Example 2. We construct a two-player belief structure. The state space is

$$\Omega = \{y_0, y_1, y_2, y_3\} \cup \{z_{-100}, \ldots, z_{-2}, z_{-1}, z_0, z_1, z_2, \ldots, z_{100}\}.$$

The possibility function b_i of player i is:

$$\begin{aligned} &b_i(y_0) = z_1 \\ &b_i(y_1) = y_1 \\ &b_i(y_2) = b_i(y_3) = \{y_2, y_3\} \end{aligned}$$

$$b_i(z_k) = \begin{cases} z_{k+1} & \text{if } k \in \{0, 2, 4, \ldots, 98\} \\ z_k & \text{if } k \in \{1, 3, 5, \ldots, 99\} \cup \{100\} \\ z_{k-1} & \text{if } k \in \{-1, -3, -5, \ldots, -99\} \\ z_k & \text{if } k \in \{-2, -4, -6, \ldots, -100\} \end{cases}$$

A visualisation of player i's beliefs appears in Figure 3.

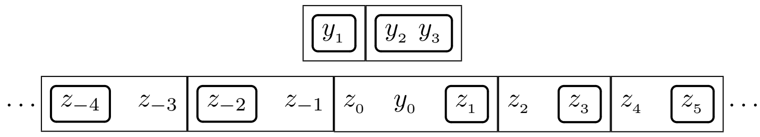

Figure 3: Player j's beliefs in Example 2.

The possibility function b_j of player j is:

$$\begin{aligned} &b_j(y_0) = b_j(y_1) = b_j(y_2) = \{y_1, y_2\} \\ &b_j(y_3) = y_3 \end{aligned}$$

$$b_j(z_k) = \begin{cases} z_{k+1} & k \in \{1, 3, 5, \ldots, 99\} \\ z_k & k \in \{2, 4, 6, \ldots, 100\} \\ z_{k-1} & k \in \{0, -2, -4, \ldots, -98\} \\ z_k & k \in \{-1, -3, -5, \ldots, -99\} \cup \{-100\} \end{cases}$$

A visualisation of player j's beliefs appears in Figure 4.

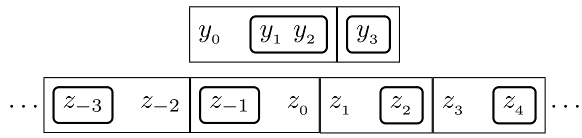

Figure 4: Player i's beliefs in Example 2.

Let $\mathcal{C}_i = \{\mathrm{con}_i(z_k)\}$ and $\mathcal{C}_j = \{\mathrm{con}_j(z_k)\}$ be families of convictions. An application of Proposition 1 shows that $\mathcal{C}_i$ and $\mathcal{C}_j$ are permission consistent. However, $\bigcap_{\mathbb{C}_i \in \mathcal{C}_i} \mathbb{C}_i \neq \bigcap_{\mathbb{C}_j \in \mathcal{C}_j} \mathbb{C}_j$, hence they are not truth equivalent.

In contrast, the conviction families $\mathcal{H}_i = \{\mathrm{con}_i(y_k) \mid k \in \{1, 2, 3\}\}$ and $\mathcal{H}_j = \{\mathrm{con}_j(y_k) \mid k \in \{1, 2, 3\}\}$ do satisfy truth equivalence.

Let the set of decisions be $D = \{d_1, d_2, d_3\}$. Defining decision functions

$$\mathbf{d}_i(\omega) = \begin{cases} d_1 & \omega \in y_0 \cup \{z_k \mid k \in \{-100, \ldots, 100\}\} \\ d_3 & \omega \in \{y_1, y_2, y_3\} \end{cases}$$

and

$$\mathbf{d}_j(\omega) = \begin{cases} d_2 & \omega \in \{z_k \mid k \in \{-100, \ldots, 100\}\} \\ d_3 & \omega \in \{y_0, y_1, y_2, y_3\} \end{cases}$$

with $d_1 \neq d_2$, shows that a belief structure can satisfy IIB without satisfying PCC. ♦

7.2. **Impossibility of Agreeing to Disagree.**

Theorem 1. *Let* $\mathbf{d}$ *be a decision function profile on a belief structure. Then* $\mathbf{d}$ *satisfies PCC if and only if it satisfies impossibility of agreeing to disagree.*

In contrast, IIB is not sufficient to guarantee the impossibility of agreeing to disagree in KD45: in the model in Example 2, $b^Q(z_0) = \{z_k \mid k \in \{-100, \ldots, -1\} \cup \{1, \ldots 100\}\}$, hence

$$Q_{ij}([\mathbf{d}_i = d_1] \cap [\mathbf{d}_j = d_2]) \neq \emptyset$$

(with $d_1 \neq d_2$), despite IIB being satisfied.

8. Agreement in Belief Structures: The Probabilistic Case

In the previous section, we showed that satisfying PCC is equivalent to the impossibility of agreeing to disagree. However, as we will show in this section, when a probabilistic belief structure is derived from a common delusional prior, PCC is not guaranteed. We can, never the less, find a condition that ensures that an analogue of the standard No Betting theorem from the literature on knowledge structures holds in belief structures.

8.1. **Standard No Betting.**

Definition 7. An n-tuple of random variables $\{f_1, \ldots, f_n\}$ is a *bet* if $\sum_{i=1}^n f_i = 0$. ♦

Definition 8. Let $(t_i)_{i \in I}$ be a probabilistic belief structure. Then a bet is an *agreeable bet* at ω (relative to (t_i)) if $E_i^{t_i}(f \mid \Pi_i(\omega)) > 0$ for all $i \in I$. A bet f is a *common knowledge agreeable bet at* ω if it is common knowledge at ω that f is an agreeable bet. ♦

The main characterisation of the existence of common priors in $S5$ knowledge models in the literature is what is sometimes known as the No Betting Theorem: a finite type space has a common prior if and only if there does not exist a common knowledge agreeable bet at any ω. In the special case of a two-player probabilistic belief structure where the random variable is the characteristic function

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

where H is an event, this characterisation implies the seminal Aumann Agreement Theorem (Aumann (1976)), which states that if it is common knowledge at a state of the world that player 1 ascribes probability η_1 to event H and player 2 ascribes probability η_2 to the same event, then $\eta_1 = \eta_2$.

In S5 knowledge structures, independence of irrelevant beliefs becomes symmetric independence of irrelevant knowledge, and Samet (2008) proves that symmetric independence of irrelevant knowledge is equivalent to the impossibility of agreeing to disagree. This immediately implies the Aumann Agreement Theorem. Let

$(t_i)_{i \in I}$ be a probabilistic belief structure, and let 1^H be the characteristic function of an event $H \subseteq \Omega$. For each player i, let the set of decisions be $D_i \mathbb{R}_0^+$, the set of non-negative real numbers, and let $\mathbf{d}_i(\omega) = E_i^{t_i}(1^H \mid \Pi_i(\omega))$. Then from the perspective of this paper, the proof of the main theorem in Aumann (1976) can be interpreted essentially as showing that when there is a common prior this set of decisions $(\mathbf{d})_i$ satisfies independence of irrelevant belief, and therefore impossibility of agreeing to disagree.

8.2. **KD45 No Betting.**

Definition 9. Let $(t_i)_{i \in I}$ be a probabilistic belief structure and $(b_i)_{i \in I}$ a belief structure induced by $(t_i)_{i \in I}$. A bet f is a *common belief agreeable bet at* ω if it is common belief at ω that f is an agreeable bet. ♦

With these definitions, we can now ask whether an analogue to the No Betting Theorem of S5 models holds in the KD45 setting. Given a probabilistic belief structure $(t_i)_{i \in I}$, does the existence of a common delusional prior imply that there is no common belief agreeable bet?

The answer to this question is no, as the following example[9] shows.

Example 3. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Consider the two-player probabilistic belief structure (t_1, t_2) defined by

$$t_1(\omega_k)(\omega_1) = \frac{1}{3}; t_1(\omega_k)(\omega_2) = \frac{1}{3}; t_1(\omega_k)(\omega_3) = \frac{1}{3},$$

and

$$t_2(\omega_k)(\omega_1) = 0; t_2(\omega_k)(\omega_2) = \frac{1}{2}; t_2(\omega_k)(\omega_3) = \frac{1}{2}$$

for $k \in \{1, 2, 3\}$:

$$t_1 = \boxed{\overbrace{\omega_1}^{\frac{1}{3}} \quad \overbrace{\omega_2}^{\frac{1}{3}} \quad \overbrace{\omega_3}^{\frac{1}{3}}},$$

$$t_2 = \boxed{\overbrace{\omega_1}^{0} \quad \overbrace{\omega_2}^{\frac{1}{2}} \quad \overbrace{\omega_3}^{\frac{1}{2}}}.$$

This induces the belief structure (b_1, b_2)

$$b_1(\omega_1) = b_1(\omega_2) = b_1(\omega_3) = \{\omega_1, \omega_2, \omega_3\},$$

and

$$b_2(\omega_1) = b_2(\omega_2) = b_2(\omega_3) = \{\omega_2, \omega_3\},$$

visualised as

$$\boxed{\boxed{\omega_1 \quad \omega_2 \quad \omega_3}}$$

$$\boxed{\omega_1 \quad \boxed{\omega_2 \quad \omega_3}}$$

For this belief structure, $\mu = (1/3, 1/3, 1/3)$ is a common delusional prior. Let $H = \{\omega_1, \omega_2\}$. Then it is common belief at every state ω that $E_1^{t_1}(1^H \mid b_1(\omega)) = 2/3$, while $E_2^{t_2}(1^H \mid b_2(\omega)) = 1/2$.

Note that this belief structure satisfies (vacuously) IIB, but not PCC. ♦

---

[9] This example is inspired by an example in Collins (1997).

To recapitulate something resembling the No Betting Theorem in belief structures, we add a new definition.

Definition 10. There is *weak common belief in truth*[10] at a state ω if there exists a state $\omega' \in b^Q(\omega)$ at which there is strong common belief in truth. ⧫

An equivalent way of stating the content of Definition 10 is as follows: there is weak common belief in truth at ω iff there exists a state $\omega' \in b^Q(\omega)$ such that

$$\bigcup_{\omega'' \in b^Q(\omega')} b_i(\omega'') = \bigcup_{\omega'' \in b^Q(\omega')} b_j(\omega'')$$

for all $i, j \in I$. This can be read intuitively as the players 'eventually' getting to strong common belief in truth as they follow chains in the common belief set.

A belief structure version of the No Betting Theorem can be attained if we assume weak common belief in truth.

Theorem 2. *Let $(t_i)_{i\in I}$ be a probabilistic belief structure over Ω and let ω be a state at which there is weak common belief in truth. Then there is a common delusional prior if and only if there is no common belief agreeable bet at ω.*

Since strong common belief in truth implies weak common belief in truth, and in a non-singular probabilistic belief structure there is strong common belief in truth at every state, Theorem 3 (which is close in content to a result appearing in Bonanno and Nehring (1999)) follows from Theorem 2 as a corollary.

Theorem 3. *Let $(t_i)_{i\in I}$ be a non-singular probabilistic belief structure over Ω. Then there is a common delusional prior if and only if there is no common belief agreeable bet at any state $\omega \in \Omega$.*

Example 4. The state space consists of $\{0, 1, 2, 3, 4, 5, 6, 7\}$. There are two players, i and j. The belief structure $((\Pi_i, b_i), (\Pi_j, b_j))$ is as follows:

Player i's beliefs are

$$\boxed{\boxed{1}}\boxed{\boxed{2}}\boxed{3\quad 4\quad \boxed{5}}\boxed{\boxed{6\quad 7}}$$

Player j's beliefs are

$$\boxed{\boxed{1\quad 2}\quad 3\quad 4}\boxed{\boxed{5\quad 6\quad 7}}$$

The states 3 and 4 are delusional states for both player i and player j, hence they perceive the same world. Note also that $b_i(3) = \{5\}$ while $b_j(3) = \{1, 2\}$, and this structure therefore does not satisfy interpersonal belief credibility. In fact, the structure can naturally be divided into two 'certainty components', $\{1, 2\}$ and $\{5, 6, 7\}$; at states 3 and 4, player i is certain that the true component is $\{5, 6, 7\}$ while player j is certain that the true component is $\{1, 2\}$.

The above belief structure can be induced by the following non-singular probabilistic belief structure (t_i, t_j):

$$t_i = \boxed{\overbrace{1}^{1}}\boxed{\overbrace{2}^{1}}\boxed{\overbrace{3}^{0}\quad \overbrace{4}^{0}\quad \overbrace{5}^{1}}\boxed{\overbrace{6}^{1/2}\quad \overbrace{7}^{1/2}}$$

---

[10] Although weak common belief in truth may seem abstract at first reading, it arises naturally in the study of interactive belief models. Concepts very similar to that of weak common belief in truth are introduced and used in Battigalli and Bonanno (1999) and Tarbush (2011).

$$t_j = \begin{array}{|cccc|ccc|}\hline \overbrace{1}^{1/2} & \overbrace{2}^{1/2} & \overbrace{3}^{0} & \overbrace{4}^{0} & \overbrace{5}^{1/3} & \overbrace{6}^{1/3} & \overbrace{7}^{1/3} \\ \hline \end{array}$$

This probabilistic belief structure has an infinite number of common delusional priors; for example,

$$\mu = (\frac{1}{7}, \frac{1}{7}, \frac{1}{14}, \frac{1}{14}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}).$$

There can therefore be no common belief disagreement.

We close by noting the following. Suppose that are working in the standard S5 knowledge model (hence that the players make 'no mistakes', that is, they revise beliefs perfectly correctly), and that the players start out with two separate priors, given by

$$\mu_i = (\frac{1}{7}, \frac{1}{7}, \frac{1}{28}, \frac{3}{28}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7})$$

and

$$\mu_j = (\frac{1}{7}, \frac{1}{7}, \frac{1}{14}, \frac{1}{14}, \frac{1}{7}, \frac{1}{7}, \frac{1}{7}).$$

Then the players will revise their beliefs into the following posteriors

$$\widehat{t_i} = \begin{array}{|c|c|ccc|cc|}\hline \overbrace{1}^{1} & \overbrace{2}^{1} & \overbrace{3}^{1/8} & \overbrace{4}^{3/8} & \overbrace{5}^{1/2} & \overbrace{6}^{1/2} & \overbrace{7}^{1/2} \\ \hline \end{array}$$

$$\widehat{t_j} = \begin{array}{|cccc|ccc|}\hline \overbrace{1}^{1/3} & \overbrace{2}^{1/3} & \overbrace{3}^{1/6} & \overbrace{4}^{1/6} & \overbrace{5}^{1/3} & \overbrace{6}^{1/3} & \overbrace{7}^{1/3} \\ \hline \end{array}.$$

Defining a bet $(f_i, -f_i)$ by

$$f_i = (1/4, 1/4, -6, 3, -1/8, 1/32, 1/32),$$

it can be checked that this bet is common knowledge agreeable at every state. But if the players make mistakes, using delusional revision with both players having deluded states at 3 and 4, then instead of $\widehat{t_i}$ and $\widehat{t_j}$ they will derive the posteriors t_i and t_j, which as we have seen have a common delusional prior precluding disagreement. ♦

9. Dynamically Agreeing to Agree

The previous section dealt with conditions related to atemporal impossibility of disagreeing when the truth axiom is not assumed. In this section we consider dynamic back-and-forth processes leading to agreement (as in the example in the introduction) in the interactive KD45 setting.

9.1. **S5 Dynamic Agreement.**

We first review the model of dynamic agreement in the standard S5 (knowledge) case, as introduced in Geanakoplos and Polemarchakis (1982), restricting attention to the two-player case for simplicity.[11]

[11] The two-player model has been extended to multi-player settings in several directions (see, for example).

We suppose that there are two players, labelled 1 and 2, who share a common prior μ over the state space Ω. Each player has an initial partition, Π_1^0 and Π_2^0, of the state space, which induces a pair of types t_1^0 and t_2^0 by standard belief revision, i.e., $t_i(\omega)(\omega') = \mu(\omega')/\mu(\Pi_i^0(\omega))$. A random variable f over Ω is given. It is assumed that it is common knowledge that the players share the common prior μ, that their respective partitions are Π_1^0 and Π_2^0, and that f is the random variable under consideration.

Fix a state $\omega \in \Omega$. We assume that there is a discrete sequence of times $\tau = 0, 1, \ldots$, and define by induction a partition element $\Pi_i^\tau(\omega)$ for each player i, with Π_i^0 given at the start. At each time τ, each player i announces $A_i^\tau(\omega)$, which is his expected value of f conditional on Π_i^τ:

$$A_i^\tau(\omega) := E_i^{t_i^\tau}(f \mid \Pi_i^\tau(\omega)).$$

From this, define

$$W_1^\tau(\omega) := \{\omega' \in \Pi_1^\tau(\omega) \mid E_2^{t_2^\tau}(f \mid \Pi_2^\tau(\omega)) = A_2^\tau(\omega)\}$$

and

$$W_2^\tau(\omega) := \{\omega' \in \Pi_2^\tau(\omega) \mid E_1^{t_1^\tau}(f \mid \Pi_1^\tau(\omega)) = A_1^\tau(\omega)\}.$$

These ingredients are used to define the next step:

$$\Pi_i^{\tau+1}(\omega) := \Pi_i^\tau(\omega) \cap W_i^\tau(\omega),$$

for $i = 1, 2$, with a corresponding $\tau_i^{\tau+1}(\omega)(\omega') = \mu(\omega')/\mu(\Pi_i^{\tau+1}(\omega))$ for $\omega' \in \Pi_i^{\tau+1}(\omega)$.

The elements of the sequence of triples

$$\{\Pi_i^0(\omega), A_i^0(\omega), W_i^0(\omega)\}, \{\Pi_i^1(\omega), A_i^1(\omega), W_i^1(\omega)\}, \ldots$$

satisfy the following properties:

(1) For each τ, $\Pi_1^0(\omega) \cap \Pi_2^0(\omega) \subseteq \Pi_1^\tau(\omega) \subseteq \Pi_1^0(\omega)$ and $\Pi_1^0(\omega) \cap \Pi_2^0(\omega) \subseteq \Pi_2^\tau(\omega) \subseteq \Pi_2^0(\omega)$.
(2) There is a time τ_0 such that for all $\tau > \tau_0$, $\Pi_i^{\tau+1}(\omega) = \Pi_i^\tau(\omega)$ and $A_i^{\tau+1}(\omega) = A_i^\tau(\omega)$.

By (2), the revision process converges after a finite number of steps to agreement among the players.

9.2. KD45 Dynamic Agreement.

In this section, we suppose again that there is a common prior μ over Ω. We also suppose that each player i has a signals set S_i and a signalling function $\sigma_i : \Omega \to S_i$, which induces a partition $\Pi_i^{\sigma_i}$ of Ω (as in Equation (9)). Again, a random variable f over Ω is under consideration. The common prior μ, the random variable f, and the signalling functions σ_1 and σ_2 are common knowledge.

As in Section 3.2, each player misreads some of his signals, and there is a function $\chi_i : \Omega \to S_i$, defined as in Equation (10), which induces a belief structure $(b_1^{\chi_1}, b_2^{\chi_2})$ with corresponding partitions $(\Theta_1^{\chi_1}, \Theta_2^{\chi_2})$. Furthermore, we assume that this belief structure satisfies interpersonal belief credibility, i.e., for all $\omega \in \Omega$, $b_1^{\chi_1}(\omega) \cap b_2^{\chi_2}(\omega) \neq \emptyset$.

We will assume that each player is ignorant of the fact that sometimes he misreads signals, and is also ignorant of the fact that the other player sometimes misreads signals. Hence, if ω is the true state, player i is convinced that the set of possible true states is $b_i^{\chi_i}(\omega)$; implementing delusional revision, he computes $t_i(\omega)(\omega') = \mu(\omega')/\mu(b_i^{\chi_i}(\omega))$ for $\omega' \in b_i^{\chi_i}(\omega)$.

But we know by Equation (13) that for any $\hat{\omega} \in b_i^{\chi_i}(\omega)$, $b_i^{\chi_i}(\omega) = \Pi_i^{\sigma_i}(\hat{\omega})$, and therefore $t_i(\omega)(\omega') = \mu(\omega')/\Pi_i^{\sigma_i}(\hat{\omega})$. Player i then reports $A_i^0(\omega) := E_i^{t_i^\tau}(f \mid \Pi_i^{\sigma_i}(\hat{\omega}))$. In particular, choose $\hat{\omega} \in b_1^{\chi_1}(\omega) \cap b_2^{\chi_2}(\omega)$, and use the reports $A_i^0(\omega)$ to define

$$W_1^0 := \{\omega' \in \Pi_1^{\sigma_i}(\hat{\omega}) \mid E_2^{t_2^\tau}(f \mid \Pi_2^{\sigma_i}(\hat{\omega}) = A_2^0(\omega))\}$$

and

$$W_2^0 := \{\omega' \in \Pi_2^{\sigma_i}(\hat{\omega}) \mid E_1^{t_1^\tau}(f\Pi_1^{\sigma_i}(\hat{\omega}) = A_1^\tau(\omega))\}.$$

The rest of the dynamic follows by induction as in the previous subsection, with $b_i^{\tau+1}(\omega) = b_i^\tau(\omega) \cap W_i^0$, and the sequences

$$\{b_i^1(\omega), A_i^1(\omega), W_i^1(\omega)\}, \{b_i^2(\omega), A_i^2(\omega), W_i^2(\omega)\} \ldots$$

satisfying the property that there is a τ_0 such that for all $\tau > \tau_0$, $b_i^{\tau+1}(\omega) = b_i^\tau(\omega)$ and $A_i^{\tau+1}(\omega) = A_i^\tau(\omega)$.

This constitutes a proof of Theorem 4.

Theorem 4. *The elements of the sequence of triples*

$$\{b_i^1(\omega), A_i^1(\omega), W_i^1(\omega)\}, \{b_i^2(\omega), A_i^2(\omega), W_i^2(\omega)\} \ldots$$

satisfy the following properties:

(1) *For each τ, $b_1^0(\omega) \cap b_2^0(\omega) \subseteq b_1^\tau(\omega) \subseteq b_1^0(\omega)$ and $b_1^0(\omega) \cap b_2^0(\omega) \subseteq b_2^\tau(\omega) \subseteq b_2^0(\omega)$.*
(2) *There is a time τ_0 such that for all $\tau > \tau_0$, $b_i^{\tau+1}(\omega) = b_i^\tau(\omega)$ and $A_i^{\tau+1}(\omega) = A_i^\tau(\omega)$.*

This revision process converges after a finite number of steps to agreement among the players.

10. Appendix: Proofs

Proof of Lemma 1. Let $(t_i)_{i\in I}$ be non-singular. If $(t_i)_{i\in I}$ has a common standard prior μ, then since μ is also a delusional prior for each t_i, it is a common standard prior.

In the other direction, let μ be a common delusional prior. Let $D \subset \Omega$ be the set of states that at which all players in I are deluded. Define a probability distribution $\widehat{\mu}$ by

$$\widehat{\mu}(\omega) = \begin{cases} 0 & \text{if } \omega \in D \\ \frac{\mu(\omega)}{\mu(\Omega \setminus D)} & \text{otherwise} \end{cases}$$

It is straight-forward to check that $\widehat{\mu}$ is a common standard prior (since $t_i(\omega) = 0$ for all i at each state $\omega \in D$). ∎

Proof of Lemma 2. Denote $F = \Omega \setminus D_i$. Each state in $\omega \in F$ is non-deluded (for player i), hence $\omega \in b_i(\omega) \subseteq F$. It follows that $F \subset B_i(F)$. If $\omega \in D_i$, then $\omega \notin b_i(\omega)$, but it is still the case that $b_i(\omega) \subseteq F$, so that $D_i \subset B_i(F)$. Since $\Omega = F \cup D_i$, $B_i(F) = \Omega$.

Next, suppose that E is a proper subset of F. Then $F \setminus E$ contains a non-deluded state ω of player i, hence $\omega \in b_i(\omega) \not\subset E$. But from that one concludes that $\omega \notin B_i(E)$. F is therefore the minimal event that satisfies $B_i(F) = \Omega$. ∎

Proof of Lemma 3.[12] Suppose that $\omega' \in b_i(\omega)$. Let $E \in \mathrm{con}_j(\omega')$. Then $\omega' \in B_j(E)$. If $\neg B_j(E) \in \mathrm{con}_i(\omega)$ then $\omega \in B_i(\neg B_j(E))$ and hence $\omega' \in b_i(\omega) \subseteq \neg B_j(E)$, a contradiction.

If $E \notin \mathrm{con}_j(\omega')$ then $\omega' \notin B_j(E)$. If $B_j(E) \in \mathrm{con}_j(\omega')$ then $\omega \in B_i(B_j(E))$ and therefore $\omega' \in b_i(\omega) \subseteq B_j(E)$, a contradition. ∎

Proof of Lemma 4. Let $\mathbb{C} \in \mathcal{C}_i(b^Q(\omega))$. Then $\mathbb{C} = \mathrm{con}_i(\omega')$ for some $\omega' \in \omega \cup b^Q(\omega)$, which holds if and only if $\mathbb{C} = \{E \mid b_i(\omega') \subseteq E\}$. But for any $\omega'' \in \Pi_i(\omega')$, $b_i(\omega'') = b_i(\omega')$, hence

$$\{E \mid b_i(\omega'') \subseteq E\} = \{E \mid b_i(\omega') \subseteq E\},$$

which means that $\mathbb{C} \in \mathcal{C}_i(b^Q(\omega))$ iff

$$\mathbb{C} \in \{\mathrm{con}_i(\omega'') \mid \omega'' \in T^i(b^Q(\omega))\}.$$

∎

Proof of Lemma 5. Fix ω. We first show that $\bigcup_{\omega' \in b^Q(\omega)} b_i(\omega') \subseteq b^Q(\omega)$. Suppose that $\omega'' \in \bigcup_{\omega' \in b^Q(\omega)} b_i(\omega')$. Then it must be the case that $\omega'' \in b_i(\omega')$ for some $\omega' \in b^Q(\omega)$. At the same time, $\omega' \in \bigcup_{j=1}^{m} b^j(\omega)$ for some integer m. Now, $b(\omega') = \bigcup_{i \in I} b_i(\omega')$, hence $\omega'' \in b(\omega')$, i.e. $\omega'' \in \bigcup_{j=1}^{m+1} b^j(\omega)$, and therefore $\omega'' \in b^Q(\omega)$.

Next, suppose that $\omega'' \in b^Q(\omega)$. Distinguish two cases:

(i) $\omega'' \in b_i(\omega'')$. Then $\omega'' \in \Pi_i(\omega'')$.
(ii) $\omega'' \notin b_i(\omega'')$. Let $\omega' \in \beta_i(\omega'')$. Then $\omega' \in b^Q(\omega)$ and $\omega'' \in \Pi_i(\omega')$, hence the proof is completed. ∎

Proof of Proposition 1. (a) We first prove the following statement. Suppose that $\mathrm{con}_i(\omega') \in \mathcal{C}_i(b^Q(\omega))$ for some state ω', and that $\Pi_j(\omega'') \cap b_i(\omega') = \emptyset$ for some state ω''. Then $\mathrm{con}_j(\omega'') \notin \mathrm{Permit}_j(\mathrm{con}_i(\omega'))$.

Indeed, let $E = b_j(\omega'')$, which implies that $B_j(E) = \Pi_j(\omega'')$, equivalently $\neg B_j(E) = \neg\Pi_j(\omega'')$. By assumption, $\Pi_j(\omega'') \cap b_i(\omega') = \emptyset$, hence $b_i(\omega') \subseteq \neg B_j(E)$. This implies that $\neg B_j(E) \in \mathrm{con}_i(\omega')$. This suffices to show that $\mathrm{con}_j(\omega'') \notin \mathrm{Permit}_j(\mathrm{con}_i(\omega'))$, which is what we wanted to show.

Therefore, every conviction $\mathrm{con}_j(\omega'')$ of j that is permitted by $\mathrm{con}_i(\omega')$ must satisfy the property that $\Pi_j(\omega'') \cap b_i(\omega') \neq \emptyset$. Thus it suffices to show that if $\Pi_j(\omega'') \cap b_i(\omega') \neq \emptyset$ then $\mathrm{con}_j(\omega'') \in \mathcal{C}_j(b^Q(\omega))$. But as we supposed that $\mathrm{con}_i(\omega') \in \mathcal{C}_i(b^Q(\omega))$, it follows that $b_i(\omega') \subset b^Q(\omega)$. Hence furthermore, if there is a state $\omega''' \in \Pi_j(\omega'') \cap b_i(\omega')$ then $b_j(\omega''') \subset b^Q(\omega)$, while $b_j(\omega''') = b_j(\omega'')$. This implies that $\mathrm{con}_j(\omega'') \in \mathcal{C}_j(b^Q(\omega))$.

(b) Let $\mathrm{con}_i(\omega) \in \mathcal{K}_i$. By the definition of $b^Q(\omega)$ (see Equation (20)), any $\omega' \in b^Q(\omega)$ satisfies $\omega' \in b_k(\omega'')$ for some $\omega'' \in \omega \cup b^Q(\omega)$ and some $k \in I$. By Lemma 3, for any $\omega' \in b_k(\omega'')$ and all $j \in I$, $\mathrm{con}_j(\omega') \in \mathrm{Permit}_j(\mathrm{con}_k(\omega))$. By the assumption of permission consistency, $\mathrm{Permit}_j(\mathrm{con}_k(\omega)) \subseteq \mathcal{K}_j$. This is sufficient to conclude that $\mathcal{C}_j(b^Q(\omega)) \subseteq \mathcal{K}_j$ for all $j \in I$. ∎

[12] This proof is very similar to the proof of Lemma 1 in Samet (2008).

Proof of Lemma 6. It suffices to show that

$$\bigcap_{\mathbb{C}_i \in \mathcal{C}_i} \mathbb{C}_i = \{E \mid S_i(\omega) \subseteq E \text{ for all } \omega \in \Omega\,|_{\mathcal{C}_i}\}.$$

But $E \in \mathrm{con}_i(\omega)$ if and only if $b_i(\omega)$ is contained in E. Hence $E \in \mathrm{con}_i(\omega)$ for all $\omega \in \Omega\,|_{\mathcal{C}_i}$ if and only if $b_i(\omega)$ is contained in E for all $\omega \in \Omega\,|_{\mathcal{C}_i}$, which is equivalent to $S_i(\omega) \subseteq E$ for all $\omega \in \Omega\,|_{\mathcal{C}_i}$. ∎

Proof of Corollary 1. Suppose that there is a state $\omega \in \Omega\,|_{\mathcal{C}_i}$ such that there exists $\omega'' \in b_i(\omega)$ with $\omega'' \notin b_j(\omega')$ for any $\omega' \in \Omega\,|_{\mathcal{C}_j}$. But then, by Equation 28, $\bigcup_{\omega \in \Omega|_{\mathcal{C}_i}} b_i(\omega)$ is an event in $\bigcap_{\mathbb{C}_i \in \mathcal{C}_i} \mathbb{C}_i$, while $\bigcup_{\omega' \in \Omega|_{\mathcal{C}_j}} b_j(\omega') \notin \bigcap_{\mathbb{C}_i \in \mathcal{C}_i} \mathbb{C}_i$. This contradicts the assumption that $\mathcal{C}_i$ and $\mathcal{C}_j$ are truth equivalent. ∎

Proof of Proposition 2 For each i, $b_i^{\mathcal{C}}$ satisfies transitivity and euclideaness because $b_i^{\mathcal{C}}(\omega) = b_i(\omega)$ by definition, and b_i satisfies those properties. To show that $\mathbf{\Pi}^{\mathcal{C}}$ is a knowledge structure, all that is left is to show that $b_i^{\mathcal{C}}$ satisfies the truth axiom; but this follows from the fact that all the states in $ND(\mathcal{C})$ are non-deluded for b_i. ∎

Proof of Lemma 7. By Proposition 2, the belief structure $\mathbf{\Pi}^{\mathcal{C}}$ is a knowledge structure. It follows that for any $\omega' \in ND(\mathcal{C})$, the common belief set $b^Q(\omega')$ of ω' in $\mathbf{\Pi}$ is identical to one of the common knowledge components in $\mathbf{\Pi}^{\mathcal{C}}$, which is $ND(\mathcal{C})$.

To complete the proof, since $b_i(\omega) \subseteq ND(\mathcal{C})$ for all i, $b(\omega) \subseteq ND(\mathcal{C})$, and then by induction $\omega' \in ND(\mathcal{C})$ for all $\omega' \in b^Q(\omega)$. ∎

Proof of Lemma 8. In one direction, suppose that there is strong common belief in truth at a state ω. Then by Lemma 7, $b^Q(\omega) \subseteq ND(C)$, i.e., each state in $b^Q(\omega)$ is non-delusional for all players. Setting $\Omega_0 = b^Q(\omega)$, it follows that $b^Q(\omega) = \bigcup_{\omega' \in \Omega_0} b_i(\omega')$.

In the other direction, suppose that there exists $\Omega_0 \subseteq \Omega$ such that $b^Q(\omega) = \bigcup_{\omega' \in \Omega_0} b_i(\omega')$ for all $i \in I$. This implies that $\bigcup_{\omega' \in \Omega_0} b_i(\omega') = \bigcup_{\omega' \in \Omega_0} b_j(\omega')$ for all $i, j \in I$. Let $\mathcal{C}_i := \{\mathrm{con}_i(\omega') \mid \omega' \in \Omega_0\}$ for all $i \in I$, and let $\mathcal{C}$ be the set of all the families $\mathcal{C}_i$ for $i \in I$. Then $\bigcap_{\mathbb{C}_i \in \mathcal{C}_i} = \bigcap_{\mathbb{C}_j \in \mathcal{C}_j}$ for all $i, j \in I$, and $\omega' \in ND(\mathcal{C})$ for all $\omega' \in \Omega_0$.

Since $b^Q(\omega) = \bigcup_{\omega' \in \Omega_0} b_i(\omega')$ for all $i \in I$, it follows that $b(\omega) \subseteq ND(C)$, which further implies that $b_i(\omega) \subseteq ND(\mathcal{C})$ for all i, hence there is strong common belief in truth at ω. ∎

Proof of Proposition 3. We prove the equivalence of items (1) through (4) step by step:

(1) All players perceive the same world $\Rightarrow$ $\mathbf{\Pi}$ satisfies interpersonal belief consistency.

Proof: By Lemma 2, the world perceived by player i is $\Omega \setminus D_i$, where D_i is the set of states at which i is deluded. It follows that if $\mathbf{\Pi}$ does not satisfy interpersonal belief consistency, and hence there is a state ω at which player i is deluded but player j is not deluded, then the world perceived by i does not contain ω but the world perceived by j does contain ω.

(2) $\mathbf{\Pi}$ satisfies interpersonal belief consistency $\Rightarrow$ for all players i and j, $\bigcap_{\mathbb{C}_i \in \mathrm{Con}_i} \mathbb{C}_i = \bigcap_{\mathbb{C}_j \in \mathrm{Con}_j} \mathbb{C}_j$.

Proof: By Lemma 6, for a player i, using Equation (28) with the family of convictions Con_i, $\bigcap_{\mathbb{C}_i \in \mathrm{Con}_i} \mathbb{C}_i = \{E \mid \bigcup_{\omega \in \Omega} b_i(\omega) \subseteq E\}$ (since in this case $\Omega\,|_{\mathcal{C}_i} = \Omega$). It follows that if $\bigcap_{\mathbb{C}_i \in \mathrm{Con}_i} \mathbb{C}_i \neq \bigcap_{\mathbb{C}_j \in \mathrm{Con}_j} \mathbb{C}_j$, then $\bigcup_{\omega \in \Omega} b_i(\omega) = \bigcup_{\omega' \in \Omega} b_j(\omega')$. But that implies that there is a state ω at which either i is deluded and j is not deluded, or j is deluded and i is not deluded.

(3) For all players i and j, $\bigcap_{\mathbb{C}_i \in \mathrm{Con}_i} \mathbb{C}_i = \bigcap_{\mathbb{C}_j \in \mathrm{Con}_j} \mathbb{C}_j \Rightarrow$ for all players i and j, $\bigcup_{\omega \in \Omega} b_i(\omega) = \bigcup_{\omega' \in \Omega} b_j(\omega')$

Proof: This is Corollary 1 applied to Con_i and Con_j.

(4) For all players i and j, $\bigcup_{\omega \in \Omega} b_i(\omega) = \bigcup_{\omega' \in \Omega} b_j(\omega') \Rightarrow$ all players perceive the same world.

Proof: By Lemma 2, the world perceived by player i is $\Omega \setminus D_i$, where D_i is the set of states at which i is deluded. But if $\bigcup_{\omega \in \Omega} b_i(\omega) = \bigcup_{\omega' \in \Omega} b_j(\omega')$, then $\Omega \setminus D_i = \Omega \setminus D_j$.

Finally, we show that all players perceive the same world iff $\mathbf{\Pi}$ is non-singular.

Suppose that $\mathbf{\Pi}$ is non-singular. Then $\mathcal{C} = (\mathrm{Con}_i)_{i \in I}$ is a set of truth equivalent families. But then for each state ω, since $b_i(\omega)$ is a set of non-delusional states for player i, $b_i(\omega) \subseteq ND(C)$, and hence there is strong common belief in truth at ω.

In the other direction, suppose that there is strong common belief in truth at every state ω. By Lemma 8, there exists Ω_0 such that $b^Q(\omega) = \bigcup_{\omega' \in \Omega_0} b_i(\omega')$ for all players $i \in I$. From this, we can deduce that every state ω that is non-delusional for a player i (hence $\omega \in b^Q(\omega)$) must also be non-delusional for every other player $j \in I$. Hence $\mathbf{\Pi}$ is non-singular. ∎

Proof of Lemma 9. Since $\mathbf{\Pi}$ satisfies interpersonal belief consistency, there is strong common belief in truth at every state.

(1) Let ω be a non-deluded state for all the players. By Lemma 8, there exists $\Omega_0 \subseteq \Omega$ such that $b^Q(\omega) = \bigcup_{\omega' \in \Omega_0} b_i(\omega')$ for all $i \in I$.

We will show that Ω_0 satisfies conditions (a) and (b) of the definition of $ND^i(\omega)$. Since ω is non-deluded for i, $\omega \in b^Q(\omega) = \bigcup_{\omega' \in \Omega_0} b_i(\omega')$, hence $b_i(\omega) \subseteq \Omega_0$. This shows that condition (a) is satisfied. Since $b^Q(\omega) = \bigcup_{\omega' \in \Omega_0} b_i(\omega')$ for all i, it follows that $\bigcup_{\omega' \in \Omega_0} b_i(\omega') = \bigcup_{\omega' \in \Omega_0} b_j(\omega')$ for all $i, j \in I$. Hence condition (b) is satisfied.

Suppose that there exists $\Omega_1 \subsetneq \Omega_0$ also satisfying (a) and (b). Then it would be the case that $b^Q(\omega) = \bigcup_{\omega' \in \Omega_1} b_i(\omega')$, contradicting $b^Q(\omega) = \bigcup_{\omega' \in \Omega_0} b_i(\omega')$. The conclusion is that $\Omega_0 = ND^i(\omega)$ for all $i \in I$.

(2) Since any $\omega' \in b_i(\omega)$ is a non-deluded state, by the proof of part (1) of this theorem $b^Q(\omega') = ND^i(\omega')$ for all such ω'. Inserting this into Equation (22), according to which $b^Q(\omega) = \bigcup_{\{\omega' \in \bigcup_{i \in I} b_i(\omega)\}} b^Q(\omega')$, we conclude that $b^Q(\omega) = \bigcup_{i \in I} ND^i(\omega)$. ∎

Proof of Proposition 4. Suppose that $\bigcap_{\mathbb{C}_i \in \mathcal{C}_i} \mathbb{C}_i = \bigcap_{\mathbb{C}_j \in \mathcal{C}_j} \mathbb{C}_j$. Let $\mathrm{con}_i(\omega) \in \mathcal{C}_i$. We first prove that if $b_j(\omega') \cap b_i(\omega) = \emptyset$, then $\mathrm{con}_j(\omega') \notin \mathrm{Permit}_j(\mathbb{C}_i)$. To see this, let $E = b_j(\omega')$. Clearly, $E \in \mathrm{con}_j(\omega')$. Suppose that $\neg B_j(E) \notin \mathrm{con}_i(\omega)$. Then $b_i(\omega) \not\subset \neg B_j(E)$, meaning that there is a state $\omega'' \in b_i(\omega)$ such that $\omega'' \in B_j(E)$. By assumption $b_j(\omega') \cap b_i(\omega) = \emptyset$, hence $\omega'' \notin b_j(\omega')$, meaning that $\omega'' \in f_j(\omega')$.

But that means that $\omega'' \notin b_j(\omega')$, but $\omega'' \in b_i(\omega)$. At the same time, $\mathrm{con}_j(\omega'') \in \mathcal{C}_j$. But then $\bigcup_{\omega\in\Omega|_{\mathcal{C}_i}} b_i(\omega) \neq \bigcup_{\omega'\in\Omega|_{\mathcal{C}_j}} b_j(\omega)$, contradicting Equation (29). The conclusion is that $\neg B_j(E) \in \mathrm{con}_i(\omega)$, showing that $\mathrm{con}_j(\omega') \notin \mathrm{Permit}_j(\mathrm{con}_i(\omega))$.

We have shown that all convictions of j that are permitted by $\mathrm{con}_i(\omega)$ must intersect $\mathrm{con}_i(\omega)$. It therefore suffices to show that if $b_j(\omega') \cap b_i(\omega) \neq \emptyset$, then $\mathrm{con}_j(\omega') \in \mathcal{C}_j$. Suppose that $\mathrm{con}_j(\omega') \notin \mathcal{C}_j$. Let $E = \neg b_j(\omega')$ and let $\omega'' \in \Omega\,|_{\mathcal{C}_j}$, i.e. $\mathrm{con}_j(\omega'') \in \mathcal{C}_j$. Since we are working in a KD45 model, this implies that $b_j(\omega') \cap b_j(\omega'') = \emptyset$. That then implies that $E \in \mathrm{con}_j(\omega'')$.

Hence $E \in \bigcap_{\mathbb{C}_i\in\mathcal{C}_i} \mathbb{C}_i$. As $\mathcal{C}_i$ and $\mathcal{C}_j$ are truth equivalent, $E \in \bigcap_{\mathbb{C}_j\in\mathcal{C}_j} \mathbb{C}_j$, and in particular $E \in \mathrm{con}_i(\omega)$. But this is impossible, since $b_i(\omega) \not\subseteq E$. ∎

Proof of Theorem 1. In the first direction, it suffices to restrict attention to any pair of players $i, j \in I$. Suppose that PCC holds and assume that $\omega \in Q_{ij}([\mathbf{d}_i = d_i] \cap [\mathbf{d}_j = d_j])$. Recall the definition in Equation (25), of

$$\mathcal{C}_i(b^Q(\omega)) := \{\mathrm{con}_i(\omega') \mid \omega' \in \omega \cup b^Q(\omega)\}.$$

We want to show that (1) and (2) in the definition of PCC hold for $\mathcal{C}_i(b^Q(\omega))$ and $\mathcal{C}_j(b^Q(\omega))$.

For (1), it suffices to show that for any $\omega' \in b^Q(\omega)$ and any ω'' such that $\mathrm{con}_j(\omega'') \notin \mathcal{C}_j(b^Q(\omega))$ one has $\mathrm{con}_j(\omega'') \notin \mathrm{Permit}_j(\mathrm{con}_i(\omega'))$. Towards this end, let $E = b^Q(\omega)$. By Lemma 5, for each $\overline{\omega} \in E$, $b_j(\overline{\omega}) \subseteq E$, hence $E \subseteq B_j(E)$. Now, $b_i(\omega') \subseteq E$ for each $\overline{\omega} \in E$, hence $\omega' \in B_i(B_j(E))$, which is another way of saying $B_j(E) \in \mathrm{con}_i(\omega')$.

By Lemma 4, it cannot be the case that $\omega'' \in T^j(b^Q(\omega))$ (where T^j is as defined in Equation 24), because if $\omega'' \in T^j(b^Q(\omega))$, then $\mathrm{con}_j(\omega'') \in \mathcal{C}_j(E)$, which contradicts one of our assumptions.

Hence $\omega'' \notin T^j(b^Q(\omega))$, which implies that $b_j(\omega'') \not\subseteq E$, and we conclude that $E \notin \mathrm{con}_j(\omega'')$.

To show (2), note that by Equation (21), $\omega \in Q_{ij}([\mathbf{d}_i = d_i] \cap [\mathbf{d}_j = d_j])$ implies that $b^Q(\omega) \subseteq C_{ij}([\mathbf{d}_i = d_i] \cap [\mathbf{d}_j = d_j])$. Since $b_i(\omega) \subseteq b^Q(\omega)$ for each i, and the conviction of i at ω and hence the decision at ω is determined by $b_i(\omega)$, $\mathbf{d}_i(\omega) = d_i$. Since $b^Q(\omega') \subseteq b^Q(\omega)$ for every $\omega' \in b^Q(\omega)$, the same reasoning gives us that $\mathbf{d}_i(\omega') = d_i$ for all i and all $\omega' \in b^Q(\omega)$. Thus, for each $\mathbb{C}_i \in \mathcal{C}_i(b^Q(\omega))$, $\mathbf{d}_i(\mathbb{C}_i) = d_i$, and a similar equality holds for j. By PCC, it follows that $d_i = d_j$, as required.

In the other direction, suppose that IAD holds, and assume that (1) and (2) in the definition of PCC hold for a collection of families of convictions $(\mathcal{C}_i)_{i\in I}$. Since $(\mathcal{C}_i)_{i\in I}$ are permission consistent, by Proposition 1 for any state ω such that $\mathrm{con}_i(\omega) \in \mathcal{C}_i$ for some player i, $\mathcal{C}_j(b^Q(\omega)) \subseteq \mathcal{C}_j$ for all $j \in I$. Hence for each $\omega' \in b^Q(\omega)$, $\mathrm{con}_i(\omega') \in \mathcal{C}_i$ for all $i \in I$. Therefore, by property (2) of PCC, $\mathbf{d}_i(\omega') = d_i$ for all $i \in I$. It follows that for any pair of players i and j, $\omega \in Q_ij([\mathbf{d}_i = d_i] \cap [\mathbf{d}_j = d_j])$, i.e, $Q_ij([\mathbf{d}_i = d_i] \cap [\mathbf{d}_j = d_j]) \neq \emptyset$. By IAD, $d_i = d_j$, showing that permission consistent consensus holds. ∎

Proof of Theorem 2. We first add a definition and a lemma, for the sake of proving the theorem.

Definition 11. Let $(t_i)_{i\in I}$ be a probabilistic belief structure over Ω with corresponding partition profile $\Pi := (\Pi_i)_{i\in I}$, and let $X \subset \Omega$ be a subset of Ω.

Define Π *restricted to* X, denoted Π^X, to be the partition profile over X given by $\Pi_i^X(\omega) := \Pi_i(\omega) \cap X$ for any state ω. Further, for each $i \in I$ let t_i^X be any type function over $(X, \mathbf{\Pi}^X)$ that satisfies the property that for any $\omega \in \Omega$, $t_i(\omega)(\Pi_i^X)t_i^X(\omega) = t_i(\omega)$. ♦

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 $\sum_{\omega \in X} t_i^X(\omega) = 1$.

For a random variable f, denote

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

A bet $\{f_1, \ldots, f_n\}$ is an *agreeable bet relative to* $(t_i^X)_i$ *at* $\omega \in X$ if $E_i^X(f \mid \omega) > 0$ for all $i \in I$. We will say that it is simply an agreeable relative to $(t_i^X)_i$ if it is an agreeable bet relative to $(t_i^X)_i$ at all states $\omega \in X$.

Lemma 10. *Let $(t_i)_{i \in I}$ be a probabilistic belief structure over Ω, let $\omega \in \Omega$ and let X be a non-empty subset of $b^Q(\omega)$, the common belief set of ω. Suppose that there exists an agreeable bet relative to $(t_i^X)_i$. Then there exists an agreeable bet relative to $b^Q(\omega)$.*

Proof. Let f be an agreeable bet relative to $(t_i^X)_i$. If $X = b^Q(\omega)$, there is nothing to prove.

Otherwise, we distinguish a few cases:

(1) Suppose that there exists a state $\omega'' \in X$ such that $b_i(\omega'') \setminus X \neq \emptyset$ for some $i \in I$. Let $\omega' \in b_i(\omega'') \setminus X$ (hence $t_i(\omega') > 0$), and let $\varepsilon := E_i^X(f_i \mid \Pi_i^X(\omega')) = E_i^X(f_i \mid \Pi_i^X(\omega''))$. By assumption, $\varepsilon > 0$ (since f is an agreeable bet relative to $(t_i^X)_i$). Set $Y := X \cup \omega'$.

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

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

and for $j \neq i$, set $\overline{f}_j(\omega') := -\overline{f}_i(\omega')/(n-1) > 0$, where $n = |I|$.

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 $(t_i^Y)_{i \in I}$.

(2) Suppose that there is a state $\omega' \in b^Q(\omega) \setminus X$ such that $b_i(\omega') \cap X \neq \emptyset$. Set $Y := X \cup \omega'$.

We distinguish two sub-cases:

(a) If $t_i(\omega') = 0$, then for all $j \in I \setminus i$ let $\overline{f}_j(\omega')$ be any arbitrary positive number, and set $\overline{f}_i(\omega') = -\sum_{j \in I \setminus i} \overline{f}_j(\omega')$. Then $\overline{f}$ is an agreeable bet relative to $(t_i^Y)_{i \in I}$.

(b) If $t_i(\omega') > 0$, let $\varepsilon := E_i^X(f_i \mid \Pi_i^X(\omega'))$. By assumption, $\varepsilon > 0$ (since $b_i(\omega') \cap X \neq \emptyset$ and f is an agreeable bet relative to $(t_i^X)_i$). From this point, define $\overline{f}_j$ for all $j \in I$ exactly as in Case 1 above, yielding an agreeable bet relative to $(t_i^Y)_{i \in I}$.

Now simply repeat this procedure as often as necessary to extend the agreeable bet to every state in the finite set $b^Q(\omega)$. ∎

Completion of the proof of Theorem 2. Let $(t_i)_{i \in I}$ be a probabilistic belief structure over Ω, and let ω be a state at which there is weak common belief in truth, and hence there is $\omega' \in b^Q(\omega)$ at which there is strong common belief in truth, i.e.,

$$\bigcup_{\omega'' \in b^Q(\omega')} b_i(\omega'') = \bigcup_{\omega'' \in b^Q(\omega')} b_j(\omega'')$$

for all $i, j \in I$. If we restrict attention solely to the states in $b^Q(\omega')$, we can consider the operators b_i for all i to constitute an S5 knowledge structure over $b^Q(\omega)$, as in the proof of Proposition 2.

In one direction, suppose that there is a common delusional prior μ. Then μ restricted to $b^Q(\omega')$ is a common (standard) prior over $b^Q(\omega')$ regarded as a knowledge structure, hence there can be no common knowledge agreeable bet at any state in $b^Q(\omega')$. If there were a common belief agreeable bet at ω, then that bet would be a common knowledge agreeable bet over $b^Q(\omega')$ regarded as a knowledge structure, which we just showed cannot happen. The contradiction establishes that there is no common belief agreeable bet at ω.

In the other direction, suppose that there is no common delusional prior. Then there can be no common (standard) prior over $b^Q(\omega')$ regarded as a knowledge structure, because if there were such a prior μ, it could be extended to a common delusional prior $\widehat{\mu}$ over all of $b^Q(\omega)$ simply by setting

$$\widehat{\mu}(\omega'') = \begin{cases} \mu(\omega'') & \text{if } \omega'' \in b^Q(\omega') \\ 0 & \text{otherwise.} \end{cases}$$

We can then apply the standard No Betting Theorem for knowledge structures to conclude that there is a common knowledge agreeable bet $\{f_1, \dots, f_n\}$ over $b^Q(\omega')$ as a knowledge structure, which is a common belief agreeable bet over $b^Q(\omega')$ as a belief structure. Applying Lemma 10, this can be extended to a common belief agreeable bet over all of $b^Q(\omega)$, which is what was needed to be shown. ∎

References

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

Bachrach, M. (1985), Some extensions of a claim of Aumann in an axiomatic model of knowledge. *J. Econ. Theory*, 37 (1), 167–190.

Battigalli, P., and G. Bonanno (1999), Recent Results on Belief, Knowledge and the Epistemic Foundations of Game Theory *Review of Economic Studies*, 64, 23–46.

Bonanno, G., and K. Nehring (1999), How to Make Sense of the Common Prior Assumption under Incomplete Information, *Int. J. Game Theory* 28, 409–434.

Collins, J., (1997), How We can Agree to Disagree. *Working paper.*

Geanakoplos, J. (1989), Game Theory Without Partitions and Applications to Speculation and Consensus, *Cowles Foundation Discussion Paper 914.*

Geanakoplos, J., and H. Polemarchakis (1982), We Can't Disagree Forever, *Journal of the Economic Theory*, (28), 192–200.

Hart, S., and Y. Tauman (2004), Market Crashes without External Shocks, *Journal of Business*, (77), 1–8.

Monderer, D., and D. Samet (1989), Approximating Common Knowledge with Common Beliefs, *Games and Economic Behavior*, 1(2), 170–190.

Morris, S. (1996), The Logic of Belief and Belief Change: A Decision Theoretic Approach, *Journal of the Economic Theory*, (69), 1–23.

Samet, D. (2008), The Sure-thing Principle and the Independence of Irrelevant Knowledge, unpublished mimeo.

Samet, D. (2010b), Agreeing to Disagree: The Non-Probabilistic Case, *Games and Economic Behavior*, (69), 169 – 174.

Samet, D. (2011), Common Belief of Rationality in Games of Perfect Information, *Working Paper*.

Tallon, J. M., J C. Vergnaud and S. Zamir (2004), Communication Among Agents: A Way to Revise Beliefs in KD45 Kripke Structures, *Journal of Applied Non-Classical Logics*, (14), 477 – 500.

Tarbush, B. (2011), Agreeing to disagree with generalised decision functions, *University of Oxford, Department of Economics, Working Paper*.