

AGREEING TO AGREE
OR
WHY WE CANNOT AGREE ON THE ODDS THAT
YOUR FIRM IS MORE PROFITABLE THAN MINE

EHUD LEHRER AND DOV SAMET

September 9, 2003

ABSTRACT. Aumann (1976) has shown that agents who have a common prior cannot have common knowledge of their posteriors for event E if these posteriors do not coincide. But given an event E , can the agents have posteriors with a common prior such that it is common knowledge that the posteriors for E *do* coincide? We show that a necessary and sufficient condition for this is the existence of a nonempty *finite* event F with the following two properties. First, it is common knowledge at F that the agents cannot tell whether or not E occurred. Second, this still holds true at F , when F itself becomes common knowledge.

1. INTRODUCTION

1.1. Agreeing to disagree and agreeing to agree. Can agents have common knowledge of their beliefs? In a seminal paper Aumann (1976) demonstrated the impossibility of *agreeing to disagree*: For *any* pair of posteriors with a common prior, if the agents' posteriors for an event E are different (= *they disagree*), then the agents cannot have common knowledge (= *agreeing*) of these posteriors. Thus, the concise answer to our opening question is: agents cannot have common knowledge of their beliefs when they are different.

But can there be other reasons why agents may fail to have common knowledge of their beliefs, even when these beliefs are the same? Obviously, agents may fail incidentally to know each others posteriors, and thus, a fortiori they cannot have common knowledge of them, even if they coincide. Our question here does not concern such an incidental lack of common knowledge of the posteriors, but rather the impossibility of having such common knowledge in *principle*. That is, we examine the impossibility of having such common knowledge regardless of the beliefs held by the agents.

We investigate the conditions under which there *exist* posteriors with a common prior, such that it is common knowledge that the agents' posteriors for an event E coincide. Using the terminology of Aumann's theorem we ask: when is *agreeing to agree* about the posteriors of E possible? We clarify this question by examining several examples.

Example 1. Each of two firms can be either profitable or losing. Thus, there are four possible states of the world (profitable, profitable), (profitable, losing), (losing, profitable) and (losing, losing). Each firm knows only how well it does.

Consider first the event E that both firms have the same financial situation. That is, $E = \{(\text{profitable}, \text{profitable}), (\text{losing}, \text{losing})\}$. There are many posteriors that have a common prior such that the firms have common knowledge that the posterior probabilities of E coincide. For instance, consider a common prior, μ , with

$$\mu(\text{profitable}, \text{profitable}) = \mu(\text{losing}, \text{losing}) = p,$$

and

$$\mu(\text{profitable}, \text{losing}) = \mu(\text{losing}, \text{profitable}) = q,$$

where $2p + 2q = 1$. Then, *in each of the four states*, both firms have the same posterior for E : $p/(p + q)$. Thus, it is common knowledge in every state that the posteriors for E are the same.

Example 2. In the state space of the previous example, consider the event $E = \{(\text{losing}, \text{profitable})\}$. Note that the only event about which there is common knowledge is the whole state space. Therefore, if for some posteriors it is common knowledge that both posteriors for E are p , then the posteriors for E must be p in all four states. However, for any posterior on this space, when firm 1 is profitable, its posterior for E is 0. Thus, the only value p can have is 0.

Consider the complement of E , denoted $\neg E$. When firm 2 is losing, its posterior for $\neg E$ is 1. Thus, by the same argument as above, the only p for which there can be common knowledge that both posteriors for $\neg E$ are p , is 1.

As opposed to the event E in Example 1, in this example the possibility of agreeing to agree about the posteriors of E , or $\neg E$, is limited to trivial ones, i.e., 0 or 1.

1.2. A necessary condition for the possibility of agreeing to agree. The difference between the events in Examples 1 and 2 suggests a simple necessary condition for agreeing to agree. In any state at which the firm's posterior for E is nontrivial, the firm cannot tell whether or

not E occurred. We say in this case that the firm is *ignorant* of E . Therefore:

If there exist posteriors with a common prior for which it is common knowledge that both posteriors for E are p for some $0 < p < 1$, then it is necessarily common knowledge that both firms are ignorant of E .

Now, in the first example the firms are ignorant of E in each state. Thus, it is common knowledge that they are ignorant of E , and the necessary condition for agreeing to agree is satisfied. In the second example, in contrast, in all states there is no common knowledge that the firms are ignorant of E . Therefore, no matter what nontrivial posteriors for E the agents may have, the event where their posteriors coincide cannot be common knowledge.

Surprisingly, this necessary condition is not sufficient as demonstrated by the next example. The analysis of this example hints at the answer to the question posed in the title of this paper.

Example 3. The profit of each of two firms can be any integer, positive or negative. Each firm is informed only of its profit. Let E be the event that firm 2's profit is higher. That is, E is the set of all states (i, j) where $i < j$. In each state no firm can tell whether or not 2's profit is higher. Therefore, it is common knowledge in each state that both firms are ignorant of E . It turns out, nevertheless, that no matter what posteriors with a common prior the firms may have, there can be no common knowledge that both posteriors for E are p , for some $0 < p < 1$.

1.3. Necessary and sufficient conditions. Our first result provides a necessary and sufficient condition for the possibility of agreeing to agree in terms of the information structure only. Although the notion of agreeing to agree involves posteriors and common priors, our characterization does not employ any probabilistic notions.

Agreeing to agree on nontrivial posteriors for E is possible, if and only if there is a nonempty *finite* event F at which it is common knowledge that the agents are ignorant of E , and moreover, this holds true at F also after F becomes common knowledge.

Using this characterization we can see why agreeing to agree is impossible in Example 3. For this we need to show that there is no nonempty finite event F that satisfies the required condition. Indeed, suppose that there exists such an event F . Let i_m be the maximal profit of 1 in all the states of F . By the property of F , 1 is still ignorant of E

after being informed of F . In particular, 1 cannot tell that $\neg E$ is the case. Therefore, there must exist a point (i_m, j_0) in F which is also in E . By the definition of E , $i_m < j_0$. By the definition of i_m , for each point $(i, j_0) \in F$, $i \leq i_m$ and hence $i < j_0$. Thus, $(i, j_0) \in E$. But this means that when firm 2's profit is j_0 , and it is informed of F , it knows E . This contradicts the assumption that it is ignorant of E .

Another characterization of the possibility of agreeing to agree is given in probabilistic terms.

Agreeing to agree about nontrivial posteriors for E is possible, if and only if it is possible with a common prior that has a finite support.

While firms cannot agree on the probability that one is more profitable than the other, as shown in the previous example, firms can still agree on the probability that they are equally profitable. To see this, consider the following example.

Example 4. In the same state space as in Example 3, define E to be the event where both firms are equally profitable. The conditions for the possibility of agreeing to agree hold in this case. Indeed, consider the finite set $F = \{(1, 1), (1, 0), (0, 1), (0, 0)\}$. It is common knowledge at every state, and in particular at every state of F , that the agents are ignorant of E . Second, when this finite set becomes common knowledge, the firms are still ignorant of E , and moreover, their ignorance is also common knowledge.

Using the set F , it is easy to construct posteriors with a common prior for which agreeing to agree holds for E . Consider the common prior μ , where $\mu(1, 1) = \mu(1, 0) = \mu(0, 1) = \mu(0, 0) = \frac{1}{4}$. When a firm's profit is either 0 or 1, the posterior for E is a half. Since any information set of a firm whose profit is neither 0 nor 1, has probability zero according to μ , one may define the posteriors in these elements in an arbitrary fashion. In particular one may define these posteriors for E to be $\frac{1}{2}$. Defining the posteriors in this way makes the fact that the posteriors for E are $\frac{1}{2}$, common knowledge at each state.

1.4. Positive common priors. In the previous example agreeing to agree is made possible by a degenerate common prior. It vanishes on all but two elements in each partition. Our next result gives a necessary and sufficient condition for the possibility of agreeing to agree with a positive common prior. That is, a common prior that assigns a positive probability to every information set.

For any given event E we define simple-structured, nonempty and finite events called *loops on E* (see Definition 4 and Figure 1). One of the conditions that define a loop implies that when a loop on E

becomes common knowledge, it is common knowledge at the loop that the agents are ignorant of E . This property links the notion of loops to the condition in our main result.

When the information set of an agent intersects a loop we say that this agent does not exclude loops. Our third result states that

Agreeing to agree is possible with a common prior which is positive on all information sets, if and only if it is common knowledge at some state that the agents do not exclude loops.

1.5. The role of finiteness. The less obvious part of our characterizations is the finiteness. It has to do with the way E and $\neg E$ are entangled together in the information structure. When there is *no* finite F with the required property, it is possible to separate E and $\neg E$ by arranging the partitions in a way that resembles the structure of the events in Example 3. When there *is* such F , the sets E and $\neg E$ are “mixed” together like the ones in Examples 1 and 4. The separation is described in Proposition 3. In Proposition 4 it is shown why this separation implies the nonexistence of posteriors that makes agreeing to agree possible. This proposition generalizes an argument made in Samet et al. (2002) for the special case of Example 3, which is also related to an intriguing puzzle due to Cover (1987).

1.6. Agreement theorems. The literature abounds with generalizations of Aumann’s agreement theorem. In this theorem the disagreement concerns the posterior of an event. Milgrom and Stokey (1982), Sebenius and Geanakoplos (1983), and Rubinstein and Wolinsky (1990) studied agreement theorems where the disagreement concerns the value or expected value of random variables.

Aumann’s theorem provides a necessary condition for the existence of a common prior. Feinberg (1995), Morris (1995), Bonanno and Nehring (1996), and Samet (1998) showed that when disagreements are extended to random variables, then the impossibility of agreeing to disagree is a necessary condition for the existence of a common prior.

McKelvey and Page (1986) and Nielsen et al. (1990) showed that when the agents have different posteriors, then not only these posteriors cannot be common knowledge, but also certain aggregates of the posteriors cannot be common knowledge.

Geanakoplos and Polemarchakis (1982), Sebenius and Geanakoplos (1983), Parikh and Krasucki (1990), and Heifetz (1996) studied the dynamics of information exchange that leads from disagreement to agreement.

In Bacharach (1985) and Samet (1990), Aumann’s theorem is generalized to a non-partition information structure. In Monderer and Samet (1989), common p -belief, rather than common knowledge, is shown to imply a bound on the probabilistic disagreement.

All the above mentioned agreement theorems study the impossibility of commonly sharing probabilistic *disagreements*. This paper, in contrast, studies the impossibility of common knowledge of probabilistic *agreements*. As it turns out, the impossibility of having common knowledge of agents’ posterior for an event may be the result not of the probabilistic disagreement, but rather the structure of the event. While the starting point of the literature on agreement theorems is the probabilistic structure juxtaposed with the knowledge structure, here it is the knowledge structure alone. We look for conditions that characterize the events for which there exists a probabilistic structure that makes it possible to have common knowledge of agreements regarding the posterior of these events.

1.7. The paper plan. In the next section, we introduce the basics of the model of knowledge and belief. The main elements of our agreeing theorems, the *ignorance*, and the *possibility of agreeing to agree* are defined in Section 3, and the main results are stated. A counter example for uncountable information structures is presented in Section 4. Two open problems are discussed in Section 5, and the proofs are in the last section.

2. PRELIMINARIES

2.1. Information structures. We fix a *state space* (Ω, \mathcal{B}) , with a set of *states* Ω and a σ -field of *events* \mathcal{B} . An *information structure* on the state space, for agents 1 and 2, is a pair (Π_1, Π_2) of countable measurable partitions of Ω . For every state ω we denote by $\Pi_i(\omega)$ the element in Π_i which contains ω .

For an event F and agent i we denote by Π_i^F the partition $\Pi_i \wedge \{F, \neg F\}$, which is the join of the partitions Π_i and $\{F, \neg F\}$.¹ It describes the information of the agent when, in addition to the information given by Π_i , she is also informed whether or not F occurred.

2.2. Knowledge and common knowledge. We say that agent i *knows* event E at ω when $\Pi_i(\omega) \subseteq E$. Thus, the event that i knows E is $K_i(E) = \{\omega \mid \Pi_i(\omega) \subseteq E\}$. The event that both agents know E is $K(E) = K_1(E) \cap K_2(E)$. The event that E is *common knowledge* is $K^\infty(E) = \bigcap_{n=1}^\infty K^n(E)$. It is the union of all the elements of the

¹The *join* of two partitions is their finest common coarsening.

join $\Pi_1 \wedge \Pi_2$ contained in E . For further discussion of knowledge and common knowledge see Aumann (1976), Monderer and Samet (1989), Geanakoplos (1994), Fagin et al. (1995), and Aumann (1999).

Note that the epistemic operators K_i, K , and K^∞ are defined with respect to the information structure (Π_1, Π_2) . In the sequel we consider also the information structure (Π_1^F, Π_2^F) on the same state space. We will make it clear when we use the epistemic operators with respect to² this latter information structure.

The event F belongs to the join $\Pi_1^F \wedge \Pi_2^F$. Thus, in the information structure (Π_1^F, Π_2^F) , $F = K^\infty(F)$, and $\neg F = K^\infty(\neg F)$. Indeed, this information structure can be defined as the one obtained from making F and $\neg F$ common knowledge.

2.3. Posteriors and priors. A *posterior* of agent i , for the information structure (Π_1, Π_2) , is a family of probability measures $\{\mu_i^\omega\}$ on (Ω, \mathcal{B}) for $i = 1, 2$ and each ω , such that $\mu_i^\omega(\Pi_i(\omega)) = 1$ and for each $\omega' \in \Pi_i(\omega)$, $\mu_i^{\omega'} = \mu_i^\omega$. A *common prior* for the posteriors of the agents is a probability measure μ on (Ω, \mathcal{B}) such that $\mu_i^\omega(\cdot) = \mu(\cdot \mid \Pi_i(\omega))$ whenever $\mu(\Pi_i(\omega)) > 0$.

3. AGREEING TO AGREE: DEFINITIONS AND THE MAIN RESULTS

3.1. Agreements. An agreement is an event that the posteriors of the agents coincide at some given probability. Formally,

Definition 1. Let E be an event in the state space (Ω, \mathcal{B}) with information structure (Π_1, Π_2) and posteriors $\{\mu_i^\omega\}$. An **agreement on E** is an event

$$(1) \quad A = \{\omega \mid \mu_1^\omega(E) = \mu_2^\omega(E) = p\},$$

for some $0 < p < 1$.

Note that this definition requires that the agreed upon posteriors be strictly between 0 and 1.

Next we formally define the possibility that there is common knowledge of an agreement.

Definition 2. Let E be an event in the state space (Ω, \mathcal{B}) with information structure (Π_1, Π_2) . **Agreeing to agree is possible for E** if there are posteriors with a common prior μ , and an agreement A on E for these posteriors, such that $\mu(K^\infty(A)) > 0$.

²From here on, w.r.t.

Definition 3. *The event that agent i is ignorant of E is*

$$I_i(E) = \neg K_i(E) \cap \neg K_i(\neg E).$$

The event that both agents are ignorant of E is $I(E) = I_1(E) \cap I_2(E)$.

Thus, $I(E)$ is the event that none of the agents can tell whether or not E occurred.³

3.2. The main results. Our first theorem states that agreeing to agree is possible for E if and only if there exists a nonempty finite event F , at which it is common knowledge that both agents are ignorant of E . Moreover, this holds true also after F becomes common knowledge. Formally and succinctly,

Theorem 1. *Agreeing to agree is possible for event E iff there exists a nonempty finite event F such that*

$$(2) \quad F \subseteq K^\infty(I(E))$$

w.r.t. both (Π_1, Π_2) and (Π_1^F, Π_2^F) .

The fact that at F it is common knowledge that both agents are ignorant of E is expressed by $F \subseteq K^\infty(I(E))$, w.r.t. (Π_1, Π_2) . The fact that it remains common knowledge that both agents are ignorant of E also after F becomes common knowledge is conveyed by $F \subseteq K^\infty(I(E))$, w.r.t. (Π_1^F, Π_2^F) . When the information structure is finite, the condition for agreeing to agree becomes simpler.

Corollary. *If the information structure is finite, then agreeing to agree is possible for E , iff*

$$(3) \quad K^\infty(I(E)) \neq \emptyset.$$

To see this, note first that since $F \neq \emptyset$, (2) implies $K^\infty(I(E)) \neq \emptyset$. On the other hand, if (3) holds, define $F = K^\infty(I(E))$. F is obviously nonempty and finite. Moreover, since F is an element of the join partition, $\Pi_1 \wedge \Pi_2$, it follows that $\Pi_i^F = \Pi_i$ for $i = 1, 2$. Therefore, (2) holds w.r.t. both, (Π_1, Π_2) and (Π_1^F, Π_2^F) .

Notice that the condition for agreeing to agree in Theorem 1 is formulated solely in terms of the information structure. The next theorem provides a condition in probabilistic terms.

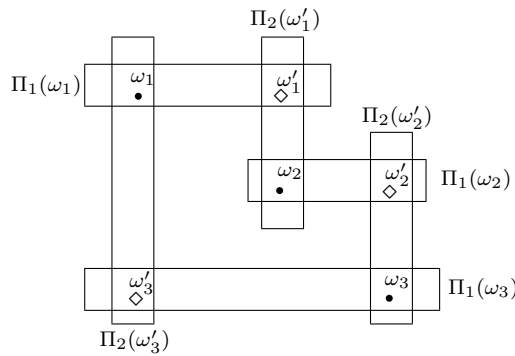
Theorem 2. *Agreeing to agree is possible for E iff it is possible with a common prior that has a finite support.*

³The negation of the operator I_i is the *knowing whether* operator, $J_i(E) = K_i(E) \cup K_i(\neg E)$, the properties of which were studied in Hart et al. (1996).

3.3. Positive priors. A common prior determines the posteriors only on elements of the partition to which it assigns positive probability; on elements of probability zero the posteriors can be defined arbitrarily. Of special interest are common priors that determine uniquely the posteriors, that is, common priors that assign a positive probability to all the elements in both partitions. We characterize next, events for which agreeing to agree is possible with such a common prior. This characterization is done in terms of events called *loops*, which play a central role also in the proof of Theorem 1.

Definition 4. A loop on E is an event L which consists of $2n$ distinct points $\omega_1, \omega'_1, \dots, \omega_n, \omega'_n$ for some integer $n \geq 1$, such that

1. $\omega_k \in E$ and $\omega'_k \in \neg E$ for any $k = 1, \dots, n$; and
2. $\omega'_k \in \Pi_1(\omega_k)$ and $\omega_k \in \Pi_2(\omega'_{k-1})$ for any $k = 1, \dots, n$, where $\omega'_0 = \omega'_n$.



A loop on E with 6 points. The dot shaped states, $\omega_1, \omega_2, \omega_3$, are in E , and the diamond shaped states, $\omega'_1, \omega'_2, \omega'_3$, are in $\neg E$.

FIGURE 1. A loop on E

The following two propositions relate loops to events F that satisfy (2).

Proposition 1. If F is a nonempty finite event such that $F \subseteq K^\infty(I(E))$ w.r.t. (Π_1^F, Π_2^F) , then F contains a loop on E .

Proposition 2. If L is a loop on E , then $L \subseteq K^\infty(I(E))$ w.r.t. (Π_1^L, Π_2^L) .

In the next theorem we state formally that agreeing to agree is possible for E with a common prior that assigns positive probability to each element of each partition, iff it is common knowledge that none of the players excludes the possibility of loops.

Theorem 3. *Agreeing to agree is possible for E with a common prior that assigns a positive probability to each element of Π_1 or Π_2 , iff $K^\infty(\neg K_1(\neg \hat{L}) \cap \neg K_2(\neg \hat{L})) \neq \emptyset$, where \hat{L} is the union of all loops on E .*

4. UNCOUNTABLE INFORMATION STRUCTURES

So far we have assumed that the partitions are countable. In order to study the question of agreeing to agree in models with uncountable partitions we need first to define in such models posteriors and common priors.

4.1. Posteriors and common prior for the uncountable case.

Posteriors $\{\mu_i^\omega\}$ are defined as in the countable case with an additional measurability requirement: $\mu_i^\omega(E)$ is measurable as a function of ω , for any fixed event E and $i = 1, 2$. Equivalently, for $i = 1, 2$, any fixed event E and p , $\{\omega \mid \mu_i^\omega(E) \geq p\}$ is a measurable set.

A probability measure μ is a common prior for these posteriors if for $i = 1, 2$ and each event E ,

$$(4) \quad \mu(E) = \int \mu_i^\omega(E) d\mu(\omega).$$

In particular, using the properties of posteriors,

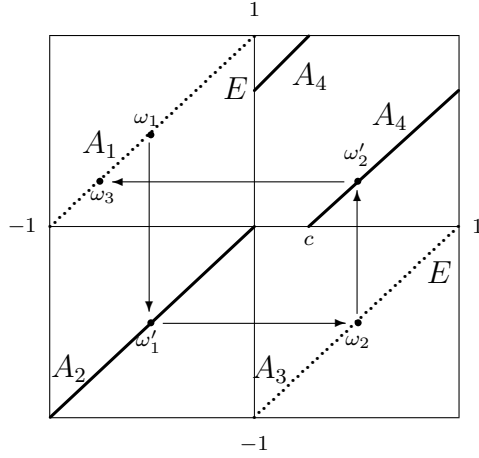
$$\mu(\Pi_i(\omega_0) \cap E) = \int_{\Pi_i(\omega_0)} \mu_i^\omega(E) d\mu(\omega) = \mu_i^{\omega_0}(E) \mu(\Pi_i(\omega_0)).$$

This implies that if $\mu(\Pi_i(\omega_0)) > 0$ then $\mu_i^\omega(E) = \mu(E \mid \Pi_i(\omega))$, which shows that (4) indeed generalizes the definition of a common prior in the countable partitions case.⁴

4.2. Theorem 1 and the uncountable case. The proof that the condition in Theorem 1 is sufficient does not make use of the countability of the information structure, and therefore it holds also for uncountable information structures. The proof that the condition is necessary relies on the countability assumption. As shown by the following example this assumption is essential: the condition is not necessary in the uncountable case.

Example 5. Consider a state space Ω that consists of the union of the following four disjoint sets A_i , $i = 1, \dots, 4$, in \mathbb{R}^2 : $A_1 = \{(x, x+1) \mid -1 < x < 0\}$, $A_2 = \{(x, x) \mid -1 < x < 0\}$, $A_3 = \{(x, x-1) \mid 0 < x < 1\}$ and $A_4 = \{(x, \varphi(x)) \mid 0 < x < 1\}$, where $\varphi(x) = x - c \pmod{1}$ for a

⁴See Samet (1999) and Samet (2000) for studies of belief systems with uncountable information structures.



The state space consists of the three diagonals A_1, A_2, A_3 and of A_4 . The latter is obtained by a rightward shift of the top-right diagonal by an irrational number c . Agent 1 is informed of the first coordinate of the state, and agent 2 of the second. The event E is the union of A_1 and A_3 . There are no loops on E and nevertheless agreeing to agree is *possible* with the common prior which is uniform on the sets A_i .

FIGURE 2. Theorem 1 fails in the uncountable case

fixed irrational c in $(0, 1)$. The σ -field on Ω is the Borel σ -field on the sets A_i .

Agent 1 is informed of the first coordinate of the state, and agent 2 of the second. Thus, each element of Π_1 contains the two points on the vertical line that contains the state. Similarly, Π_2 contains the two points on the horizontal line that includes the state.

The posterior μ_i^ω for each of the two points in $\Pi_i(\omega)$ is $\frac{1}{2}$. To see that the measurability condition is satisfied consider $P: \Omega \rightarrow \mathbb{R}$, the projection of the state space on the first coordinate. Note that both P and P^{-1} map measurable sets to measurable sets. Now, if $p \in (\frac{1}{2}, 1]$ and E is a measurable set, then $\{\omega \mid \mu_1^\omega(E) \geq p\} = P^{-1}(P(E \cap A_1) \cap P(E \cap A_2)) \cup P^{-1}(P(E \cap A_3) \cap P(E \cap A_4))$, which is measurable. On the other hand, if $p \in (0, \frac{1}{2}]$, then $\{\omega \mid \mu_1^\omega(E) \geq p\} = P^{-1}(P(E \cap A_1) \cup P(E \cap A_2)) \cup P^{-1}(P(E \cap A_3) \cup P(E \cap A_4))$, which is also measurable. A similar argument shows that the measurability condition holds for the second agent's posterior as well.

Let μ be the probability measure $\frac{1}{4} \sum_{i=1}^4 \mu_i$, where μ_i is the Lebesgue measure on A_i . It is easy to see that μ is a common prior for $\{\mu_i^\omega\}$.

Denote $E = A_1 \cup A_3$. For $i = 1, 2$ and every ω , $\mu_i^\omega(E) = \frac{1}{2}$. Thus agreeing to agree is possible for E . We show that there is no loop for E and therefore by Proposition 1, Theorem 1 does not hold in this case.

Suppose that $\omega_1 = (a - 1, a) \in A_1 \subseteq E$ is the first point in a loop. Then ω'_1 is necessarily $(a - 1, a - 1) \in A_2 \subseteq \neg E$. The next two points must be $\omega_2 = (a, a - 1) \in A_3$ and $\omega'_2 = (a, \varphi(a)) \in A_4$. Thus $\omega_3 = (\varphi(a) - 1, \varphi(a)) \in A_1$. Continuing this way, we obtain that the second coordinates of the points $\{\omega_{2k+1}\}_{k \geq 0}$ are $\{\varphi^k(a)\}_{k \geq 0}$. Since c is irrational, all the numbers in the latter sequence are distinct, and therefore so are all the points in the first sequence. A similar argument shows that each of the sequences $\{\omega'_{2k+1}\}_{k \geq 0}$, $\{\omega_{2k}\}_{k \geq 1}$, and $\{\omega'_{2k}\}_{k \geq 1}$ also consists of distinct points. These four sequences are disjoint as they belong to different sets A_i . Thus, the sequence starting with $\omega_1 \in A_1$ is infinite. This also shows that any sequence starting with $\omega_1 \in A_3$ is also infinite, and therefore there cannot be a loop L on E .

4.3. The pathology in this example. Each element of the join $\Pi_1 \wedge \Pi_2$ can be written as a doubly infinite countable sequence

$$\dots, \omega_{-k}, \omega'_{-k}, \dots, \omega_1, \omega'_1, \dots, \omega_k, \omega'_k, \dots,$$

such that $(\{\omega_k, \omega'_k\})_{k=-\infty}^{\infty} \subseteq \Pi_1$ and $(\{\omega'_k, \omega_{k+1}\})_{k=-\infty}^{\infty} \subseteq \Pi_2$. This sequence is constructed in both directions in the same way as the sequence above. On such a sequence, if considered a subspace, there is no common prior. Indeed, a common prior ν should satisfy for each k , $\nu(\omega_k) = \nu(\omega'_k)$ and also $\nu(\omega'_k) = \nu(\omega_{k+1})$. Thus, all the states must have the same probability, which is impossible.

The pathology in this example concerns the question of the existence of a common prior for given posteriors. On one hand, there exists a common prior since the measure μ satisfies (4). On the other hand, for any element of the join $\Pi_1 \wedge \Pi_2$ there is no common prior. It seems that the subspace relevant to the agents at a given state is the event of the join containing this state. This is the smallest event that is commonly known by the agents at this state. In our space, at any state there is no common prior on the relevant subspace containing it.

This example raises conceptual difficulties concerning the notion of a common prior in the case of uncountable information structures. The problem was first observed (in a different state space) and discussed in Simon (1998).

5. OPEN PROBLEMS

5.1. Uncountable information structures. The necessity of the condition in Theorem 1 is proved by negation. We show that if agreeing

to agree holds for E and the condition fails, then the prior probability of each element of Π_1 is zero. This constitutes a contradiction because the partition is countable. Here we used the countability of only one of the partitions, but as we explain next, the proof requires that both are countable.

Propositions 3 and 4 are the two main parts of the proof that the condition in Theorem 1 is necessary. In the proof of the first one we define a binary relationship on Π_1 . Using this relation we coarsen both partitions by grouping together the elements of the partitions. The second proposition uses this grouping to show that elements of Π_1 that do not intersect a loop must be assigned zero probability by the common prior.

The grouping of the elements of the partitions in the proof of Propositions 3 does not require countable partitions. However, the countability of the partitions is used to establish that the grouped elements are *measurable* sets. Although the statement of Proposition 4 refers solely to elements in Π_1 , the proof makes use also of the measurability of events resulting from the grouping procedure over Π_2 . These events are measurable due to countability. Therefore, our proofs hinge on the countability of *both* partitions.

The countability assumption in Theorem 1 is essential, because, as has been shown in the previous section, the condition in the theorem is not necessary in the uncountable case. It has been noted, though, that the common prior which enables the agreeing to agree in the previous section is pathological in the sense that it fails to induce common priors on the elements of the join of the partitions. This suggests the following problem.

Open problem 1. Consider a state space with an *uncountable* information structure. Assume that there is a common prior *which induces a common prior on each non-null element of the join of the partitions*.⁵ Does this imply that agreeing to disagree is possible if and only if the condition in Theorem 1 is satisfied?

5.2. More than two agents. All our results were formulated and proved for models with two agents. Generalizing the notion of a loop to many agents may enable a generalization of the sufficiency part of Theorem 1 to more than two agents. However, the proof of the necessity part strongly depends on having two agents.

⁵This is the case, for instance, when the whole space is the only element of the join.

The binary relationship employed in the proof of Proposition 3, as well as the grouping of the elements of the partitions, is defined only for two agents. We know of no way to generalize it to more than two agents.

Open problem 2. Find a necessary and sufficient condition in terms of the information structure for the possibility of agreeing to agree in models with more than two agents.

6. PROOFS

Proof of Proposition 1. Let F be a nonempty event that satisfies (2) w.r.t. the information structure (Π_1^F, Π_2^F) . For each $\omega \in F$ and $i = 1, 2$, we have $\Pi_i^F(\omega) \subseteq F \subseteq K^\infty(I(E)) \subseteq I(E)$ (where K^∞ and I are considered as operators w.r.t. (Π_1^F, Π_2^F)). Hence, $\Pi_i^F(\omega) \cap E \neq \emptyset$ and $\Pi_i^F(\omega) \cap \neg E \neq \emptyset$, $i = 1, 2$. Therefore, there is a sequence⁶ $\omega_1, \omega'_1, \dots, \omega_k, \omega'_k, \dots$ in F such that $\omega_k \in E$, $\omega'_k \in \neg E$, $\omega'_k \in \Pi_1^F(\omega_k)$ and $\omega_{k+1} \in \Pi_2^F(\omega'_k)$, for every $k \geq 1$. Since for each $\omega \in F$ and $i = 1, 2$, $\Pi_i^F(\omega) = \Pi_i(\omega) \cap F$, it follows that $\omega'_k \in \Pi_1(\omega_k)$ and $\omega_{k+1} \in \Pi_2(\omega'_k)$, for every $k \geq 1$.

Note that since $\omega_k \in E$ and $\omega'_k \in \neg E$, $\omega_k \neq \omega'_k$. Thus, as F is finite the sequence must contain a subsequence $\omega_{k+1}, \omega'_{k+1}, \dots, \omega_{k+n}, \omega'_{k+n}$ of $2n$ distinct points, for $n \geq 1$, such that $\omega_{k+n+1} = \omega_{k+1}$. This sequence is a loop on E . ■

Proof of Proposition 2. Consider a loop on E ,

$$L = \{\omega_1, \omega'_1, \dots, \omega_n, \omega'_n\}.$$

By the first condition of Definition 4, $L \subseteq I(E)$. By the second condition, some coarsening of $\{\{\omega_1, \omega'_1\}, \dots, \{\omega_n, \omega'_n\}\}$ is included in Π_1^F and some coarsening of $\{\{\omega'_1, \omega_2\}, \dots, \{\omega'_n, \omega_1\}\}$ is included in Π_2^F . Therefore, $L \subseteq \Pi_1 \wedge \Pi_2$. Thus, $L \subseteq K^\infty(I(E))$ w.r.t. (Π_1^F, Π_2^F) . ■

The proof of Theorem 1 requires the following two propositions.

Proposition 3. *Let $P_0 \in \Pi_1$. Suppose that for every loop L on E , $L \cap P_0 = \emptyset$. Then, there is a coarsening of Π_1 , $\{P_E, P_{\neg E}, P_0\}$, and a coarsening of Π_2 , $\{Q_E, Q_{\neg E}\}$, such that⁷*

$$(5) \quad (P_E \cup P_0) \cap Q_E \subseteq E$$

and

$$(6) \quad (P_{\neg E} \cup P_0) \cap Q_{\neg E} \subseteq \neg E.$$

⁶The sequence may be constructed inductively.

⁷We allow each of the sets $P_E, P_{\neg E}, Q_E, Q_{\neg E}$ to be empty.

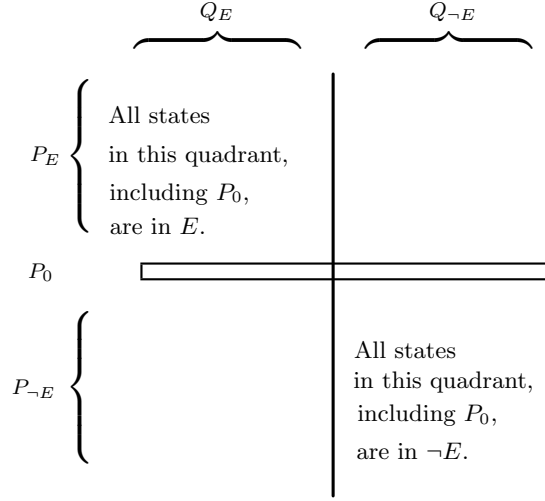
Proof. In this proof P and Q denote generic elements of Π_1 and Π_2 , respectively. We define a binary relation ρ on Π_1 : $P \rho P'$ if for some Q , $P \cap Q \cap E \neq \emptyset$ and $P' \cap Q \cap \neg E \neq \emptyset$. We denote by \succ the transitive closure⁸ of ρ .

Obviously, \succ is transitive. We show that $P_0 \not\succeq P_0$. Suppose to the contrary that $P_0 \succ P_0$. Then there is a sequence P_0, \dots, P_n , such that

$$(7) \quad P_0 = P_n \rho \dots \rho P_1 \rho P_0.$$

Moreover, by taking the shortest sequence of this kind, we can assume that P_1, \dots, P_n are distinct. By (7), there are Q_0, \dots, Q_{n-1} in Π_2 , such that for $k = 1, \dots, n$, $P_k \cap Q_{k-1} \cap E \neq \emptyset$ and $P_{k-1} \cap Q_{k-1} \cap \neg E \neq \emptyset$.

For each $k = 1, \dots, n$ choose $\omega_k \in P_k \cap Q_{k-1} \cap E$ and $\omega'_k \in Q_k \cap P_k \cap \neg E$, where $Q_n = Q_0$. All $2n$ points are distinct. Indeed, $\omega_k \neq \omega'_k$ because $\omega_k \in E$ and $\omega'_k \in \neg E$. Moreover, since $\{\omega_k, \omega'_k\} \subseteq P_k$ and P_1, \dots, P_n are all distinct, each pair $\{\omega_k, \omega'_k\}$ belongs to a different element of Π_1 . Thus, the points $\omega_1, \omega'_1, \dots, \omega_n, \omega'_n$ form a loop on E which intersects P_0 at ω_n . This contradicts the assumption, and we conclude that $P_0 \not\succeq P_0$.



The elements of Π_1 are horizontal, those of Π_2 are vertical.

FIGURE 3. An illustration for Proposition 3

Define P_E to be the union of all elements P that satisfy $P \succ P_0$. Since $P_0 \not\succeq P_0$ it follows that $P_0 \cap P_E = \emptyset$. The set P_{-E} is the complementary

⁸That is, $P \succ P'$ iff there are P_1, \dots, P_n in Π_1 , with $n \geq 2$, such that $P_1 = P$, $P_n = P'$ and for $j = 1, \dots, n-1$, $P_j \rho P_{j+1}$.

set of $P_E \cup P_0$. An element Q is a subset of $Q_{\neg E}$ iff there exists P , such that $P \subseteq P_E \cup P_0$ and $P \cap Q \cap \neg E \neq \emptyset$. The set Q_E is the complement of $Q_{\neg E}$.

By the definition of Q_E , for each $Q \subseteq Q_E$ and $P \subseteq P_E \cup P_0$, $P \cap Q \cap \neg E = \emptyset$ which proves (5). To prove (6), suppose to the contrary that for some $P \subseteq P_{\neg E} \cup P_0$ and $Q \subseteq Q_{\neg E}$, $P \cap Q \cap E \neq \emptyset$. By the definition of $Q_{\neg E}$ there is $P' \subseteq P_E \cup P_0$ such that $P' \cap Q \cap \neg E \neq \emptyset$. Thus, $P \succ P'$.

Now, if $P' = P_0$, then $P \succ P_0$. Else, $P' \subseteq P_E$, in which case $P \succ P' \succ P_0$. Thus, in either case $P \succ P_0$ which implies that $P \subseteq P_E$. This contradicts the assumption that $P \subseteq P_{\neg E} \cup P_0$ and the proof of (6) is also complete. ■

Proposition 4. *Let E be an event, $\{\mu_i^\omega\}$ be posteriors with a common prior μ , and A be the agreement $\{\omega \mid \mu_1^\omega(E) = \mu_2^\omega(E) = p\}$, for some $0 < p < 1$. If $P_0 \in \Pi_1$, $P_0 \subseteq K^\infty(A)$ and $P_0 \cap L = \emptyset$ for all loops L on E , then $\mu(P_0) = 0$.*

Proof. In case $\mu(K^\infty(A)) = 0$, the proof is complete since $P_0 \subseteq K^\infty(A)$. Otherwise, define $\nu(\cdot) = \mu(\cdot \mid K^\infty(A))$. Note that ν is also a common prior for the posteriors⁹ $\{\mu_i^\omega\}$. Moreover, if $\nu(\Pi_i(\omega)) > 0$, then $\mu_i^\omega(E) = p$.

By Proposition 3 there are partitions $\{P_E, P_{\neg E}, P_0\}$ and $\{Q_E, Q_{\neg E}\}$ that satisfy (5) and (6). The events P_E and $Q_{\neg E}$ are unions of elements of Π_1 and Π_2 respectively, and since the partitions are countable, these events are measurable. From the definitions of A and ν it follows that $\nu(E \cap P_E) = p\nu(P_E)$ and $\nu(\neg E \cap P_E) = (1 - p)\nu(P_E)$. Thus (recall, $1 - p > 0$),

$$(8) \quad \nu(E \cap P_E) = (p/(1 - p))\nu(\neg E \cap P_E).$$

By a similar argument,

$$(9) \quad \nu(E \cap Q_{\neg E}) = (p/(1 - p))\nu(\neg E \cap Q_{\neg E}).$$

By (6) the relation between the events on the left sides of (8) and (9) is $E \cap Q_{\neg E} \subseteq E \cap P_E$. Hence, by comparing the right sides of (8) and (9) we obtain (after dividing by $p/(1 - p) \neq 0$),

$$\nu(\neg E \cap P_E) \geq \nu(\neg E \cap Q_{\neg E}).$$

⁹This is so since $K^\infty(A)$ is in the join of the partitions. Therefore, $\nu(\cdot)$ induces over $K^\infty(A)$ the same posteriors as μ , while on the complement of $K^\infty(A)$, the posteriors can be defined so as to coincide with $\{\mu_i^\omega\}$.

By (5) the event on the left side is a subset of the event on the right side. Thus their difference must have probability 0. This difference is

$$\begin{aligned} (\neg E \cap Q_{\neg E}) \setminus (\neg E \cap P_E) &= \neg E \cap (Q_{\neg E} \setminus P_E) \\ &= \neg E \cap (Q_{\neg E} \cap (P_{\neg E} \cup P_0)) \\ &= Q_{\neg E} \cap (P_{\neg E} \cup P_0) \end{aligned}$$

where the last equality holds by (6).

We conclude that $\nu(Q_{\neg E} \cap (P_{\neg E} \cup P_0)) = 0$ and in particular $\nu(Q_{\neg E} \cap P_0) = 0$. By (5), $Q_{\neg E} \cap P_0 \subseteq \neg E$. Hence, $Q_{\neg E} \cap P_0 = \neg E \cap P_0$. Thus, $\nu(\neg E \cap P_0) = 0$ and therefore $\mu(\neg E \cap P_0) = 0$. However, by assumption $P_0 \subseteq K^\infty(A)$, and therefore, for every $\omega \in P_0$, $\mu_1^\omega(\neg E) = 1 - p > 0$. This is possible only when $\mu(P_0) = 0$. ■

Proof of Theorem 1. Suppose that a nonempty finite F is a subset of $K^\infty(I(E))$ w.r.t. both (Π_1, Π_2) and (Π_1^F, Π_2^F) . Thus, $K^\infty(I(E))$ w.r.t. (Π_1, Π_2) is nonempty.

From now on (in this proof), the operators K^∞ and I are used w.r.t. (Π_1, Π_2) . We construct posteriors $\{\mu_i^\omega\}$ with a common prior μ such that $\mu(K^\infty(I(E))) > 0$ and $K^\infty(I(E)) \subseteq A$, where $A = \{\omega \mid \mu_1^\omega(E) = \mu_2^\omega(E) = \frac{1}{2}\}$. This inclusion implies $K^\infty(I(E)) \subseteq K^\infty(A)$, which shows that $\mu(K^\infty(A)) > 0$ and thereby establishes the fact that agreeing to agree is possible for E .

By Propositions 1 and 2, there exists a loop $L = \{\omega_1, \omega'_1, \dots, \omega_n, \omega'_n\}$ on E such that $L \subseteq K^\infty(I(E))$. Define a common prior μ such that $\mu(\omega_k) = \mu(\omega'_k) = \frac{1}{2n}$ for $k = 1, \dots, n$. Consider $P \in \Pi_1$ such that $P \subseteq K^\infty(I(E))$.

Suppose that $P \cap L \neq \emptyset$. If $\omega_k \in P$, then also $\omega'_k \in P$. Since all the points in L are distinct, it follows that P contains the same number of points in $E \cap F$ and in $\neg E \cap F$. Thus for $\omega \in P$, $\mu_1^\omega(E) = \frac{1}{2}$.

If $P \cap L = \emptyset$, then $\mu(P) = 0$ and therefore the posteriors on P can be defined arbitrarily. Since $P \subseteq K^\infty(I(E))$ it follows that $P \subseteq I(E)$ and therefore $P \cap E \neq \emptyset$ and $P \cap \neg E \neq \emptyset$. We define the posterior μ_1^ω for $\omega \in P$ such that $\mu_1^\omega(E) = \mu_1^\omega(\neg E) = \frac{1}{2}$.

A similar argument holds for elements $Q \in \Pi_2$. This shows that $K^\infty(I(E)) \subseteq A$, as required.

The prove of the converse direction makes use of Propositions 3 and 4. Suppose that agreeing to agree is possible for E for posteriors with a common prior μ , and the agreement A (recall (1)). It means that

$$(10) \quad \mu(K^\infty(A)) > 0.$$

Obviously, $A \subseteq I(E)$ and therefore

$$(11) \quad K^\infty(A) \subseteq K^\infty(I(E)).$$

By (10) and (11) $K^\infty(I(E))$ is not empty.

Since $K^\infty(I(E))$ is a union of elements of the join $\Pi_1 \wedge \Pi_2$, it follows that for a loop L ,

$$(12) \quad \text{if } L \cap K^\infty(I(E)) \neq \emptyset \text{ then } L \subseteq K^\infty(I(E)).$$

Assume that there exists a loop L on E such that $L \subseteq K^\infty(I(E))$. By Proposition 2, $F = L$ is a nonempty finite set that satisfies (2), and the proof is complete. Otherwise, for every loop L on E , $L \not\subseteq K^\infty(I(E))$, which implies by (12) that $L \cap K^\infty(I(E)) = \emptyset$. Thus, by (11), $L \cap K^\infty(A) = \emptyset$. Therefore, by Proposition 4, for each $P_0 \in \Pi_1$ such that $P_0 \subseteq K^\infty(A)$, $\mu(P_0) = 0$. Since the partitions are countable, this implies that $\mu(K^\infty(A)) = 0$, which contradicts (10). ■

Proof of Theorem 2. If agreeing to agree is possible for E then the condition of Theorem 1 holds. As was shown in the first part of its proof, this condition implies that agreeing to agree is made possible by a common prior supported on a loop, which is a finite event. ■

Proof of Theorem 3. Assume first that the event

$$\hat{\Omega} = K^\infty(\neg K_1(\neg \hat{L}) \cap \neg K_2(\neg \hat{L}))$$

is not empty. Then, for each $P \in \Pi_i$, $i = 1, 2$, if $P \subseteq \hat{\Omega}$ then $P \subseteq \neg K_1(\neg \hat{L}) \cap \neg K_2(\neg \hat{L})$ and thus $P \cap \hat{L} \neq \emptyset$.

Since $\hat{\Omega}$ is a union of elements of the join of $\Pi_1 \wedge \Pi_2$, it follows that for a loop L , if $L \cap \hat{\Omega} \neq \emptyset$, then $L \subseteq \hat{\Omega}$. Let $\{L_1, L_2, \dots\}$ be the set of all the loops contained in $\hat{\Omega}$. For each n let μ_n be a probability distribution on L_n such that for each $\omega \in L_n$ $\mu_n(\omega) = 1/(2|L_n|)$.¹⁰ Let $\mu = \sum 2^{-n} \mu_n$. Then, for $P \in \Pi_i$, $i = 1, 2$, $\mu(P \cap E) = \sum 2^{-n} \mu_n(P \cap E) = \sum 2^{-n} (1/2) \mu_n(P) = (1/2) \mu(P)$.

Since for each $P \in \Pi_i$ which is a subset of $\hat{\Omega}$, $P \cap \hat{L} \neq \emptyset$, it follows that $\mu(P) > 0$. In addition, μ vanishes on the complement of $\hat{\Omega}$. Let μ' be any measure on Ω which vanishes on $\hat{\Omega}$ and is positive on the elements of Π_1 and Π_2 which are not subsets of $\hat{\Omega}$. Define $\nu = \mu/2 + \mu'/2$. Then ν is positive on Π_1 and Π_2 . Let $\{\nu_i^\omega\}$ be the posteriors defined by ν . Then, $\hat{\Omega} \subseteq K^\infty(A)$ for $A = \{\omega \mid \nu_1^\omega(E) = \nu_2^\omega(E) = 1/2\}$ and $\nu(\hat{\Omega}) > 0$ which shows that agreeing to agree is possible for E with the prior ν that assigns a positive probability to every element of either partition.

As for the converse direction, assume that there are posteriors $\{\mu_i^\omega\}$ with a common prior μ which is positive on every element of Π_1 or Π_2 . Furthermore, assume that A is an event of the sort $\{\omega \mid \mu_1^\omega(E) = \mu_2^\omega(E) = p\}$, for some $0 < p < 1$ that satisfies $\mu(K^\infty(A)) > 0$. It

¹⁰Note that this is the same probability as the one used in Theorem 1's proof.

is sufficient to show that $K^\infty(A) \subseteq K^\infty(\neg K_1(\neg \hat{L}) \cap \neg K_2(\neg \hat{L}))$. For this it suffices to show that for $i = 1, 2$ and any $P \in \Pi_i$ such that $P \subseteq K^\infty(A)$, $P \cap \hat{L} \neq \emptyset$. Indeed, if to the contrary, $P \cap \hat{L} = \emptyset$ for such P , then by Proposition 4, $\mu(P) = 0$. This contradicts the positivity of μ on every element of either partition. ■

REFERENCES

- Aumann, R. (1976), Agreeing to disagree, *The Annals of Statistics*, 4, pp. 1236-1239.
- Aumann, R. (1999), Interactive epistemology I: Knowledge, *International Journal of Game Theory*, 28, pp. 263-300.
- Bacharach, M. (1985), Some extensions to a claim of Aumann in an axiomatic model of knowledge, *Journal of Economic Theory*, 37, pp. 167-190.
- Bonanno, G. and K. Nehring (1996), Fundamental Agreement: A New Foundation for the Harsanyi Doctrine, Working Paper No. 96-02, University of California, Davis.
- Cover, T. (1987), Pick the largest number, in *Open Problems in Communication and Computation*, T. Cover and B. Gopinath, eds., Springer Verlag, p. 152.
- Fagin, R., J.Y. Halpern, Y. Moses, and M.Y. Vardi (1995), *Reasoning About Knowledge*, The MIT Press.
- Feinberg, Y. (1995), Converse to the Agreement Theorem, Discussion Paper No. 83, Center for Rationality and Interactive Decision Theory, The Hebrew University.
- Geanakoplos, J. (1994), Common knowledge, in *Handbook of Game Theory*, Vol II, Eds. R.J. Aumann and S. Hart, Elsevier.
- Geanakoplos, J., and H. Polemarchakis (1982), We can't disagree forever, *Journal of Economic Theory*, 28, pp. 192-200.
- Hart, S., A. Heifetz and D. Samet (1996), 'Knowing whether', 'knowing that' and the cardinality of state spaces, *Journal of Economic Theory*, 70, No. 1, pp. 249-256.
- Heifetz, A. (1996), Comment on consensus without common knowledge *Journal of Economic Theory*, 70, pp. 273-277.
- McKelvey R. and T. Page (1986), Common knowledge, consensus, and aggregate information, *Econometrica* 54, pp. 109-127.
- Milgrom, P. and N. L. Stokey (1982), Information, trade and common knowledge, *Journal of Economic Theory*, 26, No. 1, pp. 17-27.
- Monderer, D. and D. Samet (1989), Approximating common knowledge with common belief, *Games and Economic Behavior*, 1, pp. 170-190.

- Morris, S. (1995), Trade with heterogenous prior beliefs and asymmetric information, *Econometrica*, 62, pp. 1327-1347.
- Nielsen, L. T., A. Brandenburger, J. Geanakoplos, R. McKelvey, and T. Page (1990), Common knowledge of an aggregate of expectations, *Econometrica*, 58, pp. 1235-1238.
- Parikh, R., and P. Krasucki (1990), Communication, consensus and knowledge, *Journal of Economic Theory*, 52, pp. 178–189.
- Rubinstein, A. and A. Wolinsky (1990), On the logic of “Agreeing to Disagree” type of results, *Journal of Economic Theory*, 51, pp. 184-193.
- Samet, D. (1990), Ignoring ignorance and agreeing to disagree, *Journal of Economic Theory*, 52 , pp. 190-207.
- Samet, D. (1998), Common priors and the separation of convex sets, *Games and Economic Behavior*, 24, No. 1 1998, pp.172-174.
- Samet, D. (1999), Bayesianism without learning, *Research in Economics*, 53, pp.227-242.
- Samet, D. (2000), Quantified beliefs and believed quantities, *Journal of Economic Theory*, 95, pp. 169-185.
- Samet, I., D. Samet, and D. Schmeidler (2002), One observation behind two-envelope puzzles, *American Mathematical Monthly*, forthcoming.
- Sebenius, J. and J. Geanakoplos (1983), Don’t bet on it: Contingent agreements with asymmetric information, *Journal of the American Statistical Association*, 78, pp. 424-426.
- Simon, R. (1998), The common prior assumption in belief spaces, manuscript.