HOW TO MAKE SENSE OF THE COMMON PRIOR ASSUMPTION UNDER INCOMPLETE INFORMATION

Giacomo Bonanno

and

Klaus Nehring [*]

Department of Economics,
University of California,
Davis, CA 95616-8578.
E-mail: gfbonanno@ucdavis.edu, kdnehring@ucdavis.edu

February 1996. Revised, April 1998

Abstract

The Common Prior Assumption (CPA) is central to the economics of information and the foundations of game theory. Recent contributions (Dekel and Gul, 1997, Gul, 1996, Lipman, 1995) have questioned its meaningfulness in situations of incomplete information where there is no *ex ante* stage and the primitives of the model are the individuals' belief hierarchies. We address this conceptual issue by providing characterizations of two local versions of the CPA which are in terms of the primitives and, therefore, do not involve a counterfactual and problematic *ex ante* stage. The characterizations involve three notions: Comprehensive Agreement, no error of beliefs and common belief in no error. Comprehensive Agreement is defined as the absence of "agreement to disagree" about any aspect of beliefs; it is a generalization of Aumann's (1976) notion of agreement. The entire analysis is carried out *locally*, that is, with reference to the "true state" (which represents the actual profile of belief hierarchies) and does not rely on the Truth Axiom for individual beliefs.

The results are also applied to the problem of generalizing the notion of Bayesian updating to single-person, intertemporal situations without perfect recall and without given information partitions.

[*] We are grateful to Bart Lipman and two referees for helpful and constructive comments. Seminar participants at Harvard, Penn, Princeton, USC and Yale provided useful comments. We also greatly benefited from discussions with participants at the SITE Workshop on the Epistemic Foundations of Game Theory (Stanford), in particular Steve Morris.

1. Introduction.

The Common Prior Assumption (CPA) plays an important role in game theory and the economics of information. It is the basic assumption behind decision-theoretic justifications of equilibrium reasoning in games (Aumann, 1987, Aumann and Brandenburger, 1995) and no-trade results with asymmetric information (Milgrom and Stokey, 1982). Not surprisingly, the CPA has attracted its share of criticism. In models of *asymmetric information* (where there is an *ex ante* stage at which the individuals have identical information and subsequently update their beliefs in response to private signals), the controversy focuses on the plausibility or appropriateness of assuming *commonness* of the prior beliefs (see Morris, 1995). What we are interested in are situations of *incomplete information*, where there is no *ex ante* stage and where the primitives of the model are the individuals' beliefs about the external world (their first-order beliefs), their beliefs about the other individuals' beliefs (second-order beliefs), etc., i.e. their hierarchies of beliefs. In this context, the CPA is a mathematical property whose conceptual content is not clear. This has given rise to a novel and, in a way, more radical, criticism of the CPA, one that questions its very *meaningfulness* in situations of incomplete information (Dekel and Gul, 1997, Gul, 1996, Lipman, 1995). We view this skepticism as a significant contribution to our understanding of situations of incomplete information.[1] On the other hand, we shall argue that it would be mistaken to draw the negative conclusion that the CPA cannot be given a meaningful interpretation in this context. Indeed, the main results of this paper (Theorems 1 and 2), make clear what we are assuming when we accept the CPA: they provide a characterization of Harsanyi consistency in terms of properties of the belief hierarchies that are entirely unrelated to the idea of an *ex ante* stage.

The skepticism concerning the CPA in situations of incomplete information can be developed along the following lines. As Mertens and Zamir (1985) showed in their classic paper, the description of the "actual world" in terms of belief hierarchies generates a collection of "possible worlds", one of which is the actual world. This set of possible worlds, or states, gives rise to a formal similarity between situations of asymmetric information and those of incomplete information. However, while a state in the former represents a real contingency, in the latter it is "a notational device for representing

[1] For a rather different but similarly forceful criticism of the implausible *counterfactual* nature of veil-of-ignorance interpretations of the CPA see Binmore and Brandenburger (1990).

the profile of infinite hierarchies of beliefs" (Gul, 1996, p. 3). As a result, notions such as that of a common "prior", "seem to be based on giving the artificially constructed states more meaning than they have" (Dekel and Gul, 1997, p.115). When the beliefs of the individuals can be viewed *as if* they were obtained by updating a common prior on some information, they are called *Harsanyi consistent.* Harsanyi consistency is a well-defined mathematical property, but, due to the "artificial nature" of the states in situations of incomplete information, "we do not know what it is that we would be accepting if we were to accept the common prior assumption" (Gul, 1996, p.5).

The key primitive notion in our analysis is that of Comprehensive Agreement. In order to motivate it, we take as point of departure the observation that, in some special cases, it is easy to find an interpretation of Harsanyi consistency which does not involve an *ex ante* stage. In particular, in situations of *complete* information (characterized by the fact that the beliefs of each individual are commonly known) Harsanyi consistency amounts to identity of beliefs across individuals. It thus seems natural, in situations of *incomplete* information, to think of Harsanyi consistency as likewise amounting to equality of those *aspects* of beliefs *that are commonly known*. For instance, one can take as an aspect of beliefs the subjective probability of an event E, in which case Agreement reduces to the notion introduced by Aumann (1976), which says that if the subjective probability of E of each individual is common knowledge, then these probabilities must be the same. Subjective probabilities of events are rather special aspects of beliefs and are not rich enough to fully capture the conceptual content of Harsanyi consistency. Thus, in Section 3, we define *Comprehensive Agreement* as the absence of "agreement to disagree" about *any* aspect of beliefs ("belief index") in an appropriately defined general class.

In Theorem 1 (Section 3) Comprehensive Agreement is shown to be equivalent to a weak *local* notion of Harsanyi consistency called Harsanyi Quasi Consistency. This result should be thought of as a representation theorem relating conditions on belief hierarchies (Comprehensive Agreement) to a mathematical construct (Harsanyi consistency). In the special case where the Truth Axiom is postulated for individual beliefs, Theorem 1 can be viewed (with the aid of the further characterization given in Proposition 1) as a local version of a well-known result in the literature establishing the equivalence between the CPA and no trade under *asymmetric* information (Morris, 1994). While primarily conceptual, this reinterpretation is not a matter of course, as evidenced by the fact that the above-mentioned critics did not seem to perceive any relevance of this pre-existing result to the issue of the meaningfulness of the CPA under incomplete information.

Rather than viewing Agreement as a (remarkable) implication of the existence of a common prior, we propose to take it as a primitive property of belief hierarchies. On the technical level, this change of perspective has two implications. First of all, one needs to define a general concept of what Agreement is about. This is formalized in Section 3 in terms of the notion of belief index. Secondly, Comprehensive Agreement defined as equality of commonly known aspects of beliefs is a concept that applies to *pairs* of individuals. In Section 5 we point out how one can mathematically extend the results to the case of more than two individuals, based on a notion of Expectation Consistency. It is not entirely clear, however (and potentially an interesting question for future research), whether Expectation Consistency is a legitimate primitive concept under incomplete information.

Harsanyi Quasi Consistency (equivalently, Comprehensive Agreement) is too weak a notion to allow the translation to situations of incomplete information of results that are based on the Common Prior Assumption, such as Aumann's (1987) characterization of correlated equilibrium. For this one needs a stronger notion of a *local* common prior, which is defined in Section 4 and called Strong Harsanyi Consistency. The second main result of this paper (Theorem 2) provides a characterization of Strong Harsanyi Consistency in terms of the conjunction of Comprehensive Agreement, no error of individual beliefs and common belief in no error of beliefs.

In Section 6 we reinterpret our results by applying them to the case of single-person, intertemporal belief revision; note that here under the standard assumption of perfect recall the CPA coincides with Bayesian updating, and is commonly accepted as eminently plausible. We set out to formulate a general notion of intertemporal consistency of Bayesian beliefs for settings in which recall may be imperfect and/or information implicit. This is a non-trivial issue, particularly since in the absence of perfect recall there is no well-defined prior to update on. We suggest that intertemporal belief consistency can be identified with intertemporal Harsanyi consistency derived from intertemporal Agreement.

Section 7 discusses related literature. All the proofs are given in Appendix C. Appendix A clarifies the relationship between belief indices and the decision functions studied in Cave (1983), Bacharach (1985) and Rubinstein and Wolinsky (1990). In Appendix B an intermediate (purely subjective) notion of Harsanyi consistency is discussed briefly.

2. Interactive belief models

DEFINITION 1. An *interactive Bayesian model* (or Bayesian model, for short)[2] is a tuple

$$\mathcal{B} = \langle \mathrm{N},\ \Omega,\ \tau,\ \Pi,\ \phi,\ \{\mathrm{p_i}\}_{\mathrm{i}\in\mathrm{N}} \rangle$$

where

- $\mathrm{N} = \{1, ..., n\}$ is a finite set of *individuals*.
- Ω is a finite set of *states* (or possible worlds)[3]. The subsets of Ω are called *events*.
- $\tau \in \Omega$ is the "true" or "actual" state[4].
- Θ is a set of *external circumstances* or *facts of nature*.
- $\phi : \Omega \to \Theta$ is a function that specifies, for every state, the facts of Nature that are true at that state.
- for every individual $\mathrm{i}\in\mathrm{N}$, $\mathrm{p_i} : \Omega \to \Delta(\Omega)$ (where $\Delta(\Omega)$ denotes the set of probability distributions over Ω) is a function that specifies her *probabilistic beliefs,* satisfying the following property [we use the notation $\mathrm{p_{i,\alpha}}$ rather than $\mathrm{p_i}(\alpha)$]: $\forall\, \alpha, \beta \in \Omega$,

$$\text{if} \quad \mathrm{p_{i,\alpha}}(\beta) > 0 \quad \text{then} \quad \mathrm{p_{i,\beta}} = \mathrm{p_{i,\alpha}} \tag{1}$$

Thus $\mathrm{p_{i,\alpha}} \in \Delta(\Omega)$ is individual i's subjective probability distribution at state α and condition (1) says that every individual knows her own beliefs. For every $\alpha \in \Omega$, we denote by $\|\mathrm{p_i} = \mathrm{p_{i,\alpha}}\|$ the event $\{\omega \in \Omega : \mathrm{p_{i,\omega}} = \mathrm{p_{i,\alpha}}\}$. It is clear that the set $\{\|\mathrm{p_i} = \mathrm{p_{i,\omega}}\| : \omega \in \Omega\}$ is a partition of Ω; it will be referred to as individual i's *type partition*.

For every individual $\mathrm{i}\in\mathrm{N}$, i's *possibility correspondence* $\mathrm{I_i} : \Omega \to 2^{\Omega} \setminus \varnothing$, is defined as follows:

---

[2] For a similar definition see, for example, Aumann and Brandenburger (1995), Dekel and Gul (1997) and Stalnaker (1996).

[3] Finiteness of Ω is a common assumption in the literature (cf. Aumann, 1987, Aumann and Brandenburger, 1995, Dekel and Gul, 1997, Morris, 1994, Stalnaker, 1996).

[4] We have included the true state in the definition of an interactive Bayesian model in order to stress the interpretation of the model as a representation of a particular profile of hierarchies of beliefs.

$$I_i(\alpha) = \text{supp } p_{i,\alpha}.$$ [5]

Thus, for every $\alpha \in \Omega$, $I_i(\alpha)$ is the set of states that individual i considers possible at α.

From this, individual i's *belief operator* $B_i : 2^{\Omega} \to 2^{\Omega}$ is obtained as follows: $\forall E \subseteq \Omega$, $B_iE = \{\omega \in \Omega : I_i(\omega) \subseteq E\}$. B_iE can be interpreted as the event that (i.e. the set of states at which) individual i *believes for sure* that event E has occurred (i.e. attaches probability 1 to E).[6] Notice that we have allowed for false beliefs by not assuming reflexivity of the possibility correspondences ($\forall \alpha \in \Omega$, $\alpha \in I_i(\alpha)$ or, equivalently, $p_{i,\alpha}(\alpha) > 0$), which – as is well known (cf. Chellas, 1984, p. 164) – is equivalent to the *Truth Axiom:* $\forall E \subseteq \Omega$, $B_iE \subseteq E$ (if the individual believes E then E is indeed true)[7].

The common belief operator B_* is defined as follows. First, for every $E \subseteq \Omega$, let $B_eE = \bigcap_{i \in N} B_iE$, that is, B_eE is the event that everybody believes E. The event that E is commonly believed is the infinite intersection:

$$B_*E = B_eE \cap B_eB_eE \cap B_eB_eB_eE \cap B_eB_eB_eB_eE \cap \ldots$$

The corresponding possibility correspondence I_* is then defined as follows: for every $\alpha \in \Omega$,

---

[5] If $\mu \in \Delta(\Omega)$, supp μ denotes the support of μ, that is, the set of states to which μ assigns positive probability.

[6] Thus Condition (1) of Definition 1 can be stated as follows: $\forall i \in N$, $\forall \alpha \in \Omega$, $\| p_i = p_{i,\alpha} \| = B_i \| p_i = p_{i,\alpha} \|$.

[7] On the other hand, note that by Condition (1) of Definition 1, the possibility correspondence of every individual i satisfies the following properties: $\forall \alpha, \beta \in \Omega$,

Transitivity: if $\beta \in I_i(\alpha)$ then $I_i(\beta) \subseteq I_i(\alpha)$,

Euclideanness: if $\beta \in I_i(\alpha)$ then $I_i(\alpha) \subseteq I_i(\beta)$.

It is well known (see Chellas, 1984, p. 164) that transitivity of the possibility correspondence is equivalent to *positive introspection* of beliefs: $\forall E \subseteq \Omega$, $B_iE \subseteq B_iB_iE$ (if the individual believes E then she believes that she believes E) and euclideanness is equivalent to *negative introspection* of beliefs: $\forall E \subseteq \Omega$, $\neg B_iE \subseteq B_i \neg B_iE$ (if the individual does not believe E, then she believes that she does not believe E; for every event F, $\neg F$ denotes the complement of F). Furthermore, non-empty-valuedness of the possibility correspondence is equivalent to *consistency* of beliefs: $\forall E \subseteq \Omega$, $B_iE \subseteq \neg B_i \neg E$ (the individual cannot simultaneously believe E and not E).

$$I_*(\alpha) = \{ \omega \in \Omega : \alpha \in \neg B_* \neg \{\omega\}\}.$$

It is well known[8] that I_* can be characterized as the *transitive closure* of $\bigcup_{i \in N} I_i$, that is,

$\forall \alpha, \beta \in \Omega$, $\beta \in I_*(\alpha)$ if and only if there is a sequence $\langle i_1, \ldots i_m \rangle$ in N and a sequence $\langle \eta_0, \eta_1, \ldots, \eta_m \rangle$ in Ω such that: (i) $\eta_0 = \alpha$, (ii) $\eta_m = \beta$ and (iii) for every $k = 1, \ldots, m$, $\eta_k \in I_{i_k}(\eta_{k-1})$.

We will make use of the following graphical representation of a possibility correspondence $I: \Omega \to 2^{\Omega}$ (cf. Figures 1-4). States are denoted by points and for every two states α and β, $\beta \in I(\alpha)$ if and only if either (i) α and β are enclosed in the same cell (denoted by a rounded rectangle), or (ii) there is an arrow from α to the cell containing β, or (iii) there is an arrow from the cell containing α to the cell containing β.[9]

A state in a model determines, for each individual, her beliefs about the external world (her first-order beliefs), her beliefs about the other individuals' beliefs about the external world (her second-order beliefs), her beliefs about their beliefs about her beliefs (her third-order beliefs), and so on, *ad infinitum*. An entire hierarchy of beliefs about beliefs about beliefs ... about the relevant facts is thus encoded in each state of an interactive belief model. For example, consider the following model, which is illustrated in Figure 1: N = {1, 2}, $\Omega = \{\tau, \beta\}$, Θ = {spelling: Harsanyi, spelling: Harsaniy}, $\phi(\tau)$ = {spelling: Harsanyi}, $\phi(\beta)$ = {spelling: Harsaniy}, $I_1(\tau) = I_1(\beta) = \{\beta\}$, $I_2(\tau) = \{\tau\}$, $I_2(\beta) = \{\beta\}$. Thus $I_*(\tau) = \{\tau, \beta\}$ and $I_*(\beta) = \{\beta\}$. Here state τ represents the following beliefs. Individual 2 is a game-theorist who knows the correct spelling of his name (Harsanyi), while individual 1 mistakenly believes that the spelling is Harsaniy. Furthermore, individual 1 mistakenly believes that it is common belief between them that the correct spelling is Harsaniy.

---

[8] See, for example, Bonanno (1996), Halpern and Moses (1992), Lismont and Mongin (1994). These authors also show that the common belief operator can be alternatively defined by means of a finite list of axioms, rather than as an infinite conjunction.

[9] This is a general way to represent possibility correspondences as directed graphs: see Bonanno and Nehring (1997, Remark 2)

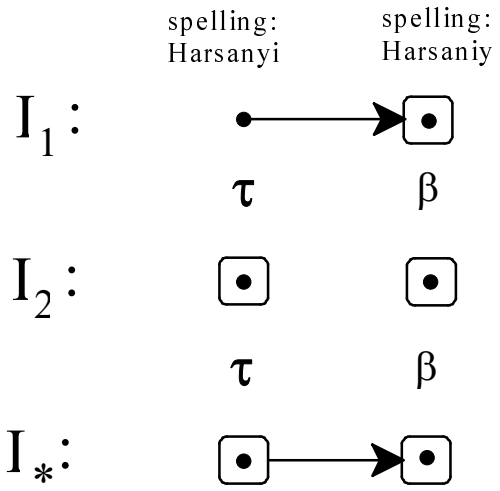

Figure 1

Conversely, given any profile of infinite hierarchies of beliefs (one for each individual) satisfying minimal coherency requirements, one can construct an interactive Bayesian model such that *at the true state* τ the beliefs of each individual $i \in N$ fully capture i's original infinite hierarchy of beliefs (see Armbruster and Boege, 1979, Boege and Eisele, 1979, Brandenburger and Dekel, 1993 and Mertens and Zamir, 1985)[10].

3. Harsanyi Quasi Consistency and Comprehensive Agreement

In this section we define a *local* version of the notion of a "common prior". The issue of how to properly formulate this notion locally does not seem to have been raised before in the literature.

In an incomplete information context, properties such as Harsanyi consistency (i.e. the existence of a common prior) are to be defined locally, that is, with respect to the true state τ. An equivalent, and mathematically more elegant, alternative is to define a property as an event, i.e. a set of states; the property is then satisfied at the true state τ if and only if τ belongs to that event. A characterization result will correspondingly be stated as the equality of two events.

[10] Finiteness of Ω, however, cannot be guaranteed in general.

DEFINITION 2. For every $\mu \in \Delta(\Omega)$, let $\mathbf{HQC}_\mu$ (for Harsanyi Quasi Consistency with respect to the "prior" μ) be the following event: $\alpha \in \mathbf{HQC}_\mu$ if and only if

(1) $\forall i \in N,\ \forall \omega \in I_*(\alpha),$ if $\mu(\|p_i = p_{i,\omega}\|) > 0$ then $p_{i,\omega} = \mu(\bullet \mid \|p_i = p_{i,\omega}\|)$ [11], and

(2) $\mu(I_*(\alpha)) > 0$.

If $\alpha \in \mathbf{HQC}_\mu$, μ is called a *local common prior at* α. Furthermore, let $\mathbf{HQC} = \bigcup_{\mu \in \Delta(\Omega)} \mathbf{HQC}_\mu$.

For example, in Figure 1 let μ be such that $\mu(\beta) = 1$. Then $\mathbf{HQC}_\mu = \{\tau, \beta\}$; hence $\mathbf{HQC} = \{\tau, \beta\}$.

The conceptual content of the notion of Harsanyi Quasi Consistency is not clear. The interpretation which is most often attached to it is the following paraphrase of Definition 2: imagine an *ex ante* stage where all the individuals had the same "information" represented by the set Ω and individual i had "prior" beliefs $\mu_i \in \Delta(\Omega)$; imagine next that, at state α, individual i is given the "information" represented by the event $\|p_i = p_{i,\alpha}\|$ and updates his prior μ_i on the basis of this information. If the "posterior" beliefs of individual i at state α coincide with $p_{i,\alpha}$ and all the individuals had the same prior beliefs, then their actual beliefs are consistent in the sense of Harsanyi (1967-68). As explained in the introduction, several authors have remarked that in a situation of incomplete information the notion of an *ex ante* stage is highly problematic. The reason for this is that the states other than τ (the true or actual state) are merely "fictitious constructs, used to clarify our understanding of the actual world" (Lipman, 1995, p. 2); thus the "prior stage is meaningless (i.e. it becomes impossible to associate the prior stage with a sensible thought experiment)" (Gul, 1996, p. 5). Our purpose is to find an alternative explication of the notion of Harsanyi consistency which does not involve a counterfactual and artificial *ex ante* stage.

Note first that in some special cases such an alternative interpretation is readily available: in particular, in the case of complete information (where the beliefs of each individual are commonly known) Harsanyi consistency amounts to identity of beliefs across individuals. Thus we propose, in situations of *incomplete* information, to think of Harsanyi consistency as likewise amounting to equality of beliefs in some appropriate sense. Clearly, it cannot be complete equality of beliefs, because of the very definition of

---

[11] $\mu(\bullet \mid \|p_i = p_{i,\omega}\|) \in \Delta(\Omega)$ is defined as follows: $\forall x \in \Omega,\ \mu(x \mid \|p_i = p_{i,\omega}\|) = \dfrac{\mu(\{x\} \cap \|p_i = p_{i,\omega}\|)}{\mu(\|p_i = p_{i,\omega}\|)}$, where, for every event $E \subseteq \Omega$, $\mu(E) = \sum_{\omega \in E} \mu(\omega)$.

incomplete information: such equality[12] precludes, of necessity, incomplete information. At most one can require equality of *aspects* of beliefs and the question is: which aspects? Taking a cue, again, from the case of complete information, it seems sound to require equality of those aspects of beliefs that are *commonly known*. For instance, one can take as an aspect of beliefs the subjective probability of an event E. In that case the condition we are looking for is simply the notion of Agreement introduced by Aumann (1976), which says that if, for a given event E, the subjective probability of E of each individual is common knowledge, then these probabilities are the same. However, subjective probabilities of events are rather special aspects of beliefs and, as illustrated in Figure 2 below, not rich enough to fully capture the conceptual content of Harsanyi consistency. Our aim in this section is to define the notion of "aspect of belief" in general and *Comprehensive Agreement* as the absence of "agreement to disagree" about any such aspect.

The nature of our argument can be illustrated with the aid of Gul's (1996) key example, which is reproduced in Figure 2. Individual 1 believes that some fact q (e.g. that it rains) is true. Individual 2 is uncertain as to whether or not q is true, but is certain that individual 1 knows whether q is true or not and, in fact, it is common knowledge that 1 knows whether or not q holds. Individual 1, in turn, is uncertain as to whether 2 assigns probability 0.4 or probability 0.5 to q being true and considers both possibilities equally likely. This, again, is common knowledge. These belief hierarchies are represented by state τ.

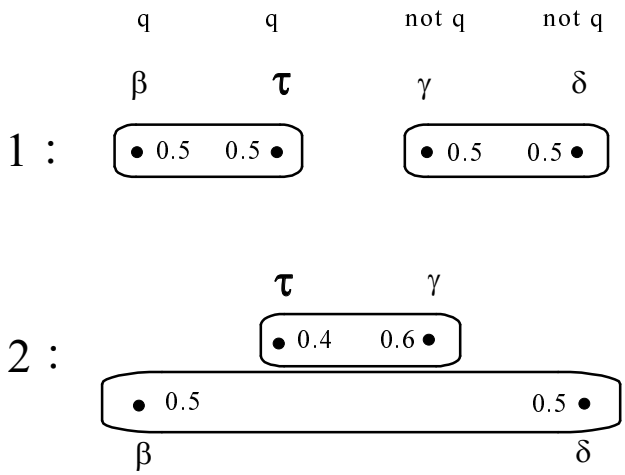

Figure 2

In this example, for no event $E \neq \Omega$ is the subjective probability of E of both individuals common belief at τ. Hence Agreement as defined by Aumann (1976) is vacuously satisfied. However, the individuals'

[12] Of beliefs on the state space Ω, i.e. equality of the belief hierarchies.

beliefs at τ cannot be derived from a common prior over the set $\{\tau, \beta, \gamma, \delta\}$ [13], that is, they are not Harsanyi consistent. The preceding discussion suggests that there ought to be *some* aspect of the individuals' beliefs about which they agree to disagree. Indeed, such an aspect is given by the expectation of the random variable Y defined as follows: $Y(\tau) = 0$, $Y(\beta) = 0.22$, $Y(\gamma) = 0.20$ and $Y(\delta) = 0.02$. It is easy to check that it is common belief at τ (true at every state) that 1's expectation is 0.11 while 2's expectation is 0.12. Given any random variable Y, one can view the expectation of Y given $p \in \Delta(\Omega)$ as an index or aspect of the beliefs represented by p. In general, a *belief index* will be any function on $\Delta(\Omega)$.

We now define formally a general notion of agreement. Agreement *as equality of beliefs* is essentially a two-person property. Hence, for the remaining part of this section, we specialize to the case where N = {1,2} (for a possible extension to the case of more than two individuals see Section 5).

DEFINITION 3. Let X be a set with at least two elements. A *proper belief index* is a function $f : \Delta(\Omega) \to X$ [14] that satisfies the following property: $\forall\, p, q \in \Delta(\Omega)$, $\forall\, x \in X$, $\forall\, a \in [0,1]$,

$$\text{if } f(p) = f(q) = x \text{ then } f(ap + (1-a)q) = x.$$

Let $\mathcal{F}$ denote the class of proper belief indices.[15]

The following are examples of proper belief indices:

(i) Let $E \subseteq \Omega$ be an arbitrary event, $X = [0,1]$ and f^E the following belief index: $f^E(p) = p(E)$; thus $f^E(p_{i,\alpha})$ is individual i's subjective probability of event E at state α.

(ii) Let $Y : \Omega \to \mathbb{R}$ be a random variable, $X = \mathbb{R}$ and f_Y be the belief index given by $f_Y(p) = \sum_{\omega \in \Omega} Y(\omega) p(\omega)$; thus $f_Y(p_{i,\alpha})$ is i's subjective expectation of Y at state α.

---

[13] Indeed such a common prior μ would have to satisfy the following four, jointly inconsistent, conditions: $\mu(\tau) = \mu(\beta)$, $\mu(\gamma) = \mu(\delta)$, $\mu(\beta) = \mu(\delta)$ and $\mu(\gamma) = 1.5\, \mu(\tau)$.

[14] It may seem that a belief index f depends on the set of states Ω. However, this is not so: one should think of f as being defined on the "universal belief space" (cf. Mertens and Zamir, 1985). Indeed, all that matters is the restriction of f to $I_*(\tau)$.

[15] See Appendix A for a discussion of the relationship between belief indices and the "decision functions" considered in the literature (Bacharach, 1985, Cave, 1983, Rubinstein and Wolinsky, 1990).

(iii) Let A be a set of actions, $X = 2^A$ and $U : A \times \Omega \to \mathbb{R}$ a utility function. Define the belief index $f_U : \Delta(\Omega) \to 2^A$ as follows: $f_U(p) = \arg\max_{a \in A} \sum_{\omega \in \Omega} U(a,\omega)\, p(\omega)$; thus $f_U(p_{i,\alpha})$ is the set of actions that maximize individual i's expected utility at state α.

Given a proper belief index $f : \Delta(\Omega) \to X$ and an individual $i \in N$, define $f_i : \Omega \to X$ by $f_i(\omega) = f(p_{i,\omega})$ and, for every $x \in X$, denote the event $\{\omega \in \Omega : f_i(\omega) = x\}$ by $\|f_i = x\|$.

DEFINITION 4. Given a Bayesian model and a proper belief index $f : \Delta(\Omega) \to X$, at $\alpha \in \Omega$ there is *Agreement for f* or *f-Agreement* if and only if, for all $x_1, x_2 \in X$,

$$\text{if} \quad \alpha \in B_*\left(\|f_1 = x_1\| \cap \|f_2 = x_2\|\right) \quad \text{then} \quad x_1 = x_2 .$$

That is, if at α it is common belief that individual 1's belief index is x_1 and individual 2's index is x_2, then $x_1 = x_2$. Let ***f*-Agree** be the set of states where there is *f*-Agreement:

$$\textbf{\textit{f}-Agree} \;=\; \bigcap_{\substack{x_1, x_2 \in X \\ x_1 \neq x_2}} \neg B_* \left(\|f_1 = x_1\| \cap \|f_2 = x_2\| \right).$$

EXAMPLES. (i) Consider a model where there is *complete information*, that is, $\forall E \subseteq \Omega$, $\forall i \in \{1,2\}$, $B_i E \subseteq B_* B_i E$. Then $\alpha \in$ ***id*-Agree** [where *id* is the identity function on $\Delta(\Omega)$] if and only if $p_{1,\alpha} = p_{2,\alpha}$ if and only $\alpha \in$ **HQC**. (ii) In the example of Figure 2, f_Y **-Agree** $= \varnothing$ [where Y is the random variable defined there and, for every $p \in \Delta(\Omega)$, $f_Y(p) = \sum_{\omega \in \Omega} Y(\omega)\, p(\omega)$].

The notion of proper belief index defines in a general way what disagreement may be *about*; it is essential to an appropriate definition of "agreement in general" *under incomplete information*, and is lacking from the literature, which has dealt with generalizations of Aumann's (1976) theorem under asymmetric information[16]. Properness is necessary to ensure that public (i.e. commonly believed)

[16] Bacharach (1985), Cave (1983), Geanakoplos (1989), Geanakoplos and Polemarchakis (1982), Samet (1990), Rubinstein and Wolinsky (1990).

inequality of the value of the belief index can indeed be interpreted as genuine disagreement, rather than as a byproduct of asymmetric information. To clarify this point, consider the following example: $\Omega = \{\tau, \beta\}$, $p_{1,\tau} = p_{1,\beta} = \begin{pmatrix} \tau & \beta \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$, $p_{2,\tau} = \begin{pmatrix} \tau & \beta \\ 1 & 0 \end{pmatrix}$ and $p_{2,\beta} = \begin{pmatrix} \tau & \beta \\ 0 & 1 \end{pmatrix}$. Let f be the *improper* belief index defined by: $f(p) = |1 - 2p(\tau)|$. Then $\tau \in B_*\left(\|f_1 = 0\| \cap \|f_2 = 1\|\right)$. This public inequality of the value of f merely reflects the public fact that individual 2 knows the true state whereas individual 1 does not, and therefore cannot be viewed as genuine disagreement. This example is generalized in Corollary A.1 in Appendix A.

A comprehensive notion of agreement is obtained from the entire class of proper belief indices.

DEFINITION 5. Let **CA** (for Comprehensive Agreement) be the following event:[17]

$$\mathbf{CA} = \bigcap_{f \in \mathcal{F}} f\text{-}\mathbf{Agree}.$$

The following theorem characterizes Comprehensive Agreement as equivalent to Harsanyi Quasi Consistency[18]. The key step in the proof of Theorem 1 is the observation that $\alpha \in \mathbf{HQC}$ is equivalent to non-emptiness of the intersection of the convex hull of the sets of commonly possible beliefs of the individuals at α: $\alpha \in \mathbf{HQC}$ if and only if $\bigcap_{i \in N} \text{co}\, P_i(\alpha) \neq \varnothing$ where $\text{co}\, P_i(\alpha)$ is the convex hull of $P_i(\alpha) = \{p_{i,\omega} \in \Delta(\Omega) : \omega \in I_*(\alpha)\}$ (see Lemma C.3 in Appendix C).

THEOREM 1. CA = HQC.

---

[17] Comprehensive Agreement searches, as it were, logical space for disagreement about anything. Any such (dis)agreement is an *actual property* of individual belief profiles at the true state τ, rather than being merely a "behavioral" *consequence* of such beliefs in a hypothetical "opinion game", on roughly the same footing, for instance, as absence of common knowledge trade (as in Geanakoplos, 1989, 1992). Being an actual property, it makes sense to view Comprehensive Agreement as a *primitive* qualitative condition on the individuals' belief profiles, rather than merely as a welcome but coincidental fact which certain profiles happen to exhibit.

[18] A qualitative version of Theorem 1 – in terms of "union-consistent" qualitative belief indices and qualitative agreement – is given in Bonanno and Nehring (1997, Proposition 3).

Note that Comprehensive Agreement has been defined as a fact about agents' beliefs. It may obtain without this being known to every individual. On the other hand, "disagreement", defined as the negation of Comprehensive Agreement, is equivalent to common belief in disagreement, that is, $\neg\mathbf{CA} = B_*(\neg\mathbf{CA})$.

The following proposition makes the notion of Comprehensive Agreement more transparent by establishing its equivalence to agreement on two-valued proper belief indices, which in turn are those with a betting interpretation. Let $\mathcal{J}_2 \subseteq \mathcal{J}$ be the class of proper belief indices $f : \Delta(\Omega) \to X$ such that: (1) $X = \{0, 1\}$, and (2) $f^{-1}(1)$ is closed.

PROPOSITION 1. (i) $\mathbf{CA} = \bigcap_{f \in \mathcal{J}_2} f\text{-}\mathbf{Agree}$;

(ii) $f \in \mathcal{J}_2$ if and only if there exists a random variable $Y : \Omega \to \mathbb{R}$ such that, for all $p \in \Delta(\Omega)$, $f(p) = \begin{cases} 1 & \text{if } \sum_{\omega \in \Omega} Y(\omega)\, p(\omega) \geq 0 \\ 0 & \text{otherwise} \end{cases}$.

In view of Proposition 1, Theorem 1 is a local re-interpretation (for situations of incomplete information) and a generalization (to the case where false beliefs are allowed) of a result by Morris (1994, Lemma A.3, p. 1343) concerning the equivalence between the existence of a common prior and the absence of common knowledge trade in situations of asymmetric information. A related result can also be found in Feinberg (1995, Proposition 2, p. 5)[19].

4. Strong Harsanyi Consistency requires Truth

The notion of Harsanyi Quasi Consistency is rather weak: it allows the "common prior" to assign zero probability to the true beliefs of *all* the individuals (even if none of the individuals has false beliefs: see example below) and it is compatible with some individuals believing that there is agreement to disagree.

[19] See Section 7 for further discussion.

EXAMPLE. Consider the model of Figure 3. Let $\mu \in \Delta(\Omega)$ be such that $\mu(\beta) = 1$. Then $\mathbf{HQC}_\mu = \Omega$. Thus at τ Harsanyi Quasi Consistency is satisfied even though the type (beliefs) of each individual is assigned zero probability by μ. Note that at τ *both individuals have correct beliefs* ($\tau \in I_1(\tau) \cap I_2(\tau)$).

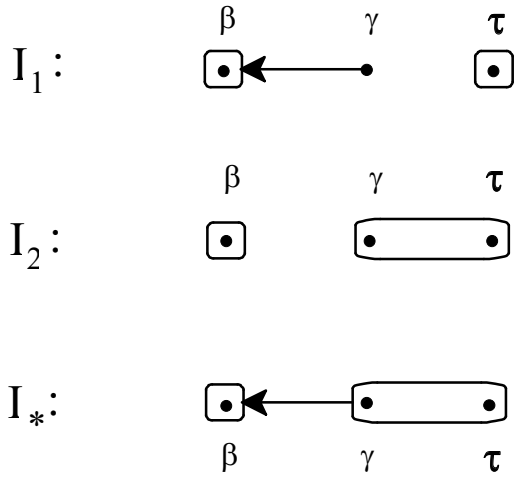

Figure 3

Not surprisingly, Harsanyi Quasi Consistency is too weak a notion to allow the translation to situations of incomplete information of results that are based on the Common Prior Assumption, such as Aumann's (1987) characterization of correlated equilibrium (see Bonanno and Nehring, 1998a, Section 4). In order to strengthen the notion of Harsanyi Quasi Consistency one needs to tighten the connection between the implied prior and the true beliefs / state. The following definition does so by requiring the prior to assign positive probability to the true state.

DEFINITION 6. For every $\mu \in \Delta(\Omega)$, let $\mathbf{SHC}_\mu$ (for Strong Harsanyi Consistency with respect to the "prior" μ) be the following event: $\alpha \in \mathbf{SHC}_\mu$ if and only if

(1) $\alpha \in \mathbf{HQC}_\mu$, and

(2) $\mu(\alpha) > 0$.

That is, $\mathbf{SHC}_\mu = \mathbf{HQC}_\mu \cap \text{supp } \mu$. Furthermore, let $\mathbf{SHC} = \bigcup_{\mu \in \Delta(\Omega)} \mathbf{SHC}_\mu$.

For example, in the model of Figure 1, while **HQC** = $\{\tau, \beta\}$, **SHC** = $\{\beta\}$. It is easily verified that the common prior at α is *locally unique*, that is, if $\alpha \in \mathbf{SHC}_{\mu} \cap \mathbf{SHC}_{\nu}$ then $\mu(\cdot \mid I_*(\alpha)) = \nu(\cdot \mid I_*(\alpha))$. An analogous claim cannot be made for **HQC**.

It is shown in Appendix C (Corollary C.1) that if $\alpha \in \mathbf{SHC}_{\mu}$ then $\mu(\omega) > 0$ for all $\omega \in I_*(\alpha) \cup \{\alpha\}$. As a result, the belief hierarchies at α are entirely determined by the local common prior at α and by the possibility correspondences I_i.

A Bayesian model is *globally strongly Harsanyi consistent* if **SHC** = Ω. It is easily seen that in this case, for some $\mu \in \Delta(\Omega)$, $\mathbf{SHC}_{\mu} = \Omega$.[20] Thus, global Strong Harsanyi Consistency is equivalent to the existence of a *global common prior* μ. Any global common prior has "full marginal support" (FMS), i.e. $\mu(\|p_i = p_{i,\alpha}\|) > 0$, $\forall i \in N$, $\forall \alpha \in \Omega$.

A *common-prior model of Asymmetric Information* is a tuple $\mathcal{A} = \langle N, \Omega, \tau, \Theta, \phi, \{\mathcal{T}_i\}_{i \in N}, \mu \rangle$ where N, Ω, τ, Θ and ϕ are as in Definition 1, $\mathcal{T}_i$ is a partition of Ω (individual i's information partition) and $\mu \in \Delta(\Omega)$ is the individuals' common prior. The (ex interim) Bayesian model $\mathcal{B} = \langle N, \Omega, \tau, \Theta, \phi, \{p_i\}_{i \in N} \rangle$ is *compatible with* $\mathcal{A}$ if, for all $i \in N$ and $\alpha \in \Omega$, $T_i(\alpha) = \|p_i = p_{i,\alpha}\|$ (where $T_i(\alpha)$ is the cell of $\mathcal{T}_i$ that contains α) and $\forall \omega \in \Omega$, $\forall i \in N$, if $\mu(\|p_i = p_{i,\omega}\|) > 0$ then $p_{i,\omega} = \mu(\cdot \mid \|p_i = p_{i,\omega}\|)$. If μ has FMS then $\mathcal{B}$ is unique and has μ as a global common prior. If μ does not have FMS, and $\mathcal{B}$ is compatible with $\mathcal{A}$, then $\mu(\mathbf{SHC}_{\mu}) = 1$, that is, ex interim beliefs are strongly Harsanyi consistent with ex ante probability 1. It need not be the case, though, that **HQC** = Ω.

Having ensured the positive "prior" probability of the true state, Strong Harsanyi Consistency allows an immediate local translation (to situations of incomplete information) of probability one results based on the Common Prior Assumption obtained in an asymmetric information context. In particular, in a strategic set-up such as the one of Aumann and Brandenburger (1985), Aumann's (1987)

---

[20] **SHC** = Ω holds if and only if $\forall \alpha \in \Omega$, $\exists \mu(\alpha) \in \Delta(\Omega)$ such that $\alpha \in \mathbf{SHC}_{\mu(\alpha)}$. Letting $\mu = \sum_{\alpha \in \Omega} \frac{1}{|\Omega|} \mu(\alpha)$ we have that, $\forall \alpha \in \Omega$, $\alpha \in \mathbf{SHC}_{\mu}$, that is, $\mathbf{SHC}_{\mu} = \Omega$.

characterization of correlated equilibrium translates into the local and non-probabilistic statement that common belief of rationality and **SHC** at the true state imply that the strategy profile at the true state belongs to the support of some correlated-equilibrium distribution (see Bonanno and Nehring, 1998a, Theorem 3). Theorem 2 below replaces **SHC** with locally meaningful assumptions[21]. While **SHC** is more satisfactory in its implications, we have introduced a separate notion of **HQC** for two reasons: to determine the common-prior implications of Comprehensive Agreement in isolation – without reliance on Truth-type conditions which do not seem to be justifiable on rationality grounds alone (as, for instance, in the example of Figure 1)[22] – and to emphasize an element of arbitrariness in the formalization of the Common Prior Assumption that arises in an incomplete-information setting.

Let **T** (for Truth) be the following event:

$$\mathbf{T} = \bigcap_{i \in N} \bigcap_{E \in 2^{\Omega}} (\neg B_i E \cup E)$$

Thus $\alpha \in \mathbf{T}$ if and only if at α every individual has correct beliefs (for every event E and every individual i, if $\alpha \in B_i E$ then $\alpha \in E$)[23]. For example, in Figure 1 $\tau \notin \mathbf{T}$ and in Figure 3, although $\tau \in \mathbf{T}$, $\tau \notin B_* \mathbf{T} = \{\beta\}$

THEOREM 2. **(i)** **for any n,** $\mathbf{SHC} = \mathbf{HQC} \cap \mathbf{T} \cap B_* \mathbf{T}$,

(ii) **hence for n = 2,** $\mathbf{SHC} = \mathbf{CA} \cap \mathbf{T} \cap B_* \mathbf{T}$

Thus the gap between **HQC** and **SHC** is filled by the requirement that the individuals' beliefs be correct and that this fact be common belief (we call this joint condition the Truth Axiom). This condition is not just an internal consistency assumption on beliefs but rather adds external requirements of conformity between individual beliefs and facts of nature. As a consequence the notion of Strong Harsanyi Consistency is more than a consistency requirement proper. This observation motivates an intermediate notion, which is a purely subjective strengthening of Harsanyi Quasi Consistency: it

---

[21] For a *non-epistemic* justification of correlated equilibrium based on no-arbitrage arguments, see Nau and McCardle (1990).

[22] Sometimes the Common Prior Assumption is motivated by appeal to the "Harsanyi doctrine" (see, for example, Aumann, 1987).

[23] It is well-known that $\alpha \in \mathbf{T}$ if and only if, for every $i \in N$, $\alpha \in I_i(\alpha)$.

requires only that the beliefs of each individual be accounted for (i.e. be given positive probability) by the common prior. This intermediate notion (called **HC**) is discussed in Appendix B, where it is shown that **HC** = B_***SHC** and **HQC** = $\neg B_* \neg$**HC.**

By Theorem 2, if $\alpha \in$ **HQC** $\cap \neg$**SHC** then at α the Truth Axiom is violated. However, one can improve Theorem 2 significantly by noting that **HQC** (hence **CA**) itself incorporates "Truth conditions". To see this, let $\mathbf{T}^*$ (for Truth *of* common belief) and $\mathbf{T}_{CB}$ (for Truth *about* common belief) be the following events:

$$\mathbf{T}^* = \bigcap_{E \in 2^\Omega} (\neg B_* E \cup E).$$

$$\mathbf{T}_{CB} = \bigcap_{i \in N} \bigcap_{E \in 2^\Omega} (\neg B_i B_* E \cup B_* E)$$

Thus $\alpha \in \mathbf{T}^*$ if and only if at α whatever is commonly believed is true (for every event E, if $\alpha \in B_* E$ then $\alpha \in E$)[24]. Truth of common belief is qualitatively weaker than Truth: since it is always true that $B_* \mathbf{T}^* = \Omega$, $\mathbf{T}^*$ can be viewed as Truth shorn of any intersubjective implications. $\mathbf{T}_{CB}$, on the other hand, captures the notion that individuals are correct in their beliefs about what is commonly believed: $\alpha \in \mathbf{T}_{CB}$ if and only if, for every event E, if, at α, individual i believes that E is commonly believed, then, at α, E is indeed commonly believed (for every event E and individual i, if $\alpha \in B_i B_* E$ then $\alpha \in B_* E$)[25]. With the aid of these two notions, Theorem 2 can be strengthened as follows:[26]

SHC = **CA** $\cap \mathbf{T}^* \cap B_* \mathbf{T}_{CB}$. Thus Strong Harsanyi Consistency is equivalent to the conjunction of Comprehensive Agreement, Truth *of* common belief and common belief in Truth *about* common belief. Even with this improved characterization it remains true that **SHC** is a significantly stronger

---

[24] It is straightforward that $\alpha \in \mathbf{T}^*$ if and only if, $\alpha \in I_*(\alpha)$.

[25] Within a belief and knowledge framework, it is shown in Bonanno and Nehring (1998b) that $\mathbf{T}_{CB}$ can be interpreted as an intersubjective notion of caution.

[26] In Bonanno and Nehring (1998b, Proposition 5) it is shown that $\mathbf{T} \cap B_* \mathbf{T} = \mathbf{Q} \cap \mathbf{T}^* \cap B_* \mathbf{T}_{CB}$ where **Q** is the property of qualitative agreement, which is implied by **CA**. Using this result and Theorem 2, we get that **SHC** $\supseteq$ **CA** $\cap \mathbf{T}^* \cap B_* \mathbf{T}_{CB}$. The converse inclusion follows from Theorem 2 and the fact that $\mathbf{T} \subseteq \mathbf{T}^* \cap \mathbf{T}_{CB}$.

property than **HQC**: while the latter does not rely on any event being commonly believed by the individuals (Comprehensive Agreement stating merely the *absence* of common belief in certain kinds of events), the former involves positive common belief assumptions in an essential way.[27] We have noted above the weakness of $\mathbf{T}^*$ compared to **T**; likewise, $B_*\mathbf{T}_{CB}$ is significantly weaker than $B_*\mathbf{T}$. In particular, in contrast to $B_*\mathbf{T}$, Truth about common belief is vacuously satisfied in situations of complete information. In addition (again in contrast to belief in the correctness of others' beliefs), Truth about common belief has no agreement-type implications (see Bonanno and Nehring, 1997, Remark 9).

$\mathbf{T}_{CB}$ ensures that an individual's belief in *f*-disagreement implies *de facto* existence of *f*-disagreement (hence, by contraposition, also that Comprehensive Agreement can obtain only if it is known to both individuals). Given $\mathbf{T}_{CB}$ (but not necessarily otherwise), individuals can communicate their *f*-disagreement without revealing any information about their beliefs, in a Geanakoplos-Polemarchakis-style dialogue[28], and thus, by "disagreeing forever" about *f*, empirically confirm their silently assumed disagreement.

5. Extension to more than two agents

While (Strong/Quasi) Harsanyi Consistency has been defined for the general case of n individuals, Comprehensive Agreement as equality of commonly known aspects of beliefs is restricted to two individuals.

One may wonder whether there is a way of extending the above characterization of Harsanyi

---

[27] To further understand the above characterization of **SHC**, let $\mathbf{NI}^*$ (for Negative Introspection of common belief) be the following event: $\mathbf{NI}^* = \bigcap_{E \in 2^\Omega} (B_*E \cup B_*\neg B_*E)$. Thus $\alpha \in \mathbf{NI}^*$ if and only if, for every event E, if, at α, E is not commonly believed then, at α, it is commonly believed that E is not common belief. It can be shown (see Bonanno and Nehring, 1998a) that $\mathbf{NI}^* = \mathbf{T}_{CB} \cap B_*\mathbf{T}_{CB}$ and therefore $\mathbf{T}^* \cap B_*\mathbf{T}_{CB} = \mathbf{T}^* \cap B_*\mathbf{NI}^* = \mathbf{T}^* \cap B_*\mathbf{T}^* \cap \mathbf{NI}^* \cap B_*\mathbf{NI}^*$. Hence the gap between the weakest (**HQC**) and the strongest (**SHC**) versions of Harsanyi consistency is bridged by effectively assuming *locally* the knowledge axioms for common belief that do not come with its definition, namely Negative Introspection and Truth. This is significantly weaker than assuming locally the knowledge axiom for individual belief, i.e. $\mathbf{T} \cap B_*\mathbf{T}$, as in Theorem 2.

[28] See Geanakoplos and Polemarchakis (1982).

consistency to the case of more than two individuals. A possible avenue is suggested by Proposition 1, according to which – in the two-person case – Comprehensive Agreement is equivalent to *Expectation Consistency*, defined as the nonexistence of a random variable $Y : \Omega \to \mathbb{R}$ such that it is common belief that individual 1's expectation of Y is positive and individual 2's expectation of $-Y$ is positive. This can be generalized to the case of n individuals as follows: Expectation Consistency is satisfied at a state $\alpha \in \Omega$ if and only if there do not exist random variables $Y_i : \Omega \to \mathbb{R}$ (i = 1, …, n) such that: (1) $\forall\, \omega \in \Omega$, $\sum_{i \in N} Y_i(\omega) = 0$, and (2) at α it is common belief that, for every individual i, i's subjective expectation of Y_i is positive, that is, $\alpha \in B_*(\|E_1 > 0\| \cap \ldots \cap \|E_n > 0\|)$, where $\|E_i > 0\| = \{\omega \in \Omega : \sum_{x \in \Omega} Y_i(x) p_{i,\omega}(x) > 0\}$. Replacing Comprehensive Agreement with Expectation Consistency, the characterization results of Theorem 1 and of Theorem 2(ii) hold for the case of any number of individuals[29].

Expectation Consistency seems conceptually rather less satisfactory than Comprehensive Agreement. In particular, since Expectation Consistency refers to *different* belief indices for different individuals, it cannot be understood as a generalization of the notion of "equality of beliefs", in contrast to Comprehensive Agreement. As a result, it is not clear whether Expectation Consistency can be meaningfully elucidated without implicit reference to an *ex ante* stage. Finally, and relatedly, it is not clear how Expectation Consistency can be axiomatically justified (as it has been in the case of two individuals by Proposition 1) when there are more than two individuals.

6. Intertemporal application: Bayesian updating without a prior

In this section we reinterpret our results by applying them to the case of single-person, intertemporal belief revision. In contrast to multi-person settings, the Harsanyi doctrine (which states that differences in beliefs ought to be attributed to differences in information) has largely gone unchallenged in this context. In fact, under the standard hypothesis of perfect recall (a sequence of information partitions such that the partition at time t+1 is a refinement of the partition at time t) the

[29] The direct proof strategy of Theorem 1 can be used here as well: this has been shown by Samet (1996a), who provides an elegant proof of the required characterization of the non-emptiness of the intersection of a finite number of closed convex subsets of the unit simplex.

Harsanyi doctrine can be identified with the assumption of Bayesian updating. Yet the assumption of perfect recall is empirically often unrealistic; its relaxation has been the object of recent investigation[30]. Our purpose in this section is to formulate a general notion of intertemporal consistency of Bayesian beliefs without making restrictive assumptions on the nature of belief formation. This is a non-trivial issue, particularly since in the absence of perfect recall there is no well-defined prior to update on! Below, we propose to identify intertemporal belief consistency with intertemporal Harsanyi consistency derived from intertemporal Agreement.

The most general way of representing an individual's evolution of beliefs over time is precisely in terms of a Bayesian model (cf. Definition 1), where the set N is now interpreted as a set of dates and, for $t \in N$, the event $B_t E$ represents the event that at date t the individual believes E. The true state τ encodes the actual evolution of the individual's beliefs over time, that is, the facts believed by the individual at every date, as well as her beliefs about her past and future beliefs. For example, the model of Figure 1 represents the situation where initially (i.e. at date 1) the individual wrongly believes the spelling to be Harsaniy and anticipates maintaining this belief in the future, but in fact at date 2 she switches to the correct belief that the spelling is Harsanyi. Furthermore, at date 2, she remembers that she previously held a different belief. This example is particularly interesting in that it captures the type of belief revision that underlies much of equilibrium analysis in extensive games, where updating on zero probability events is crucial. The example confirms that in a truly general analysis of interactive belief systems one cannot and should not postulate the Truth Axiom.

Consider now the example of Figure 3. Let q be a fact of Nature which is true at τ and γ and false at β. In a single-person, intertemporal interpretation of this model the individual initially correctly believes q and anticipates maintaining the belief in q in the future, but correctly anticipates not remembering his previous beliefs: at date 2 he does not remember whether at date 1 he believed q or not q. Thus we have an example where perfect recall is lacking.

Under the single-person intertemporal interpretation, the common belief operator captures the notion of *intertemporally evident belief:* $B_* E$ is the event that at every date the individual believes E and believes that she believed E in the past and will believe E in the future and so on. If f is a belief

[30] See, for example, Battigalli (1995), Halpern (1995), Piccione and Rubinstein (1994).

index and $\alpha \notin f$ **-Agree** then at α it is intertemporally evident to the individual that the value of the index at date 1 is different from the value of the index at date 2. Comprehensive Agreement rules this out for every proper belief index and therefore can be viewed as a generalization of the principle of reflection and of dynamic Dutch book arguments[31]. In this context our main results (Theorems 1 and 2) can be interpreted as providing a justification of "Bayesian updating without a prior"; note that the truth-like conditions such as $B_* \mathbf{T}_{CB}$ are non-trivial here in that they rule out the absence of *actual* surprises.

7. Related literature

In an interesting recent paper, Lipman (1995) has shown that the existence of a common prior imposes almost no restrictions on the first k levels ($k < \infty$) of a belief hierarchy. From the point of view of this paper, Lipman's result may be viewed as a meta-proposition stating that any characterization of Harsanyi consistency must involve *infinite*-order conditions. This suggests that characterizing conditions cannot be much simpler than the ones offered in Theorems 1 and 2, which are infinite-order conditions of a simple kind. Note that Lipman's result *is* compatible with the existence of a characterization, unlike ours, in terms of arbitrarily deep (unbounded) finite levels of belief hierarchies (such a result has been recently provided by Samet, 1996b).

In the context of asymmetric information, Morris (1991, 1994) established the equivalence between the existence of a common prior and the absence of common knowledge trade. As remarked in Section 3, in view of Proposition 1, the equivalence between Harsanyi Quasi Consistency and Comprehensive Agreement can be viewed as a local translation of Morris's result to the case of incomplete information. Morris does not discuss the relevance of this result to incomplete information settings. The focus of his paper is on the extent to which no trade results hold in an asymmetric information framework when there are heterogeneous priors. Morris's results establish the relationship between interim efficiency properties (which are closely related to agreement properties) and consistency conditions on prior beliefs.

The closest work to ours is the independent contribution of Feinberg (1995, 1996), which nicely complements ours by providing infinite and syntactic versions of a characterization relating the

[31] See, for example, Goldstein (1983), van Fraassen (1984), Maher (1993).

CPA to Agreement. Feinberg, like Morris, does not raise the issue of the conceptual content of the CPA under incomplete information[32]. In contrast to Feinberg, we have been concerned in this paper with the definition of an appropriate general notion of Agreement under incomplete information, and with the role of "Truth-type" assumptions in the characterization of alternative local notions of Harsanyi consistency. The main differences between our approach and Feinberg's are the following:

- Feinberg has a dual knowledge/belief framework (his results are about what is commonly *known* about individuals' beliefs, rather than about what is commonly believed)[33].
- If the Truth Axiom is postulated for individual beliefs (i.e. beliefs have "full support") then Feinberg's common knowledge operator and our common belief operator coincide and therefore Feinberg's result (1995, Proposition 2, p.5) coincides with our Theorem 1 (concerning our Theorem 2, note that when the Truth Axiom holds, **HQC** = **SHC**, hence Theorem 2 collapses into Theorem 1).
- When the Truth Axiom does not hold (individual beliefs do not have "full support") then Feinberg's definition of Agreement, if formulated locally, would be equivalent to $\mathbf{HQC}^{\mathbf{K}}$ which is obtained by replacing the common belief operator B_* in our definition of **HQC** with the common knowledge operator (where the knowledge operator of individual i is obtained from the type partition $\{\| p_i = p_{i,\omega} \| : \omega \in \Omega\}$). However, $\mathbf{HQC}^{\mathbf{K}}$ is substantially weaker than **HQC**, as the example of Figure 4 shows.

---

[32] Indeed, his goal seems to have been to establish a converse to the Agreement Theorem in Aumann's (1976) original set-up: "The purpose of this study is to investigate the converse situation, i.e. whether the lack of a common prior implies the existence of common knowledge disagreement" (1995, p. 2).

[33] So does Samet (1996a). Samet (1996b) assumes the Truth Axiom for beliefs, thus identifying beliefs and knowledge.

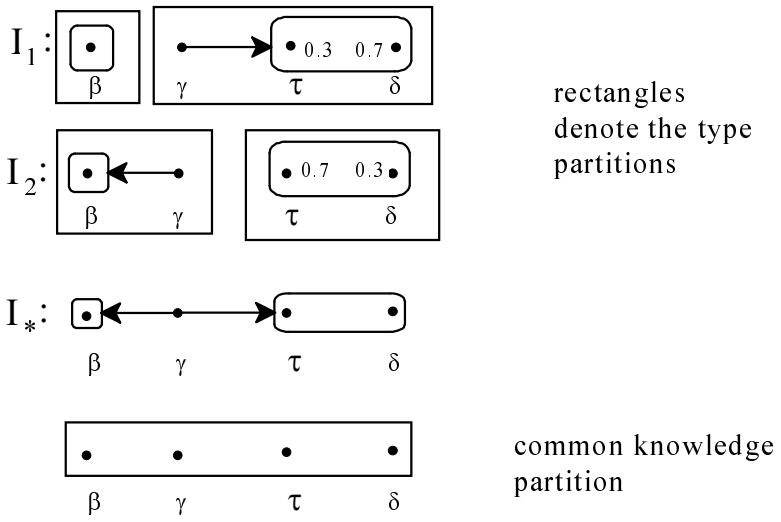

Figure 4

Here, according to Feinberg's definition, there is Agreement everywhere (in particular at τ) whereas, according to our definition, at τ individuals 1 and 2 agree to disagree. Note that here behavioral implications would be determined by the belief system $\{\tau, \delta\}$. Hence **HQC** is the relevant concept, not $\mathbf{HQC}^{\mathbf{K}}$.[34] Note also that in this example $\tau \in \mathbf{T} \cap B_*\mathbf{T}$, which shows that the gap between $\mathbf{HQC}^{\mathbf{K}}$ and **HQC** cannot be bridged by local assumptions on the *belief* hierarchies in the manner of Theorem 2.[35] Indeed the example illustrates that *any* belief hierarchy can be embedded in a knowledge/belief hierarchy that satisfies $\mathbf{HQC}^{\mathbf{K}}$ at the true state (we thank Bart Lipman for this observation).

- Finally, Feinberg defines Agreement as expectation consistency. As noted in Section 5, Agreement as "equality of beliefs" makes perfect sense under incomplete information while it is not clear that expectation consistency can be understood without reference to a prior.

[34] Of course, we do not suggest that a dual knowledge/belief framework is indispensable for his results; if anything, the analysis of this paper would seem to suggest the contrary.

[35] The analogous bridging assumption for $\mathbf{HQC}^{K}$ would be $\tau \in K_*\mathbf{T}$, that is, full support of individual beliefs.

APPENDIX A

Decision functions and belief indices

The received approach to generalizing the notion of agreement beyond Aumann's (1976) original one is based on the concept of *decision function* $D : 2^{\Omega} \to A$, where A is a set of *actions*. A decision function either is interpreted as describing individuals' behavior in response to hypothetically received information (as in Bacharach, 1985) or can be viewed as a mathematical construct (the "predicate functions" of Rubinstein and Wolinsky, 1990). Both interpretations rely on an *asymmetric information* context, for which they are designed[36]. In particular the behavioral explication relies heavily on the interpretation of events as real contingencies (i.e. as information)[37]. That decision functions (however interpreted) are entities of a rather different kind than belief indices clearly emerges from the fact that the assignment of *equal* decision functions to both individuals is, and needs to be, justified by an *assumption* of like-mindedness, i.e. in Bayesian terms, by their having a common prior. By contrast, identity of the belief index in the definition of Comprehensive Agreement simply reflects the requirement that agreement (or disagreement) be about *one* thing. Indeed, a decision function is naturally viewed as a derived entity describing the value of a belief index evaluated at a posterior that is obtained from a (fixed) prior and a (variable) information set. Recall that a decision function D satisfies the *Sure-Thing Principle* (STP) if, whenever $\{P_k\}_{k\in K}$ is a partition of E and $D(P_k) = a$ for all $k \in K$, then $D(E) = a$ (cf. Bacharach, 1985). As shown by the following two propositions, a decision function based on a *proper* belief index (cf. Definition 3) satisfies the STP while an improper one generally does not.

PROPOSITION A.1. Fix an arbitrary $\mu \in \Delta(\Omega)$ with full support (i.e. $\mu(\omega) > 0$, $\forall\, \omega \in \Omega$). Given a belief index $f : \Delta(\Omega) \to X$ define the decision procedure $D_f^{\mu} : 2^{\Omega} \to X$ as follows: $\forall\, E \subseteq \Omega$, $D_f^{\mu}(E) = f(\mu(\cdot \mid E))$. If f is proper, then D_f^{μ} satisfies the Sure Thing Principle.

---

[36] In such a context, both Bacharach (1985) and Rubinstein and Wolinsky (1990) emphasize the possibility of non-Bayesian generalizations of Aumann's (1976) result.

[37] However, even under asymmetric information, a behavioral interpretation of decision functions is not unproblematic, as pointed out by Moses and Nachum (1990).

Proof. Let $\mu \in \Delta(\Omega)$ have full support. Choose an arbitrary $f \in \mathcal{F}$ and let $D_f^\mu : 2^\Omega \to X$ be given by $D_f^\mu(\mathrm{E}) = f(\mu(\cdot \mid \mathrm{E}))$. We want to show that D_f^μ satisfies the Sure Thing Principle. Let $\varnothing \neq \mathrm{E} \subseteq \Omega$ be an arbitrary event and $\{\mathrm{E}_1, ..., \mathrm{E}_m\}$ be a partition of E. Suppose that $D_f^\mu(\mathrm{E}_k) = x_0$ for every k = 1, ..., m. Let $X_0 = f^{-1}(x_0)$ and $\mu_k = \mu(\cdot \mid \mathrm{E}_k)$. Then $\mu_k \in X_0$ for all k = 1, ..., m. Since $f \in \mathcal{F}$, X_0 is a convex set. It follows that, since $\mu(\cdot \mid \mathrm{E}) = \frac{1}{\mu(\mathrm{E})} \sum_{k=1}^{m} \mu(E_k)\ \mu(\cdot \mid E_k)$, $\mu(\cdot \mid \mathrm{E}) \in X_0$. Hence, $D_f^\mu(\mathrm{E}) = f(\mu(\cdot \mid \mathrm{E})) = x_0$. ■

To obtain the converse of Proposition A.1, let $\Xi = \{0,1\}$. Let $\mathrm{marg}_\Omega : \Delta(\Omega \times \Xi) \to \Delta(\Omega)$ be the function that associates with every $\overline{\mu} \in \Delta(\Omega \times \Xi)$ the marginal of $\overline{\mu}$ on Ω: $\forall\ \omega \in \Omega$, $\mathrm{marg}_\Omega(\overline{\mu})(\omega) = \sum_{\xi \in \Xi} \overline{\mu}(\omega, \xi)$. Given a belief index $f : \Delta(\Omega) \to X$ define $\bar{f} : \Delta(\Omega \times \Xi) \to X$ as: $\bar{f} = f \circ \mathrm{marg}_\Omega$. Finally, given $\overline{\mu} \in \Delta(\Omega \times \Xi)$ and $f : \Delta(\Omega) \to X$ define the decision procedure $\overline{D}_{\bar{f}}^{\overline{\mu}} : 2^{\Omega \times \Xi} \to X$ as follows:

$\forall\ \overline{E} \subseteq \Omega \times \Xi$, $\overline{D}_{\bar{f}}^{\overline{\mu}}(\overline{E}) = \bar{f}(\overline{\mu}(\cdot \mid \overline{E}))$.

PROPOSITION A.2. Suppose that, for every $\overline{\mu} \in \Delta(\Omega \times \Xi)$ that has full marginal support on Ξ (that is, for all $\xi \in \Xi$, $\overline{\mu}(\Omega \times \{\xi\}) > 0$), $\overline{D}_{\bar{f}}^{\overline{\mu}}$ satisfies the Sure Thing Principle. Then f is a proper belief index.

Proof. Let $\mathrm{p}, \mathrm{q} \in \Delta(\Omega)$, $f : \Delta(\Omega) \to X$ and $x \in X$ be such that $f(\mathrm{p}) = f(\mathrm{q}) = x$. Fix an arbitrary $a \in [0,1]$. We want to show that $f(a\mathrm{p} + (1-a)\mathrm{q}) = x$. Let $\overline{\Omega} = \Omega \times \{0,1\}$ and define $\overline{p}, \overline{q}, \overline{\mu} \in \Delta(\overline{\Omega})$ as follows: $\forall\ \omega \in \Omega$, $\overline{p}(\omega,0) = \mathrm{p}(\omega)$, $\overline{p}(\omega,1) = 0$, $\overline{q}(\omega,0) = 0$, $\overline{q}(\omega,1) = q(\omega)$, $\overline{\mu} = a\overline{p} + (1-a)\overline{q}$. Let $\overline{E}_0 = \Omega \times \{0\}$ and $\overline{E}_1 = \Omega \times \{1\}$. Then $\mathrm{marg}_\Omega(\overline{\mu}(\cdot \mid \overline{E}_0)) = \mathrm{p}$ and $\mathrm{marg}_\Omega(\overline{\mu}(\cdot \mid \overline{E}_1)) = \mathrm{q}$. Let $\overline{f} = f \circ \mathrm{marg}_\Omega$ and $\overline{D}_{\bar{f}}^{\overline{\mu}} : 2^{\Omega \times \Xi} \to X$ be defined by $\overline{D}_{\bar{f}}^{\overline{\mu}}(\overline{E}) = \bar{f}(\overline{\mu}(\cdot \mid \overline{E}))$. Then $\overline{D}_{\bar{f}}^{\overline{\mu}}(\overline{E}_0) = f(\mathrm{p}) = x$ and $\overline{D}_{\bar{f}}^{\overline{\mu}}(\overline{E}_1) = f(\mathrm{q}) = x$. Since $\{\overline{E}_0, \overline{E}_1\}$ is a partition of $\overline{\Omega}$ and $\overline{D}_{\bar{f}}^{\overline{\mu}}$ satisfies the Sure Thing Principle, $\overline{D}_{\bar{f}}^{\overline{\mu}}(\overline{\Omega}) = x$. But $\overline{D}_{\bar{f}}^{\overline{\mu}}(\overline{\Omega}) = \bar{f}(\overline{\mu}) = f(\mathrm{marg}_\Omega(\overline{\mu})) = f(a\mathrm{p} + (1-a)\mathrm{q})$. ■

The following corollary shows that properness is a necessary requirement for the notion of Comprehensive Agreement if the latter is to be an implication of Strong Harsanyi Consistency.

COROLLARY A.1. Let $f : \Delta(\Omega) \to X$ be an *im*proper belief index (that is, $f \notin \mathcal{F}$). Then there exists a Bayesian model with state space $\Omega \times \{0,1\}$ for which **SHC** = $\Omega \times \{0,1\}$ and $\bar{f}$ **-Agree** = $\varnothing$.

Proof. Let $f : \Delta(\Omega) \to X$ be an improper belief index, that is, $f \notin \mathcal{F}$. Then there exist p,q$\in\Delta(\Omega)$ with p ≠ q, $x \in X$ and $a \in (0,1)$ such that f(p) = f(q) = x and f(ap + (1 – a)q) = y ≠ x. Construct the following Bayesian model with state space $\overline{\Omega} = \Omega \times \{0,1\}$: $\forall\ \overline{\omega} \in \overline{\Omega}$, $\overline{p}_{1,\overline{\omega}} = \overline{\mu}$, $\forall\ \overline{\omega} \in \overline{E}_0$, $\overline{p}_{2,\overline{\omega}} = \overline{p}$ and $\forall\ \overline{\omega} \in \overline{E}_1$, $\overline{p}_{2,\overline{\omega}} = \overline{q}$, where $\overline{p}$, $\overline{q}$ and $\overline{\mu}$ are as in the proof of Proposition A.2. In this model **SHC** = $\overline{\Omega}$ and $B_*(\| \overline{f}_1 = y \| \cap \| \overline{f}_2 = x \|) = \overline{\Omega}$. Since x ≠ y, $\overline{f}$ **-Agree** = $\varnothing$. ■

From the point of view of the present paper, the generalizations of Aumann's (1976) result by Bacharach (1985), Cave (1983) and Rubinstein and Wolinsky (1990) can be translated to a Bayesian incomplete information context as follows.

PROPOSITION A.3. Let $\mu \in \Delta(\Omega)$ and f be any belief index such that D_f^μ (cf. Proposition A.1) satisfies the Sure Thing Principle. If $\tau \in \mathbf{HQC}_\mu \cap \mathbf{T} \cap B_*\mathbf{T}$ then $\tau \in f$ **-Agree**.

Proof. Let $\tau \in \mathbf{HQC}_\mu \cap \mathbf{T} \cap B_*\mathbf{T}$. Then by Theorem 2, $\tau \in \mathbf{SHC}_\mu$. Hence, by Lemmas C.2 and C.6 (see Appendix C), the reduced model with $\Omega = I_*(\tau) = I_*(\tau) \cup \{\tau\}$ is a partition model. Thus, by a standard application of Bacharach's result (1985), if $\tau \in B_*(\| D^\mu_{f,1} = a \| \cap \| D^\mu_{f,2} = b \|)$ then $a = b$, where $\| D^\mu_{f,1} = a \| = \{\omega \in \Omega : D^\mu_f(I_i(\omega)) = a\}$ and similarly for $\| D^\mu_{f,2} = b \|$. It follows from the definition of D_f^μ that $\tau \in f$ **-Agree**. ■

The sufficiency part of Theorem 1 (Section 3) drops the Truth Axiom (i.e. the clause $\tau \in \mathbf{T} \cap B_*\mathbf{T}$), and – more importantly – replaces the Sure-Thing-Principle requirement on the *doubly non-primitive* entity D_f^μ by that of properness of f.

APPENDIX B

An intermediate notion: Harsanyi Consistency

As noted in Section 4, Harsanyi Quasi Consistency is a very weak property, in particular, it allows the "common prior" to assign zero probability to the true beliefs of *all* the individuals (see Figure 3). On the other hand, Strong Harsanyi Consistency imposes an external requirement of conformity between individual beliefs and facts of nature and therefore cannot be seen as an internal consistency assumption on beliefs. An intermediate notion, Harsanyi Consistency, requires only that the true beliefs of each individual be given positive probability by the common prior.[38]

DEFINITION B.1. For every $\mu \in \Delta(\Omega)$, let $\mathbf{HC}_\mu$ (for Harsanyi Consistency with respect to the "prior" μ) be the following event: $\alpha \in \mathbf{HC}_\mu$ if and only if

(1) $\alpha \in \mathbf{HQC}_\mu$, and

(2) $\forall i \in N,\ \mu(\|p_i = p_{i,\alpha}\|) > 0$.

Furthermore, let $\mathbf{HC} = \bigcup_{\mu \in \Delta(\Omega)} \mathbf{HC}_\mu$.

The following result is proved in Bonanno and Nehring (1996).

PROPOSITION B.1. (i) $\mathbf{HQC} = \neg B_* \neg \mathbf{HC}$; (ii) $\mathbf{HC} = B_* \mathbf{SHC} = B_* \mathbf{HC}$.

From (ii) of Proposition B.1 we get that Harsanyi Consistency holds at a state if and only if at that state everybody believes that it holds ($\mathbf{HC} = \bigcap_{i \in N} B_i \mathbf{HC}$). It follows that a necessary and sufficient condition for Harsanyi Consistency *not* to be satisfied is that (at least) one individual fail to believe that it holds. In other words, any failure of Harsanyi Consistency can be located in *one* individual's belief hierarchy alone. Thus, Proposition B.1 may throw some light on Gul's (1996) puzzlement in this regard.

From the above proposition and Theorem 2 the following corollary is obtained.

COROLLARY B.1. $\mathbf{HC} = B_* \mathbf{CA} \cap B_* \mathbf{T}_{CB}$.

---

[38] Note, however, that Definition B.1 is consistent with the local common prior assigning zero probability to the *true state*.

APPENDIX C: proofs

DEFINITION C.1. For every $\mu \in \Delta(\Omega)$, let $\mathbf{COM}_\mu$ (for Compatibility with the "prior" μ) be the following event: $\alpha \in \mathbf{COM}_\mu$ if and only if, $\forall\, i \in N$, $\mu(\|p_i = p_{i,\alpha}\|) > 0$ implies $p_{i,\alpha} = \mu(\cdot \mid \|p_i = p_{i,\alpha}\|)$. Thus (cf. Definition 2) $\alpha \in \mathbf{HQC}_\mu$ if and only if (1) $\alpha \in B_*\mathbf{COM}_\mu$ and (2) $\mu(I_*(\alpha)) > 0$.

LEMMA C.1. Let $\beta \in \mathbf{COM}_\mu \cap \text{supp}\ \mu$. Then, $\forall\, i \in N$, $\beta \in I_i(\beta)$ and $I_i(\beta) \subseteq \text{supp}\ \mu$.

Proof. Let $\beta \in \mathbf{COM}_\mu \cap \text{supp}\ \mu$. Fix an arbitrary $i \in N$. Since $\beta \in \|p_i = p_{i,\beta}\|$ and $\mu(\beta) > 0$, $\mu(\|p_i = p_{i,\beta}\|) > 0$. Hence, since $\beta \in \mathbf{COM}_\mu$, $p_{i,\beta} = \mu(\cdot \mid \|p_i = p_{i,\beta}\|)$. Thus $p_{i,\beta}(\beta) > 0$, that is, $\beta \in I_i(\beta)$. Furthermore, fix an arbitrary $\gamma \in I_i(\beta)$. Then $p_{i,\beta}(\gamma) > 0$ and therefore, since $p_{i,\beta} = \mu(\cdot \mid \|p_i = p_{i,\beta}\|)$, $\mu(\gamma) > 0$. Thus $I_i(\beta) \subseteq \text{supp}\ \mu$. ■

LEMMA C.2. Let $\mu \in \Delta(\Omega)$, $\alpha \in \mathbf{HQC}_\mu$ and $S = I_*(\alpha) \cap \text{supp}\ \mu$. Then, for every $i \in N$, $\{I_i(\omega) : \omega \in S\}$ is a partition of S.

Proof. Fix an arbitrary $i \in N$. By transitivity and euclideanness of I_i (cf. Footnote 7), any two elements of $\{I_i(\omega) : \omega \in S\}$ are disjoint. Thus we only need to show that $\bigcup_{\omega \in S} I_i(\omega) = S$. First we show that $S \subseteq \bigcup_{\omega \in S} I_i(\omega)$. Fix an arbitrary $\beta \in S$. Since $\alpha \in \mathbf{HQC}_\mu$, $\alpha \in B_*\mathbf{COM}_\mu$ (cf. Definition C.1) hence $\beta \in \mathbf{COM}_\mu$. Hence by Lemma C.1 $\beta \in I_i(\beta)$. Thus $S = \bigcup_{\omega \in S} \{\omega\} \subseteq \bigcup_{\omega \in S} I_i(\omega)$. Next we show that $\bigcup_{\omega \in S} I_i(\omega) \subseteq S$. Fix an arbitrary $\beta \in S$. By definition of I_*, $I_i(\beta) \subseteq I_*(\beta)$ and by transitivity of I_*, $I_*(\beta) \subseteq I_*(\alpha)$. Thus $I_i(\beta) \subseteq I_*(\alpha)$. Since $\beta \in \mathbf{COM}_\mu$ (shown above), by Lemma C.1, $I_i(\beta) \subseteq \text{supp}\ \mu$. Hence $I_i(\beta) \subseteq S$. Thus $\bigcup_{\omega \in S} I_i(\omega) \subseteq S$. ■

REMARK C.1. A possibility correspondence $I : \Omega \to 2^\Omega$ is *secondary reflexive* if $\forall\, \alpha, \beta \in \Omega$, $\beta \in I(\alpha)$ implies $\beta \in I(\beta)$. Secondary reflexivity is implied by euclideanness (cf. Footnote 7). Hence for every $i \in N$, I_i is secondary reflexive. It follows from the definition of I_* that I_* is also secondary reflexive.

LEMMA C.3 (Convex hull characterization of **HQC**). Let $\alpha \in \Omega$ and $\mu \in \Delta(\Omega)$ be such that $\text{supp}\ \mu \subseteq I_*(\alpha)$. Then

$$\text{(i) } \alpha \in \mathbf{HQC}_\mu \text{ if and only if } \mu \in \bigcap_{i \in N} \text{co}\, P_i(\alpha),$$

where $\text{co}\, P_i(\alpha)$ is the convex hull of $P_i(\alpha) = \{p_{i,\omega} \in \Delta(\Omega) : \omega \in I_*(\alpha)\}$, and, therefore,

$$\text{(ii) } \alpha \in \mathbf{HQC} \text{ if and only if } \bigcap_{i \in N} \text{co}\, P_i(\alpha) \neq \varnothing.$$

Proof. (i) Let $\alpha \in \mathbf{HQC}_\mu$. By Lemma C.2, since $\text{supp}\ \mu \subseteq I_*(\alpha)$, $\{I_i(\omega) : \omega \in \text{supp}\ \mu\}$ is a partition of $\text{supp}\ \mu$. The sets $\{I_i(\omega)\}_{\omega \in \text{supp}\ \mu}$ are of the form $\|p_i = q\|$ for appropriate $q \in P_i(\alpha)$. It

follows from this and Definition 2 that $\mu = \sum_{q \in P_i(\alpha)} q\, \mu(\|p_i = q\|)$ hence that $\mu \in \bigcap_{i \in N} \text{co}\, P_i(\alpha)$.

Conversely, let $\mu \in \bigcap_{i \in N} \text{co}\, P_i(\alpha)$. Fix an arbitrary $i \in N$ and let $\mu = \sum_{q \in P_i(\alpha)} a_q^i\, q$ for appropriate $a_q^i \geq 0$ adding up to 1. Consider any $q \in P_i(\alpha)$ such that $\mu(\|p_i = q\|) > 0$, i.e. (since the supports of any two elements of $P_i(\alpha)$ are disjoint) $a_q^i > 0$. Then, by the dijointness of the supports of the elements of $P_i(\alpha)$, $q = \mu(\cdot \mid \text{supp } q) = \mu(\cdot \mid \|p_i = q\|)$. The latter equality follows from the fact that $\omega \in \|p_i = q\| \setminus \text{supp } q$ implies $\mu(\omega) = 0$.

(ii) Let $\alpha \in \mathbf{HQC}$. Then there exists a $\nu \in \Delta(\Omega)$ such that $\alpha \in \mathbf{HQC}_\nu$. Let $\mu = \nu(\cdot \mid I_*(\alpha))$. Then $\alpha \in \mathbf{HQC}_\mu$ and supp $\mu \subseteq I_*(\alpha)$. Hence by part (i) $\mu \in \bigcap_{i \in N} \text{co}\, P_i(\alpha)$. ■

PROOF OF THEOREM 1 (the proof is illustrated in Figure C.1). Suppose that $\alpha \notin$ ***f*-Agree** for some $f \in \mathcal{F}$, that is, there exist $x_1, x_2 \in X$ such that $x_1 \neq x_2$ and $\alpha \in B_*(\|f_1 = x_1\| \cap \|f_2 = x_2\|)$, that is, $I_*(\alpha) \subseteq \|f_1 = x_1\| \cap \|f_2 = x_2\|$. Then $f^{-1}(x_1) \supseteq P_1(\alpha)$ and $f^{-1}(x_2) \supseteq P_2(\alpha)$. Since $x_1 \neq x_2$, $f^{-1}(x_1) \cap f^{-1}(x_2) = \varnothing$. By definition of $\mathcal{F}$, $f^{-1}(x_1)$ and $f^{-1}(x_2)$ are convex sets. Hence $f^{-1}(x_1) \supseteq \text{co}\, P_1(\alpha)$ and $f^{-1}(x_2) \supseteq \text{co}\, P_2(\alpha)$. Thus $\text{co}\, P_1 \cap \text{co}\, P_2 = \varnothing$. Hence, by Lemma C.3, $\alpha \notin \mathbf{HQC}$.

Conversely suppose that $\alpha \notin \mathbf{HQC}$. Then, by Lemma C.3, $\text{co}\, P_1(\alpha) \cap \text{co}\, P_2(\alpha) = \varnothing$. Since $\text{co}\, P_i(\alpha)$ is convex and compact, by the separating hyperplane theorem (see, for example, Border, 1985, p. 11), there exists a $c \in \mathbb{R}^{|\Omega|}$ with $c \neq 0$ and an $a \in \mathbb{R}$ such that, for all $p \in \text{co}\, P_1(\alpha)$ and $q \in \text{co}\, P_2(\alpha)$, $c \cdot p > a > c \cdot q$. Let $f: \Delta(\Omega) \to \{0,1\}$ be given by: $f(p) = \begin{cases} 1 & \text{if } p \cdot c \geq a \\ 0 & \text{otherwise} \end{cases}$. Then $f \in \mathcal{F}$ and $I_*(\alpha) \subseteq \|f_1 = 1\| \cap \|f_2 = 0\|$. Thus $\alpha \notin$ ***f*-Agree**. ■

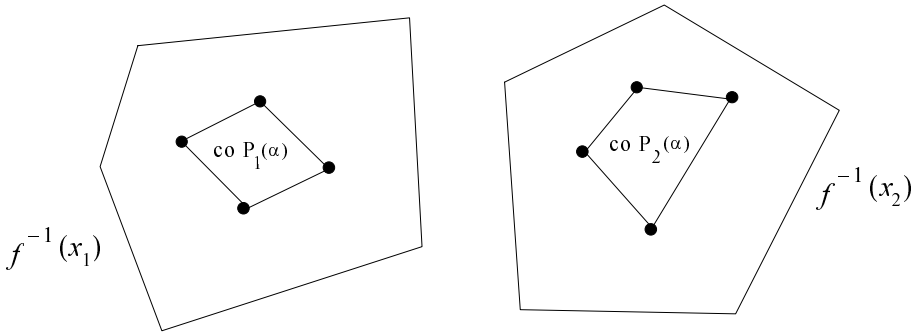

Figure C.1

PROOF OF PROPOSITION 1. Define $\mathcal{F}_2$**-Agree** as $\bigcap_{f \in \mathcal{F}_2}$ ***f*-Agree** and recall that **CA** = $\bigcap_{f \in \mathcal{F}}$ ***f*-Agree**. (i) That **CA** $\subseteq$ $\mathcal{F}_2$**-Agree** is clear, since $\mathcal{F} \supseteq \mathcal{F}_2$. Thus we only need to show

that the violation of Comprehensive Agreement implies the violation of $\mathcal{J}_2$-Agreement. The second part of the proof of Theorem 1 shows that if $\alpha \notin$ **HQC**, then there exists an $f \in \mathcal{J}_2$ such that $\alpha \notin f$-**Agree.** But, by Theorem 1, **HQC** = **CA.**

(ii) Let $Y : \Omega \to \mathbb{R}$ and $f : \Delta(\Omega) \to \{0,1\}$ be such that $f(p) = 1$ if and only if $\sum_{\omega\in\Omega} Y(\omega)\, p(\omega) \geq 0$. We want to show that $f \in \mathcal{J}_2$, that is, (1) $f^{-1}(0)$ and $f^{-1}(1)$ are convex sets and (2) $f^{-1}(1)$ is a closed set. Let $p,q \in f^{-1}(0)$ and $\lambda \in [0,1]$. Let $r = \lambda p + (1-\lambda) q$. We want to show that $f(r) = 0$, that is, $\sum_{\omega\in\Omega} Y(\omega) r(\omega) < 0$. Now, $\sum_{\omega\in\Omega} Y(\omega) r(\omega) = \sum_{\omega\in\Omega} Y(\omega)[\lambda\, p(\omega) + (1-\lambda) q(\omega)]$ $= \lambda\, a + (1-\lambda) b$, where $a = \sum_{\omega\in\Omega} Y(\omega)\, p(\omega)$ and $b = \sum_{\omega\in\Omega} Y(\omega)\, q(\omega)$. By hypothesis, $a < 0$ and $b < 0$. Hence $\lambda a + (1-\lambda) b < 0$, that is, $f(r) = 0$. The proof that $f^{-1}(1)$ is convex is similar. The proof that $f^{-1}(1)$ is closed is straightforward.

Conversely, let $f : \Delta(\Omega) \to \{0,1\}$ belong to $\mathcal{J}_2$. We want to show that there exists a $Y : \Omega \to \mathbb{R}$ such that, for all $p \in \Delta(\Omega)$, $f(p) = 1$ if and only if $\sum_{\omega\in\Omega} Y(\omega)\, p(\omega) \geq 0$. Since $f^{-1}(0)$ and $f^{-1}(1)$ are disjoint convex sets, by the separating hyperplane theorem, there exists a $\mathbf{c} \in \mathbb{R}^{|\Omega|}$ with $\mathbf{c} \neq 0$ and an $a \in \mathbb{R}$ such that $f^{-1}(1) \subseteq \{y \in \mathbb{R}^{|\Omega|} : y \cdot \mathbf{c} \geq a\}$ and $f^{-1}(0) \subseteq \{z \in \mathbb{R}^{|\Omega|} : z \cdot \mathbf{c} < a\}$. Let $\nu : \Omega \to \{1, 2, ..., |\Omega|\}$ be a one-to-one mapping (a numbering of the elements of Ω) and define $Y : \Omega \to \mathbb{R}$ as follows: $Y(\omega) = \mathbf{c}_{\nu(\omega)} - a$ where $\mathbf{c}_{\nu(\omega)}$ is the $\nu(\omega)$th coordinate of $\mathbf{c}$. Then, for every $p \in \Delta(\Omega)$, $\sum_{\omega\in\Omega} Y(\omega)\, p(\omega) = \mathbf{y}_p \cdot \mathbf{c} - \mathbf{y}_p \cdot \mathbf{a} = \mathbf{y}_p \cdot \mathbf{c} - a \geq 0$ if and only if $p \in f^{-1}(1)$, where $\mathbf{y}_p \in \mathbb{R}^{|\Omega|}$ is the vector whose jth coordinate is $p(\omega)$ for that ω for which $\nu(\omega) = j$ and $\mathbf{a} \in \mathbb{R}^{|\Omega|}$ is the vector in which every coordinate is equal to a. ■

LEMMA C.4. Let $\mu \in \Delta(\Omega)$. If $\alpha \in \mathbf{HQC}_\mu \cap \text{supp}\ \mu$ then, $\forall\, i \in N$, $\alpha \in I_i(\alpha)$. Thus (cf. Footnote 23) $\mathbf{HQC}_\mu \cap \text{supp}\ \mu \subseteq \mathbf{T}$.

Proof. Let $\alpha \in \mathbf{HQC}_\mu \cap \text{supp}\ \mu$ and fix an arbitrary $i \in N$. Since $\alpha \in \|p_i = p_{i,\alpha}\|$, $\mu(\|p_i = p_{i,\alpha}\|) > 0$. Choose an arbitrary $\beta \in I_i(\alpha)$. Then $p_{i,\beta} = p_{i,\alpha}$ and therefore $\mu(\|p_i = p_{i,\beta}\|) > 0$. Since $I_i(\alpha) \subseteq I_*(\alpha)$, $\beta \in I_*(\alpha)$. Hence, since $\alpha \in B_*\mathbf{COM}_\mu$ (cf. Definition C.1), $\beta \in \mathbf{COM}_\mu$. Thus $p_{i,\beta} = \mu(\cdot \mid \|p_i = p_{i,\beta}\|)$. Since $\mu(\alpha) > 0$, $p_{i,\beta}(\alpha) > 0$. From this and the equality $p_{i,\beta} = p_{i,\alpha}$ we get $p_{i,\alpha}(\alpha) > 0$, that is, $\alpha \in I_i(\alpha)$. ■

LEMMA C.5. Let $\mu \in \Delta(\Omega)$. Then $\mathbf{HQC}_\mu \cap \text{supp}\ \mu \subseteq B_*\left(\mathbf{HQC}_\mu \cap \text{supp}\ \mu\right)$.

Proof. Fix $\mu \in \Delta(\Omega)$. In view of the definition of B_*, it is sufficient to show that

$$\forall\, i \in N, \qquad \mathbf{HQC}_\mu \cap \text{supp}\ \mu \subseteq B_i\left(\mathbf{HQC}_\mu \cap \text{supp}\ \mu\right). \tag{C.1}$$

Let $\alpha \in \mathbf{HQC}_\mu \cap \text{supp}\ \mu$ and fix an arbitrary $i \in N$. We want to show that $I_i(\alpha) \subseteq \mathbf{HQC}_\mu \cap \text{supp}\ \mu$. Since $\alpha \in \mathbf{HQC}_\mu$, $\alpha \in B_*\mathbf{COM}_\mu$ (cf. Definition C.1). By definition of B_*, $B_*\mathbf{COM}_\mu \subseteq B_i B_*\mathbf{COM}_\mu$. Thus $\alpha \in B_i B_*\mathbf{COM}_\mu$, that is,

$$I_i(\alpha) \subseteq B_*\mathbf{COM}_\mu. \tag{C.2}$$

By Lemma C.4, $\alpha \in I_i(\alpha)$ and by definition of I_*, $I_i(\alpha) \subseteq I_*(\alpha)$. Thus $\alpha \in I_*(\alpha)$ and, since $\alpha \in B_*\mathbf{COM}_\mu$, $\alpha \in \mathbf{COM}_\mu$. Hence $\alpha \in \mathbf{COM}_\mu \cap$ supp μ. It follows from Lemma C.1 that

$$I_i(\alpha) \subseteq \text{supp } \mu. \tag{C.3}$$

From (C.2) and (C.3) we get that $I_i(\alpha) \subseteq B_*\mathbf{COM}_\mu \cap$ supp μ. Fix an arbitrary $\beta \in I_i(\alpha)$. By (C.3), $\mu(\beta) > 0$. By Secondary Reflexivity of I_i (cf. Remark C.1), $\beta \in I_i(\beta)$. By definition of I_*, $I_i(\beta) \subseteq I_*(\beta)$. Thus $\mu(I_*(\beta)) > 0$. By (C.2), $\beta \in B_*\mathbf{COM}_\mu$. Hence (cf. Definition C.1) $\beta \in \mathbf{HQC}_\mu$. Thus

$$I_i(\alpha) \subseteq \mathbf{HQC}_\mu. \tag{C.4}$$

The conjunction of (C.4) and (C.3) yields (C.1). ■

COROLARY C.1. Let $\alpha \in \mathbf{HQC}_\mu \cap$ supp μ. Then, $\forall\, \omega \in I_*(\alpha) \cup \{\alpha\}$, $\mu(\omega) > 0$.

LEMMA C.6. Let $\alpha \in \mathbf{T} \cap B_*\mathbf{T}$. Then, $\forall\, \beta \in I_*(\alpha)$, $I_*(\alpha) = I_*(\beta)$.

Proof. Let $\alpha \in \mathbf{T} \cap B_*\mathbf{T}$ and fix an arbitrary $\beta \in I_*(\alpha)$. That $I_*(\beta) \subseteq I_*(\alpha)$ follows from transitivity of I_*. Thus we only need to show that $I_*(\alpha) \subseteq I_*(\beta)$. By definition of I_*, since $\beta \in I_*(\alpha)$, there exists a sequence $\langle j_1, \ldots j_m \rangle$ in N and a sequence $\langle \eta_0, \eta_1, \ldots, \eta_m \rangle$ in Ω such that: $\eta_0 = \alpha$, $\eta_m = \beta$ and, for every $k = 1, \ldots, m$, $\eta_k \in I_{j_k}(\eta_{k-1})$. By transitivity and euclideanness of I_{j_1} (cf. Footnote 7), since $\eta_1 \in I_{j_1}(\alpha)$, $I_{j_1}(\alpha) = I_{j_1}(\eta_1)$. Since $\alpha \in \mathbf{T}$, $\alpha \in I_{j_1}(\alpha)$. Thus $\alpha \in I_{j_1}(\eta_1)$. Since $\eta_1 \in I_*(\alpha)$ and $\alpha \in B_*\mathbf{T}$, $\eta_1 \in I_{j_2}(\eta_1)$. Since $\eta_2 \in I_{j_2}(\eta_1)$, by transitivity and euclideanness of I_{j_2}, $I_{j_2}(\eta_1) = I_{j_2}(\eta_2)$. Thus $\eta_1 \in I_{j_2}(\eta_2)$. Repeating this argument m times, we get that, $\forall\, k = 0, \ldots, m$, $\eta_k \in I_{j_{k+1}}(\eta_{k+1})$. Thus, by definition of I_*, $\eta_0 \in I_*(\eta_m)$, that is, $\alpha \in I_*(\beta)$. Hence by transitivity of I_* $I_*(\alpha) \subseteq I_*(\beta)$. ■

LEMMA C.7 (see Bonanno and Nehring, 1997)**.** For every event $E \subseteq \Omega$, $B_*E = B_*B_*E$.

REMARK C.2. It is well-known that if $B : 2^\Omega \to 2^\Omega$ is the belief operator obtained for a possibility correspondence $I : \Omega \to 2^\Omega$ (that is, $\forall\, E \in 2^\Omega$, $BE = \{\omega \in \Omega : I(\omega) \subseteq E\}$) then it satisfies the following properties. Necessity: $B\Omega = \Omega$; Conjunction: $B(E \cap F) = BE \cap BF$; Monotonicity: if $E \subseteq F$ then $BE \subseteq BF$.

PROOF OF THEOREM 2. First we show that $\mathbf{SHC} \subseteq \mathbf{HQC} \cap \mathbf{T} \cap B_*\mathbf{T}$. Let $\alpha \in \mathbf{SHC}$. Then there exists a $\mu \in \Delta(\Omega)$ such that $\alpha \in \mathbf{SHC}_\mu$. Since $\mathbf{SHC}_\mu = \mathbf{HQC}_\mu \cap$ supp μ, by Lemma C.4 $\mathbf{SHC}_\mu \subseteq \mathbf{T}$. By Monotonicity of B_* (cf. Remark C.2), $B_*\mathbf{SHC}_\mu \subseteq B_*\mathbf{T}$. By Lemma C.5, $\mathbf{SHC}_\mu \subseteq B_*\mathbf{SHC}_\mu$. Thus $\mathbf{SHC}_\mu \subseteq B_*\mathbf{T}$. Hence $\alpha \in \mathbf{HQC}_\mu \cap \mathbf{T} \cap B_*\mathbf{T} \subseteq \mathbf{HQC} \cap \mathbf{T} \cap B_*\mathbf{T}$.

For the converse, let $\alpha \in \mathbf{HQC} \cap \mathbf{T} \cap B_*\mathbf{T}$. Then there exists a $\mu \in \Delta(\Omega)$ such that $\alpha \in \mathbf{HQC}_\mu \cap \mathbf{T} \cap B_*\mathbf{T}$. Since $\alpha \in \mathbf{HQC}_\mu$, $\mu(I_*(\alpha)) > 0$, that is, there exists a $\beta \in I_*(\alpha)$ such that $\mu(\beta) > 0$. Furthermore, $\alpha \in B_*\mathbf{COM}_\mu$ (cf. Definition C.1). By Lemma C.7, $B_*\mathbf{COM}_\mu = B_*B_*\mathbf{COM}_\mu$. Thus $I_*(\alpha) \subseteq B_*\mathbf{COM}_\mu$. Hence $\beta \in B_*\mathbf{COM}_\mu$. By secondary reflexivity of I_* (cf. Remark C.1), $\beta \in I_*(\beta)$. Thus $\mu(I_*(\beta)) > 0$. Hence $\beta \in \mathbf{HQC}_\mu \cap$ supp $\mu = \mathbf{SHC}_\mu$. Since $\alpha \in \mathbf{T}$, $\alpha \in I_i(\alpha)$ for all $i \in N$. Hence $\alpha \in I_*(\alpha)$. By Lemma C.6, since $\beta \in I_*(\alpha)$, $I_*(\alpha) = I_*(\beta)$. Thus $\alpha \in I_*(\beta)$. By Lemma C.5, $\mathbf{SHC}_\mu \subseteq B_*\mathbf{SHC}_\mu$. Hence $\beta \in B_*\mathbf{SHC}_\mu$, that is, $I_*(\beta) \subseteq \mathbf{SHC}_\mu$. Thus $\alpha \in \mathbf{SHC}_\mu$. ■

REFERENCES

Armbruster, W. and W. Boege (1979), Bayesian game theory, in *Game Theory and Related Topics*, ed. by O. Moeschlin and D. Pallaschke, Amsterdam, North-Holland.

Aumann, R. (1976), Agreeing to disagree, *Annals of Statistics,* 4, 1236-1239.

Aumann, R. (1987), Correlated equilibrium as an expression of Bayesian rationality, *Econometrica*, 55, 1-18.

Aumann, R. and A. Brandenburger (1995), Epistemic conditions for Nash equilibrium, *Econometrica*, 63, 1161-1180

Bacharach, M. (1985), Some extensions of a claim of Aumann in an axiomatic model of knowledge, *Journal of Economic Theory*, 37, 167-190.

Battigalli, P. (1995), Dynamic consistency and imperfect recall, mimeo, Princeton University.

Binmore, K. and A. Brandenburger (1990), Common knowledge and game theory, in K. Binmore, *Essays on the foundations of game theory,* Oxford, Basil Blackwell.

Boege, W. and T. Eisele (1979), On solutions of Bayesian games, *International Journal of Game Theory,* 8, 193-215.

Bonanno, G. (1996), On the logic of common belief, *Mathematical Logic Quarterly*, 42, 305-311.

Bonanno, G. and P. Battigalli (1997), The logic of belief persistence, *Economics and Philosophy*, 13, 39-59.

G. Bonanno and K. Nehring (1996), Making sense of the common prior assumption under incomplete information, Working Paper, University of California, Davis.

Bonanno, G. and Nehring, K. (1997), Assessing the Truth Axiom under incomplete information, Working Paper, University of California, Davis [forthcoming in *Mathematical Social Sciences*].

Bonanno, G. and Nehring, K. (1998a), On the logic and role of Negative Introspection of Common Belief, *Mathematical Social Sciences*, 35, 17-36.

Bonanno, G. and Nehring, K. (1998b), Intersubjective consistency of knowledge and beliefs, Working Paper, University of California, Davis.

Border, K. (1985), *Fixed point theorems with applications to economics and game theory,* Cambridge University Press, Cambridge.

Brandenburger, A. and E. Dekel (1993), Hierarchies of beliefs and common knowledge, *Journal of Economic Theory*, 59, 189-198.

Cave, J. (1983), Learning to agree, *Economics Letters*, 12, 147-152.

Chellas, B. (1984), *Modal logic*, Cambridge University Press, Cambridge.

Dekel, E. and F. Gul (1997), Rationality and knowledge in game theory, in: Kreps D. M. and K. F. Wallis (eds.), *Advances in Economics and Econometrics*, vol. 1, Cambridge: Cambridge University Press, pp. 87-172.

Feinberg, Y. (1995), A converse to the Agreement Theorem, Discussion Paper # 83, Center for Rationality and Interactive Decision Theory, Jerusalem.

Feinberg, Y. (1996), Characterizing the existence of a common prior via the notion of disagreement, Ph. D. thesis, Hebrew University of Jerusalem.

van Fraassen, B. (1984), Belief and the will, *Journal of Philosophy*, 81, 235-56.

Geanakoplos, J. (1989), Game theory without partitions and applications to speculation and consensus, Cowles Foundation Discussion Paper 914, Yale University.

Geanakoplos, J. (1992), Common knowledge, *Journal of Economic Perspectives*, 6, 53-82.

Geanakoplos, J. and M. Polemarchakis (1982), We can't disagree for ever, *Journal of Economic Theory*, 26, 363-390.

Goldstein, M. (1983), The prevision of a prevision, *Journal of the American Statistical Association,* 78, 817-19.

Gul, F. (1996), A comment on Aumann's Bayesian view, mimeo, Northwestern University [forthcoming in *Econometrica*].

Halpern, J. (1995), On ambiguities in the interpretation of game trees, mimeo, Almaden Res. Center.

Halpern, J. and Y. Moses (1992), A guide to completeness and complexity for modal logics of knowledge and belief, *Artificial intelligence*, 54, 319-379.

Harsanyi, J. (1967-68), Games with incomplete information played by "Bayesian players", Parts I-III, *Management Science*, 8, 159-182, 320-334, 486-502.

Lipman, B. (1995), Approximately common priors, mimeo, University of Western Ontario.

Lismont, L. and P. Mongin (1994), On the logic of common belief and common knowledge, *Theory and Decision*, 37 (1), 75-106.

Maher, P. (1993), *Betting on theories*, Cambridge University Press.

Mertens, J-F. and S. Zamir (1985), Formulation of Bayesian analysis for games with incomplete information, *International Journal of Game Theory,* 14, 1-29.

Milgrom, P. and N. Stokey (1982), Information, trade and common knowledge, *Journal of Economic Theory*, 26, 17-27.

Morris, S. (1991), The role of beliefs in economic theory, Ph. D. dissertation, Yale University.

Morris, S. (1994), Trade with heterogeneous prior beliefs and asymmetric information, *Econometrica*, 62, 1327-1347.

Morris, S. (1995), The common prior assumption in economic theory, *Economics and Philosophy*, 11, 227-53.

Moses, Y. and G. Nachum (1990), Agreeing to disagree after all, in R. Parikh (Ed.), *Theoretical aspects of reasoning about knowledge: proceedings of the third conference*, Morgan Kaufmann, San Francisco, pp. 151-168.

Nau, R. and K. McCardle (1990), Coherent behavior in noncooperative games, *Journal of Economic Theory*, 50, 424-444.

Piccione M. and A. Rubinstein (1994), On the interpretation of decision problems with imperfect recall, mimeo, Tel Aviv University.

Rubinstein, A. and A. Wolinsky (1990), On the logic of "Agreeing to Disagree" type of results, *Journal of Economic Theory*, 51, 184-193.

Samet, D. (1990), Ignoring ignorance and agreeing to disagree, *Journal of Economic Theory*, 52, 190-207.

Samet, D. (1996a), Common priors and separation of convex sets, mimeo, Tel Aviv university.

Samet, D. (1996b), Common priors and Markov chains, mimeo, Tel Aviv university.

Sebenius, J. and J. Geanakoplos (1983), Don't bet on it: contingent agreements with asymmetric information, *Journal of the American Statistical Association,* 78:382, 424-26.

Stalnaker, R. (1996), Knowledge, belief and counterfactual reasoning in games, forthcoming in *Economics and Philosophy*.