in: Itzhak Gilboa (ed.), *Theoretical aspects of rationality and knowledge* (TARK 1998), Morgan Kaufman, San Francisco, 1998, pp.147-160.

UNDERSTANDING COMMON PRIORS UNDER INCOMPLETE INFORMATION

Extended Abstract

Giacomo Bonanno Klaus Nehring

Department of Economics,
University of California,
Davis, CA 95616-8578, USA

e-mail: gfbonanno@ucdavis.edu,
kdnehring@ucdavis.edu

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). Recently several authors (Dekel and Gul, 1997, Gul, 1996, Lipman, 1995) have questioned whether the CPA is meaningful in 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. The main results of this paper (Theorems 1 and 2) 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 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 (where 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 we define *Comprehensive Agreement* as the absence of "agreement to disagree" about *any* aspect of beliefs in an appropriately defined general class.

In Theorem 1 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 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.

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 3 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.

Comprehensive Agreement defined as equality of commonly known aspects of beliefs is a concept that applies to *pairs* of individuals. In Section 4 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, whether Expectation Consistency is a legitimate primitive concept under incomplete information. In Section 5 we provide a reinterpretation of our results in the context of single-person, intertemporal belief revision.

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 CPA to Agreement. Feinberg, like Morris (1994), does not raise the issue of the conceptual content of the CPA under incomplete information. More recent related contributions are Halpern (1998) and Samet (1996b).

Proofs as well as further discussion and results can be found in Bonanno and Nehring (1996).

1. Interactive belief models

DEFINITION 1. An *interactive Bayesian model* is a tuple

$$\mathcal{B} = \langle N, \ \Omega, \ \tau, \ \Theta, \ \phi, \ \{p_i\}_{i\in N} \rangle, \quad \text{where}$$

- $N = \{1, ..., n\}$ is a finite set of *individuals*.
- Ω is a finite set of *states* (or possible worlds). The subsets of Ω are called *events.*
- $\tau \in \Omega$ is the "true" or "actual" state.
- 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 that are true at that state.
- for every individual i∈N, $p_i : \Omega \to \Delta(\Omega)$ (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 $p_{i,\alpha}$ rather than $p_i(\alpha)$]: $\forall\, \alpha, \beta \in \Omega$,

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

Thus $p_{i,\alpha} \in \Delta(\Omega)$ 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 $\|p_i = p_{i,\alpha}\|$ the event $\{\omega \in \Omega : p_{i,\omega} = p_{i,\alpha}\}$. It is clear that the set $\{\|p_i = p_{i,\omega}\| : \omega \in \Omega\}$ is a partition of Ω*;* it will be referred to as individual i's *type partition*.

For every individual i∈N, i's *possibility correspondence* $P_i : \Omega \to 2^{\Omega}$, is defined as follows [if $\mu \in \Delta(\Omega)$, $supp(\mu)$ denotes the support of μ, that is, the set of states to which μ assigns positive probability]:

$$P_i(\alpha) = supp(p_{i,\alpha}).$$

Thus, for every $\alpha \in \Omega$, $P_i(\alpha)$ 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: $\forall E \subseteq \Omega$, $B_iE = \{\omega \in \Omega : P_i(\omega) \subseteq E\}$. B_iE can be interpreted as the event that (i.e. the set of states at which) individual i *believes* (with certainty) that event E has occurred (i.e. attaches probability 1 to E). Notice that we have allowed for false beliefs by not assuming reflexivity of the possibility correspondences ($\forall \alpha \in \Omega$, $\alpha \in P_i(\alpha)$ or, equivalently, $p_{i,\alpha}(\alpha) > 0$), which – as is well known – is equivalent to the *Truth Axiom* : $\forall E \subseteq \Omega$, $B_iE \subseteq E$ (if the individual believes E then E is indeed true).

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: $B_*E = B_eE \cap B_eB_eE \cap B_eB_eB_eE \cap \ldots$ The corresponding possibility correspondence P_* is then defined as follows: for every $\alpha \in \Omega$, $P_*(\alpha) = \{\omega \in \Omega : \alpha \in \neg B_*\neg\{\omega\}\}$. It is well known that P_* is the *transitive closure* of $\bigcup_{i \in N} P_i$, that is,

> $\forall \alpha, \beta \in \Omega$, $\beta \in P_*(\alpha)$ 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 $k = 0, \ldots, m-1$, $\eta_{k+1} \in P_{i_{k+1}}(\eta_k)$.

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.[1] For example, consider the

---

[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

following model, which is illustrated in Figure 1: N = {1, 2}, $\Omega = \{\tau, \beta\}$, Θ = {spelling: Harsan*yi*, spelling: Harsan*iy*}, $\phi(\tau)$ = {spelling: Harsan*yi*}, $\phi(\beta)$ = {spelling: Harsan*iy*}, $P_1(\tau) = P_1(\beta) = \{\beta\}$, $P_2(\tau) = \{\tau\}$, $P_2(\beta) = \{\beta\}$. Thus $P_*(\tau) = \{\tau, \beta\}$ and $P_*(\beta) = \{\beta\}$. Here state τ represents the following beliefs. Individual 2 is a game-theorist who knows the correct spelling of his name (Harsan*yi*), while individual 1 mistakenly believes that the spelling is Harsan*iy*. Furthermore, individual 2 (mistakenly) believes that it is common belief between them that the correct spelling is Harsan*iy*.

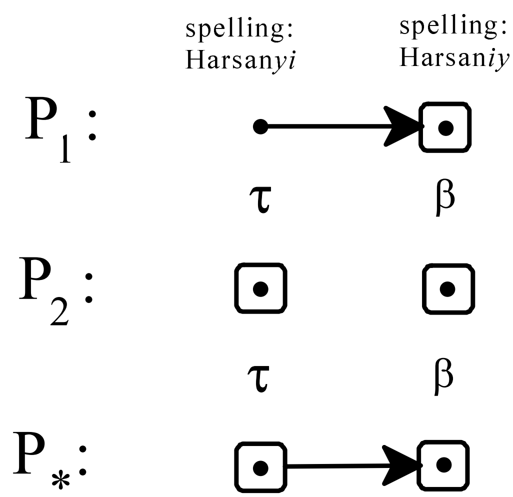

Figure 1

2. Harsanyi Quasi Consistency and Comprehensive Agreement

In this section we define a *local* version of Harsanyi consistency (i.e. the existence of a "common prior"). In an incomplete information context, properties of belief hierarchies ought 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.

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).

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: $\forall \alpha \in \Omega$, $\alpha \in \mathbf{HQC}_\mu$ if and only if

(1) $\forall i \in N$, $\forall \omega \in P_*(\alpha)$, if $\mu(\|p_i = p_{i,\omega}\|) > 0$ then $p_{i,\omega} = \mu(\cdot \mid \|p_i = p_{i,\omega}\|)$[2], and

(2) $\mu(P_*(\alpha)) > 0$

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 $\mathbf{HQC}_\mu = \mathbf{HQC} = \{\tau, \beta\}$.

Note that our definition of Harsanyi Quasi Consistency is entirely in terms of belief hierarchies. By contrast, related contributions in the literature (Feinberg, 1995, 1996, Halpern, 1998, Samet, 1996b) make use of a dual knowledge/belief framework and provide results about what is commonly *known* about individuals' beliefs, rather than about what is commonly believed. The definition given in the literature, if formulated locally, would be equivalent to $\mathbf{HQC}^K$ which is obtained by replacing the common belief correspondence P_* in Definition 2 with the common knowledge partition (where the knowledge partition of individual i is the type partition $\{\|p_i = p_{i,\omega}\| : \omega \in \Omega\}$). It can be shown (see Bonanno and Nehring, 1996) that $\mathbf{HQC}^K$ is substantially weaker than $\mathbf{HQC}$ and the gap between $\mathbf{HQC}^K$ and $\mathbf{HQC}$ cannot be bridged by local assumptions on the *belief* hierarchies in the manner of Theorem 2.[3]

The conceptual content of the notion of Harsanyi Quasi Consistency is not clear. The

---

[2] $\mu(\cdot \mid \|p_i = p_{i,\omega}\|) \in \Delta(\Omega)$ 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 $E \subseteq \Omega$, $\mu(E) = \sum_{\omega \in E} \mu(\omega)$.

[3] The bridging assumption for $\mathbf{HQC}^K$ analogous to the one provided in Theorem 2 below would be $\tau \in K_*\mathbf{T}$, that is, full support of individual beliefs (the event $\mathbf{T}$ is defined in Section 3 and K_* denotes the common knowledge operator).

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 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.* 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.

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 4).

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

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

Let $\mathscr{F}$ denote the class of proper belief indices.[4]

For example, if $E \subseteq \Omega$ is an arbitrary event, the function $f: \Delta(\Omega) \rightarrow [0,1]$ defined by $f^{E}(p) = p(E)$ is a proper belief index (thus $f^{E}(p_{i,\alpha})$ is individual i's subjective probability of event E at state α).

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[5]. Properness is necessary to ensure that public (i.e. commonly believed) 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.

Given a proper belief index $f: \Delta(\Omega) \rightarrow X$ and an individual $i \in N$, define $f_i : \Omega \rightarrow X$ 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 $\|f_i = x\|$.

---

[4] 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 $P_{*}(\tau)$.

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

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

$$\text{if} \quad \alpha \in \mathrm{B}_*\big(\|f_1 = x_1\| \cap \|f_2 = x_2\|\big) \quad \text{then} \quad x_1 = x_2 .$$

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

$$f\textbf{-Agree} = \bigcap_{\substack{x_1, x_2 \in X \\ x_1 \neq x_2}} \neg\mathrm{B}_*\big(\|f_1 = x_1\| \cap \|f_2 = x_2\|\big).$$

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

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

The following theorem characterizes Comprehensive Agreement as equivalent to Harsanyi Quasi Consistency. The key step in the proof of Theorem 1 is the observation that $\alpha \in$ **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 $\bigcap_{i \in N}$ co $\Pi_i(\alpha) \neq \emptyset$ where co $\Pi_i(\alpha)$ is the convex hull of $\Pi_i(\alpha) = \{p_{i,\omega} \in \Delta(\Omega) : \omega \in \mathrm{P}_*(\alpha)\}$.

THEOREM 1. **CA** = **HQC.**

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 $\mathcal{F}_2 \subseteq \mathcal{F}$ be the class of proper belief indices $f: \Delta(\Omega) \rightarrow X$ such that: (1) $X = \{0, 1\}$, and (2) $f^{-1}(1)$ is closed.

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

(ii) $f \in \mathscr{F}_2$ if and only if there exists a random variable $Y : \Omega \to \mathbb{R}$ such that, $\forall\, p \in \Delta(\Omega)$,

$$f(p) = \begin{cases} 1 & \text{if } \sum_{\omega \in \Omega} Y(\omega)\, p(\omega) \geq 0 \\ 0 & \text{otherwise} \end{cases}.$$

3. Strong Harsanyi Consistency requires Truth

The notion of Harsanyi Quasi Consistency is rather weak, in particular 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). Hence it is not surprising that 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, 1998, 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: $\forall\, \alpha \in \Omega$, $\alpha \in \mathbf{SHC}_\mu$ if and only if (1) $\alpha \in \mathbf{HQC}_\mu$, and (2) $\mu(\alpha) > 0$. Furthermore, let $\mathbf{SHC} = \bigcup_{\mu \in \Delta(\Omega)} \mathbf{SHC}_\mu$.

For example, in the model of Figure 1, while $\mathbf{HQC} = \{\tau, \beta\}$, $\mathbf{SHC} = \{\beta\}$. 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 $\mu(\cdot \mid P_*(\alpha)) = \nu(\cdot \mid P_*(\alpha))$. An analogous claim cannot be made for **HQC**.

The notion of Strong Harsanyi Consistency allows a 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, Aumann's (1987) 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, 1998, Theorem 3). Theorem 2 below replaces **SHC** with locally meaningful assumptions.

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 $\alpha \in E$)[6]. For example, in Figure 1 $\mathbf{T} = B_* \mathbf{T} = \{\beta\}$.

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. This condition (namely, $\mathbf{T} \cap B_* \mathbf{T}$), however, is stronger than needed, as shown in Bonanno and Nehring (1997).

4. Extension to more than two individuals

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 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 $Y_i : \Omega \to \mathbb{R}$ ($i \in N$) such that: (1) $\forall\, \omega \in \Omega$, $\sum_{i \in N} Y_i(\omega) = 0$, and (2) at α it is common belief

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

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.[7]

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, 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.

5. Intertemporal application: Bayesian updating without a prior

Our results have an interesting interpretation for 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 Harsanyi doctrine can be identified with the assumption of Bayesian updating 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

---

[7] 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. See also Feinberg (1995, 1996).

beliefs about her past and future beliefs.[8] 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 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. 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_***T** are non-trivial here in that they rule out the absence of *actual* surprises.

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.

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

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

[8] 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.

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

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.

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.

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

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

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

Halpern, J. (1998), Characterizing the Common Prior Assumption, mimeo, Cornell University.

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.

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

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.