Abstract

This paper studies strategic games in which the beliefs of each player are represented by a set of probability measures satisfying a parametric specialization that is called ϵ-contamination. That is, beliefs are represented by a set of probability measures, where every measure in the set has the form $(1-\epsilon)p^*+\epsilon p$, p^* being the benchmark probability measure, p being a contamination, and ϵ reflecting the amount of error in p^* that is deemed possible. Under a suitably modified common prior assumption, if beliefs about opponents' action choices are common knowledge, then beliefs satisfy some properties that can be interpreted as agreement and stochastic independence. *Journal of Economic Literature* Classification Numbers: C72, D81.

Epistemic Conditions for Agreement and Stochastic Independence of ϵ-Contaminated Beliefs

Kin Chung Lo *

Department of Economics, York University,

North York, Ontario, Canada M3J 1P3 †

This Version: August, 1998

*I especially thank a referee for providing detailed suggestions and comments that lead to substantial improvements. I also thank Larry Epstein for valuable comments and the Social Sciences and Humanities Research Council of Canada for financial support.

†E-mail address: kclo@yorku.ca

1. INTRODUCTION

According to the subjective expected utility model (henceforth the expected utility model), the beliefs of a decision maker are represented by a probability measure. However, the Ellsberg Paradox (Ellsberg, 1961) and related experimental findings (summarized by Camerer and Weber, 1992) demonstrate that in situations where a decision maker only possesses ambiguous information about the set of states of the world, the decision maker is typically averse to ambiguity and the probabilistic representation of beliefs is too restrictive. As a result, generalizations of the expected utility model have been proposed. In the Choquet expected utility model (Schmeidler, 1989), the decision maker's beliefs are represented by a capacity (that is, a non-additive probability measure). In the multiple priors model (Gilboa and Schmeidler, 1989), beliefs are represented by a set of probability measures. Also see Bewley (1986, 1987, 1988, 1989) for a related model and study of the relevance of ambiguity to various economic problems. Following Gilboa and Schmeidler (1993), we reserve the term "ambiguous beliefs" for beliefs that are representable by either a capacity or a set of probability measures.

Nash equilibrium has been the central solution concept in game theory. Adopting the view that choosing an action in a game is a decision problem under uncertainty, Nash equilibrium presumes that players' preferences are represented by the expected utility model (see Brandenburger, 1992, pp.90-92). In order to study the implications of ambiguity in strategic situations, this equilibrium concept has been generalized to allow for preferences to be representable by the Choquet expected utility and multiple priors models. See Dow and Werlang (1994), Eichberger and Kelsey (1994), Klibanoff (1993), Lo (1995, 1996), Marinacci (1995) and Mukerji (1995). The issue in this literature is how the following three features of Nash equilibrium should be generalized.

- *Best response property.* If an action of player i is in the support of player j's beliefs, then the action must be optimal for player i given i's beliefs.
- *Agreement of beliefs.* For each player, all the other players hold the same beliefs about the action choice of that player.
- *Stochastic independence of beliefs.* Each player believes that the action choices of the other players are stochastically independent.

The above papers focus attention on generalizing the best response property of Nash equilibrium. The issue of agreement and stochastic independence has been very much neglected. In particular, Dow and Werlang (1994), Lo (1995) and Marinacci (1995) avoid the issue by considering only two person games. Even though the other papers consider n-person games, they typically adopt definitions of agreement and stochastic independence of ambiguous beliefs in an arbitrary manner. This phenomenon can be easily explained. When beliefs are probabilistic, there is one standard and well received definition of agreement and stochastic independence. The definition is as follows. Players' beliefs about an event agree if they attach the same probability to that event. The beliefs of a player satisfy stochastic independence if beliefs are represented by a product measure. When beliefs are ambiguous, many different definitions of agreement and stochastic independence have been proposed. They possess some mathematical properties, but none of them can be clearly characterized in terms of preference.

However, even though the definition of stochastic independence of probabilistic beliefs enjoys respectable preference characterization in the single agent setting (see Blume *et al.*, 1991, p. 74, Axiom 6), whether it is legitimate to assume that beliefs are stochastically independent in multi agent strategic situations is a separate question (see Brandenburger, 1992, pp. 87-89). This concern is addressed by Aumann and Brandenburger (1995). They demonstrate that in a strategic game, if players' beliefs about opponents' action choices are derived from a common prior and if beliefs are common knowledge, then beliefs satisfy agreement and stochastic independence.

Since many definitions of agreement and stochastic independence of ambiguous beliefs have been suggested, the need to carry out the exercise of Aumann and Brandenburger for the case of ambiguous beliefs is even stronger. The only result in this direction appears in Lo (1996, p. 471, Proposition 7). However, that result is very restrictive because players' beliefs are assumed to satisfy "complete ignorance" (see the end of Section 5.3 for details). In this paper, we extend Lo's result to an important parametric class of ambiguous beliefs: ϵ-contamination. Roughly speaking, ϵ-contaminated beliefs are represented by a set of probability measures, where every measure in the set has the form $(1-\epsilon)p^*+\epsilon p$, p^* being the benchmark probability measure, p being a "contamination", and ϵ reflecting the amount of error in p^* that is deemed possible. For brevity, we will simply say in this paper that beliefs are *contaminated* if the parametric restriction of ϵ-contamination is satisfied. This class of beliefs is a special case of the class of belief functions (Dempster (1967) and Shafer (1976)), which is in turn a special case of convex capacities (Schmeidler, 1989).

There are reasons for studying contaminated beliefs. In the multiple priors model, the set of probability measures is unrestricted except by technical conditions. Therefore, the model may be too general for some settings. The class of contaminated beliefs is not only intuitively appealing, it also possesses sufficient structure so that tractable analysis is often possible. As a result, it has been thoroughly studied and its usefulness is well established in the robust statistics literature (see Berger (1984, 1985), Berger and Berliner (1986) and Wasserman (1990), for example). In the economics literature, many papers adopt the class of contaminated beliefs either as the primary model or at least as an illustrative example. See, for instance, Dow and Werlang (1992, p. 202; 1994, pp. 313-314), Eichberger and Kelsey (1994), Epstein (1997, pp. 15-16), Epstein and Wang (1994, pp. 288-289), Epstein and Zhang (1997) and Mukerji (1995). Finally, when beliefs are restricted to be contaminated, various rules for updating ambiguous beliefs proposed in the literature are equivalent (see Section 4.1 for details). As a result, the controversy on which updating rule is more appropriate than the others can be avoided.

More precisely, the aim of this paper is to search for a notion of agreement and stochastic independence of contaminated beliefs by asking the following question:

> Suppose that players' beliefs are contaminated. Under a suitably modified common prior assumption, does common knowledge of beliefs about opponents' action choices imply that beliefs satisfy any property that we may want to call "agreement and stochastic independence"?

We provide a positive answer to the above question. Recall that the key to the fundamental result of Aumann (1976) on agreement of beliefs and also to the result of Aumann and Brandenburger

(1995) is that conditional probabilistic beliefs that are updated from a prior using Bayes rule satisfy the *sure-thing principle*: Given any two disjoint events D and E, if the probability of an event is equal to α conditional on D, and similarly, if the probability of the event is also α conditional on E, then the probability of the event conditional on $D \cup E$ is also α. See Geanakoplos (1992, pp. 66-67). Roughly speaking, the key to our positive answer is that conditional contaminated beliefs (to be defined precisely in Section 4.1) also satisfy a sure-thing principle. Since the more general subclasses of ambiguous beliefs do not have such a property, we believe that the result in this paper cannot be extended further.

Although the focus of this paper is agreement and stochastic independence, the main result described in the preceding paragraph has an unexpected corollary which constitutes a separate contribution. The corollary is that all the definitions of support of ambiguous beliefs proposed in the literature coincide with each other. Therefore, under the above assumptions, the controversy on how the support of ambiguous beliefs is defined does not arise. Finally, we demonstrate that if players' rationality is mutual knowledge, then their beliefs constitute an equilibrium.

The plan of this paper is as follows. Section 2 establishes notation for the paper, provides a review of the multiple priors model and contaminated beliefs, and defines strategic games. Section 3 discusses the meaning of agreement and stochastic independence of beliefs in terms of preference and reviews existing definitions in terms of ambiguous beliefs. Section 4 presents the main result and Section 5 discusses its implications. An appendix contains proofs of all propositions.

2. PRELIMINARIES

2.1 Notation

The following notation applies throughout the paper.

- For any set Y, the cardinality of Y is denoted by $|Y|$. For any set Z, $Z \subseteq Y$ means that Z is a subset of Y, and $Z \subset Y$ means that Z is a strict subset of Y.
- For any finite set Y, the set of all probability measures on Y is denoted by $M(Y)$. For any $\psi \in M(Y)$, the support of ψ is denoted by supp ψ. For any $\Psi \subseteq M(Y)$, supp $\Psi \equiv \cup_{\psi \in \Psi}$supp ψ. That is, supp $\Psi \subseteq Y$ is the union of the supports of all the probability measures in Ψ.
- Suppose $Y = \times_{l=1}^{L} Y_l$. For any $\psi \in M(Y)$, the marginal probability measure of ψ on Y_l is denoted by $\text{marg}_{Y_l}\psi$. For any $\Psi \subseteq M(Y)$, $\text{marg}_{Y_l}\Psi \equiv \{\psi_l \in M(Y_l) : \exists \psi \in \Psi \text{ such that } \psi_l = \text{marg}_{Y_l}\psi\}$. That is, $\text{marg}_{Y_l}\Psi$ is the set of marginal probability measures on Y_l as one varies over the set of probability measures in Ψ.
- Suppose $Y = \times_{l=1}^{L} Y_l$. For any $Z \subseteq Y$, $\text{proj}_{Y_l}Z \equiv \{y_l \in Y_l : \exists z \in Z$ such that the lth co-ordinate of z is $y_l\}$. That is, $\text{proj}_{Y_l}Z$ is the projection of Z on Y_l.

2.2 Multiple Priors Model and Contaminated Beliefs

This section contains a review of the multiple priors model and contaminated beliefs.

Let S be a finite set of states of the world, X a set of outcomes, and $\mathcal{F}$ the set of acts from S to X. For notational simplicity, $x \in X$ also denotes the constant act that yields x for all $s \in S$. The decision maker has a preference ordering $\succeq$ over $\mathcal{F}$. According to the multiple priors model, the representation for $\succeq$ consists of a von Neuman Morgenstern (vNM) index $u : X \to \mathbf{R}$ and a closed and convex set $\mathcal{P}$ of probability measures on S. The utility function $U : \mathcal{F} \to \mathbf{R}$ representing $\succeq$ is given by

$$U(f) \equiv \min_{p \in \mathcal{P}} \sum_{s \in S} u(f(s))p(s). \tag{1}$$

The intuition of the model is as follows. The decision maker's beliefs over the state space S may be too vague to be represented by a probability measure and are represented instead by a set $\mathcal{P}$ of probability measures. The decision maker is averse to ambiguity in the sense that he evaluates an act by computing the minimum expected utility over the probability measures in $\mathcal{P}$.

In this paper, we focus on the following parametric specialization of $\mathcal{P}$: there exist an event $E \subseteq S$, a probability measure $p^* \in M(E)$ and a real number $\epsilon \in [0,1]$ such that

$$\mathcal{P} = \{(1-\epsilon)p^* + \epsilon p : p \in M(E)\}. \tag{2}$$

Say that $\mathcal{P}$ is *contaminated* if it satisfies (2). The probability measure p^* may be thought of as the benchmark probability measure governing the relative likelihood of events in S. However, the decision maker is not certain about p^* in the sense that it is contaminated or perturbed with weight ϵ by the probability measures in the set $M(E)$. Note that the event E is allowed, but is not required, to be the whole state space S. Finally, observe that with p^* and E fixed, the set $\mathcal{P}$ increases in the sense of set inclusion as ϵ increases, modeling increased ambiguity aversion. In particular, if $\epsilon = 0$, then $\mathcal{P} = \{p^*\}$; if $\epsilon = 1$, then $\mathcal{P} = M(E)$.

2.3 Strategic Games

In this section, we define strategic games in which each player behaves according to the single person decision theory in Section 2.2.

A *strategic game form* is denoted as $A \equiv \times_{i=1}^{n} A_i$, where A_i is a finite set of actions for player i. Throughout, the indices i, j and k vary over distinct players in $\{1, \ldots, n\}$. Elements in A_i and $\times_{j \neq i} A_j$ are denoted by a_i and a_{-i}, respectively. Given a strategic game form A, a *strategic game* is denoted as $u \equiv \{u_i\}_{i=1}^{n}$, where $u_i : A \to \mathbf{R}$ is a vNM index representing player i's preference ordering over A.

Since player i is uncertain about the action choices of the other players, the relevant state space for i is $\times_{j \neq i} A_j$. Consistent with the above single person decision theory, i's preference ordering over the set of acts defined on $\times_{j \neq i} A_j$ is represented by the multiple priors model defined in (1). Every action $a_i \in A_i$ can be identified as an act over $\times_{j \neq i} A_j$ and i's utility of taking a_i is equal to

$$\min_{\phi_i \in \Phi_i} \sum_{a_{-i} \in \times_{j \neq i} A_j} u_i(a_i, a_{-i}) \phi_i(a_{-i}),$$

where Φ_i is a closed and convex set of probability measures on $\times_{j \neq i} A_j$. Consistent with (2), if Φ_i is contaminated, then there exist $E_{-i} \subseteq \times_{j \neq i} A_j$, $\phi_i^* \in M(E_{-i})$ and $\epsilon_i \in [0,1]$ such that

$$\Phi_i = \{(1-\epsilon_i)\phi_i^* + \epsilon_i \phi_i : \phi_i \in M(E_{-i})\}.$$

It should be emphasized that in the absence of further assumptions, the event E_{-i} may not be a product space and ϕ_i^* may not be a product measure. In particular, supp ϕ_i^* may be any subset of E_{-i}.

3. AGREEMENT AND STOCHASTIC INDEPENDENCE

In Section 3.1, we present some preference based conditions that can be interpreted as agreement and stochastic independence. They serve as some criteria for evaluating the existing definitions of agreement and stochastic independence of ambiguous beliefs which we present in Section 3.2. We also suggest informally the definition of agreement and stochastic independence of contaminated beliefs which we justify in Section 4. For notational simplicity, assume that a game has three players.

3.1 Conditions in Terms of Preference

The preference based conditions for agreement and stochastic independence make use of the following two choice theoretic notions from Savage (1954). Recall the single agent setting in Section 2.2, where S denotes a set of states of the world, $\mathcal{F}$ the set of acts defined on S and $\succeq$ a preference ordering over $\mathcal{F}$. Given any event $E \subseteq S$, say that E is *null* if for all $f, f', g \in \mathcal{F}$,

$$\begin{bmatrix} f(\omega) & \text{if} & \omega \in E \\ g(\omega) & \text{if} & \omega \not\in E \end{bmatrix} \sim \begin{bmatrix} f'(\omega) & \text{if} & \omega \in E \\ g(\omega) & \text{if} & \omega \not\in E \end{bmatrix}. \tag{3}$$

In words, an event E is null if the decision maker does not care about payoffs for states that are in E.[1] If $\succeq$ is represented by the multiple priors model having $\mathcal{P}$ as the underlying set of probability measures, then

$$E \text{ is null if and only if } p(E) = 0 \quad \forall p \in \mathcal{P}. \tag{4}$$

The second choice theoretic notion concerns the decision maker's beliefs about the relative likelihood of events. Given any two events $D, E \subseteq S$, say that D is *more likely than* E if for all $x^*, x \in X$ with $x^* \succ x$,

$$\begin{bmatrix} x^* & \text{if} & s \in D \\ x & \text{if} & s \not\in D \end{bmatrix} \succeq \begin{bmatrix} x^* & \text{if} & s \in E \\ x & \text{if} & s \not\in E \end{bmatrix}. \tag{5}$$

In words, D is more likely than E if the decision maker prefers to bet on D rather than on E.[2] If

[1] Alternative definitions of null events have been proposed for the Choquet expected utility and multiple priors models. See Ryan (1998) for a summary. If an event is null in the sense of (3), then it must also be null according to all the alternative definitions. The converse is not true. Therefore, the preference based condition for stochastic independence defined in (10) below will be weaker if any of those definitions is adopted.

[2] The following way to induce a more likely than relation from preference seems equally legitimate: Given any $D, E \subseteq S$, D is *more likely than* E if for all $x^*, x \in X$ with $x^* \succ x$,

$$\begin{bmatrix} x & \text{if} & s \in E \\ x^* & \text{if} & s \not\in E \end{bmatrix} \succeq \begin{bmatrix} x & \text{if} & s \in D \\ x^* & \text{if} & s \not\in D \end{bmatrix}. \tag{6}$$

That is, the decision maker prefers to bet against E rather than against D. If $\succeq$ is represented by the multiple priors model, then D is more likely than E according to (6) if and only if $\max_{p \in \mathcal{P}} p(D) \geq \max_{p \in \mathcal{P}} p(E)$. In this paper, we adopt (5). However, all the results are valid if we adopt (6).

$\succeq$ is represented by the multiple priors model, then

$$D \text{ is more likely than } E \text{ if and only if } \min_{p \in \mathcal{P}} p(D) \geq \min_{p \in \mathcal{P}} p(E). \tag{7}$$

Going back to the context of strategic games, consider the beliefs of players i and j about the action choice of player k. The following is a basic condition for agreement of beliefs: Given any two events $D, E \subseteq A_k$,

$$\begin{aligned} &i \text{ believes that } D \text{ is more likely than } E \\ &\quad \text{if and only if } j \text{ believes that } D \text{ is more likely than } E. \end{aligned} \tag{8}$$

If i and j's preferences are represented by the multiple priors model, then (7) implies that (8) can be rewritten as follows: For all $D, E \subseteq A_k$,

$$\begin{aligned} &\min_{\sigma_k \in \mathrm{marg}_{A_k} \Phi_i} \sigma_k(D) \geq \min_{\sigma_k \in \mathrm{marg}_{A_k} \Phi_i} \sigma_k(E) \\ &\quad \text{if and only if } \min_{\sigma_k \in \mathrm{marg}_{A_k} \Phi_j} \sigma_k(D) \geq \min_{\sigma_k \in \mathrm{marg}_{A_k} \Phi_j} \sigma_k(E). \end{aligned} \tag{9}$$

Consider player i's beliefs about the action choices of players j and k. The following is a basic condition capturing the intuition that i believes that what j is going to do is independent of what k is going to do: Given any two events $D_j, E_j \subseteq A_j$ and any two events $D_k, E_k \subseteq A_k$ that are nonnull from i's point of view,[3]

$$\begin{aligned} &i \text{ believes that } D_j \times D_k \text{ is more likely than } E_j \times D_k \\ &\quad \text{if and only if } i \text{ believes that } D_j \times E_k \text{ is more likely than } E_j \times E_k. \end{aligned} \tag{10}$$

Again, if i's preference is represented by the multiple priors model, then we can use (4) and (7) to rewrite (10) as follows: For all $D_j, E_j \subseteq A_j$ and $D_k, E_k \subseteq A_k$ such that

$$\min_{\sigma_k \in \mathrm{marg}_{A_k} \Phi_i} \sigma_k(D_k) > 0 \text{ and } \min_{\sigma_k \in \mathrm{marg}_{A_k} \Phi_i} \sigma_k(E_k) > 0,$$

$$\begin{aligned} &\min_{\phi_i \in \Phi_i} \phi_i(D_j \times D_k) \geq \min_{\phi_i \in \Phi_i} \phi_i(E_j \times D_k) \\ &\quad \text{if and only if } \min_{\phi_i \in \Phi_i} \phi_i(D_j \times E_k) \geq \min_{\phi_i \in \Phi_i} \phi_i(E_j \times E_k). \end{aligned} \tag{11}$$

3.2 Existing Definitions in Terms of Ambiguous Beliefs

Eichberger and Kelsey (1994. p. 15, Definition 4.3), Lo (1996, p. 453, Definition 4) and Mukerji (1995, p. 20, Definition 7) adopt the following definition of agreement of ambiguous beliefs:

$$\mathrm{marg}_{A_k} \Phi_i = \mathrm{marg}_{A_k} \Phi_j. \tag{12}$$

[3]The notation $D_j \times D_k$ in (10) refers to the event that j chooses an action from D_j and k chooses an action from D_k. Similarly for $E_j \times D_k$, $D_j \times E_k$ and $E_j \times E_k$.

That is, the beliefs of players i and j about the action choice of player k are represented by the same set of probability measures. Clearly, (12) implies (9) and therefore the preference based condition for agreement in (8). The converse is not true. For instance, suppose $A_3 = \{L, R\}$,

$$\begin{aligned} \text{marg}_{A_3}\Phi_1 &= \{\sigma_3 \in M(A_3) : 0.4 \leq \sigma_3(L) \leq 0.6\} \text{ and} \\ \text{marg}_{A_3}\Phi_2 &= \{\sigma_3 \in M(A_3) : 0.3 \leq \sigma_3(L) \leq 0.7\}. \end{aligned} \tag{13}$$

Clearly, $\text{marg}_{A_3}\Phi_1 \neq \text{marg}_{A_3}\Phi_2$ and therefore (12) is violated. However, both players believe that L and R are equally likely.

In fact, the sets of probability measures in (13) can be rewritten as

$$\begin{aligned} \text{marg}_{A_3}\Phi_1 &= \{(1-\epsilon_1)\sigma_3^* + \epsilon_1\sigma_3 : \sigma_3 \in M(A_3)\} \text{ and} \\ \text{marg}_{A_3}\Phi_2 &= \{(1-\epsilon_2)\sigma_3^* + \epsilon_2\sigma_3 : \sigma_3 \in M(A_3)\}, \end{aligned}$$

where $\sigma_3^* = (L, 0.5; R, 0.5)$, $\epsilon_1 = \frac{1}{5}$ and $\epsilon_2 = \frac{2}{5}$. That is, $\text{marg}_{A_3}\Phi_1$ and $\text{marg}_{A_3}\Phi_2$ share the same benchmark probability measure and contamination, and their difference is due to the fact that $\epsilon_1 \neq \epsilon_2$. In Section 4, we will explain why this phenomenon may occur.

Klibanoff (1993) adopts the following definition of agreement:

$$\text{marg}_{A_k}\Phi_i \cap \text{marg}_{A_k}\Phi_j \neq \emptyset. \tag{14}$$

That is, the set of probability measures representing player i's beliefs and that representing player j's beliefs have a nonempty intersection. In contrast to the definition of agreement in (12), the one in (14) does not imply the preference based condition in (8). For instance, suppose $A_3 = \{L, R\}$,

$$\begin{aligned} \text{marg}_{A_3}\Phi_1 &= \{\sigma_3 \in M(A_3) : 0 \leq \sigma_3(L) \leq 0.6\} \text{ and} \\ \text{marg}_{A_3}\Phi_2 &= \{\sigma_3 \in M(A_3) : 0.6 \leq \sigma_3(L) \leq 1\}. \end{aligned}$$

The probability measure $(L, 0.6; R, 0.4)$ is contained in both $\text{marg}_{A_3}\Phi_1$ and $\text{marg}_{A_3}\Phi_2$. Therefore, (14) is satisfied. However, player 1 believes that R is strictly more likely than L but player 2 believes the opposite. Therefore, (8) is violated.

We now turn to existing definitions of stochastic independence of ambiguous beliefs. Almost all of them imply that Φ_i satisfies the following *independent product property*: for all events $E_j \subseteq A_j$ and $E_k \subseteq A_k$,

$$\min_{\phi_i \in \Phi_i} \phi_i(E_j \times E_k) = \min_{\sigma_j \in \text{marg}_{A_j}\Phi_i} \sigma_j(E_j) \min_{\sigma_k \in \text{marg}_{A_k}\Phi_i} \sigma_k(E_k). \tag{15}$$

It is well known that if Φ_i is a singleton, then Φ_i is completely determined by the right hand side of (15). Therefore, both Ghirardato (1997, p. 270) and Hendon *et al.* (1996, p. 97) regard the independent product property as basic that any definition of stochastic independence of ambiguous beliefs should satisfy. The definition proposed by Gilboa and Schmeidler (1989, pp. 150-151), which is adapted to the context of strategic games by Lo (1996, p. 453, Definition 4), also satisfies independent product property. Similarly for the definitions in Eichberger and Kelsey (1994, p. 14, Definition 4.2) and Mukerji (1995, p. 8, Definition 2).

Clearly, the independent product property implies (11) and therefore the preference based condition for stochastic independence in (10). However, except for a few trivial cases, it is incompatible with contaminated beliefs.[4]

Proposition 1. *Suppose there exist $E_{-i} \subseteq A_j \times A_k$ with $|\text{proj}_{A_j} E_{-i}| > 1$ and $|\text{proj}_{A_k} E_{-i}| > 1$, $\phi_i^* \in M(E_{-i})$, and $\epsilon_i \in (0,1)$ such that $\Phi_i = \{(1-\epsilon_i)\phi_i^* + \epsilon_i \phi_i : \phi_i \in M(E_{-i})\}$. Then Φ_i does not satisfy the independent product property.*

The following example, which is presented in a setting that will be formally defined in Section 4.1, illustrates the intuition behind Proposition 1.

Example 1. Suppose player 1 is facing a set $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ of states of the world, where every state $\omega \in \Omega$ contains the following specification of actions $\mathbf{a}_2(\omega) \in A_2$ and $\mathbf{a}_3(\omega) \in A_3$ taken by players 2 and 3, respectively.

	ω_1	ω_2	ω_3	ω_4
$\mathbf{a}_2$	U	D	U	D
$\mathbf{a}_3$	L	R	R	L

Suppose player 1's beliefs about Ω are represented by the set of probability measures

$$\Lambda_1 = \{(1-\epsilon_1)\mu_1^* + \epsilon_1 \mu_1 : \mu_1 \in M(\Omega)\}, \tag{16}$$

where $\mu_1^* = (\omega_1, 0.25; \omega_2, 0.25; \omega_3, 0.25; \omega_4, 0.25)$ and $\epsilon_1 \in (0,1)$. Given the action functions $\mathbf{a}_2$ and $\mathbf{a}_3$, every probability measure in Λ_1 induces a probability measure on $\{U,D\} \times \{L,R\}$ in the obvious manner. If we do it for every probability measure in Λ_1, we obtain the set of probability measures

$$\Phi_1 = \left\{ (1-\epsilon_1) \prod_{j=2,3} \sigma_j^* + \epsilon_1 \phi_1 : \phi_1 \in M(\{U,D\} \times \{L,R\}) \right\}, \tag{17}$$

where $\sigma_2^* = (U, 0.5; D, 0.5)$ and $\sigma_3^* = (L, 0.5; R, 0.5)$.

Given Φ_1 in (17), 1's marginal beliefs about 2 and 3's action choices are represented by

$$\begin{aligned} \text{marg}_{A_2} \Phi_1 &= \{(1-\epsilon_1)\sigma_2^* + \epsilon_1 \sigma_2 : \sigma_2 \in M(\{U,D\})\} \text{ and} \\ \text{marg}_{A_3} \Phi_2 &= \{(1-\epsilon_1)\sigma_3^* + \epsilon_1 \sigma_3 : \sigma_3 \in M(\{L,R\})\}, \end{aligned} \tag{18}$$

respectively. Note that Φ_1, $\text{marg}_{A_2}\Phi_1$ and $\text{marg}_{A_3}\Phi_1$ are all contaminated with the *same* degree of contamination ϵ_1. The reason for this is as follows. The number ϵ_1 represents player 1's degree of uncertainty about Ω. Note that the action choices of players 2 and 3 are both functions on Ω. Therefore, 1's degree of uncertainty about the action choices of 2 and 3, or each of their action choices separately, should also be represented by ϵ_1. In contrast, given 1's marginal beliefs in (18), the independent product property requires player 1, when considering the overall action choices of

[4] It follows that although Eichberger and Kelsey (1994) and Mukerji (1995) require all the marginal beliefs $\text{marg}_{A_j}\Phi_i$ to be contaminated, the overall beliefs Φ_i are *not* contaminated.

his two opponents, to "double count" ϵ_1. For example, take the actions U and L. Let Ψ_1 be any set of probability measures on $A_2 \times A_3$ such that $\text{marg}_{A_2}\Psi_1 = \text{marg}_{A_2}\Phi_1$, $\text{marg}_{A_3}\Psi_1 = \text{marg}_{A_3}\Phi_1$, and the independent product property is satisfied. Then we have

$$\begin{aligned}\min_{\psi_1 \in \Psi_1} \psi_1(U \times L) &= \min_{\sigma_2 \in \text{marg}_{A_2}\Phi_1} \sigma_2(U) \min_{\sigma_3 \in \text{marg}_{A_3}\Phi_1} \sigma_3(L) \\ &= (1-\epsilon_1)^2(0.25),\end{aligned}$$

which is strictly less than

$$\min_{\phi_1 \in \Phi_1} \phi_1(U \times L) = (1-\epsilon_1)(0.25).$$

That is, the independent product property ignores the fact that player 1's uncertainty about each of his opponents is in fact originated from the same source Ω.

Although Φ_1 in (17) does not satisfy the independent product property, it has the following features of "stochastic independence": The benchmark probability measure $\prod_{j=2,3}\sigma_j^*$ is a product measure and the contamination is defined on the product space $\{U,D\} \times \{L,R\}$. In Section 4, we will describe precisely the conditions under which these features arise. $\diamondsuit$

The only existing definition of stochastic independence of ambiguous beliefs which do not imply the independent product property is due to Klibanoff (1993). He only requires

$$\Phi_i \text{ to contain at least one product measure.} \tag{19}$$

This definition is compatible with contaminated beliefs. For instance, suppose $A_2 = \{U,D\}$, $A_3 = \{L,R\}$ and

$$\Phi_1 = \{0.5\phi_1^* + 0.5\phi_1 : \phi_1 \in M(A_2 \times A_3)\},$$

where $\phi_1^* = (UL, 0.5; UR, 0; DL, 0; DR, 0.5)$. Then Φ_1 satisfies (19) because it contains the product measure $(UL, 0.25; UR, 0.25; DL, 0.25; DR, 0.25)$. However, we have

$$\min_{\phi_1 \in \Phi_1} \phi_1(UL) = \min_{\phi_1 \in \Phi_1} \phi_1(DR) = 0.25 \quad \text{and} \quad \min_{\phi_1 \in \Phi_1} \phi_1(UR) = \min_{\phi_1 \in \Phi_1} \phi_1(DL) = 0.$$

That is, player 1 believes that UL is more likely than DL, but UR is less likely than DR. This demonstrates that, even if Φ_i is contaminated, (19) does not imply the preference based condition for stochastic independence in (10).

4. EPISTEMIC CONDITIONS

4.1 Generalization of the Common Prior Assumption

In this section, we first establish a decision theoretic framework for discussing epistemic matters. Then we formulate a generalization of the common prior assumption for contaminated beliefs.

Fix a strategic game form A. The model that we use to discuss epistemic matters is denoted as $\{\Omega, H_i, \Delta_i, \mathbf{u}_i, \mathbf{a}_i\}$. Each of its components is explained as follows.

- All the players are facing a finite set Ω of states of the world.
- Player i's information structure is represented by a partition H_i of Ω. For every state $\omega \in \Omega$, $H_i(\omega) \in H_i$ denotes the partitional element which contains ω.
- Player i's beliefs at ω are represented by a closed and convex set $\Delta_i(\omega)$ of probability measures on $H_i(\omega)$.
- Player i's payoff function at ω is $\mathbf{u}_i(\omega) : A \to \mathbf{R}$.
- The action taken by player i at ω is $\mathbf{a}_i(\omega) \in A_i$.

Note that Δ_i, $\mathbf{u}_i$ and $\mathbf{a}_i$ are functions on Ω. To respect the partitional information structure, they are measurable with respect to H_i. Also note that at every $\omega \in \Omega$, $\mathbf{u}(\omega) \equiv \{\mathbf{u}_1(\omega), \ldots, \mathbf{u}_n(\omega)\}$ constitutes a strategic game.

The above model is essentially the same as the standard model that is used to discuss epistemic conditions for Bayesian solution concepts. See, for instance, Dekel and Gul (1997, p. 17) and Osborne and Rubinstein (1994, p. 76). The only difference is that in the Bayesian case, the set $\Delta_i(\omega)$ of probability measures at every state $\omega \in \Omega$ and for every player i is required to be a singleton.

The object that is of ultimate interest is player i's beliefs about opponents' action choices. At every state $\omega \in \Omega$, i's beliefs over $\times_{j \neq i} A_j$ are represented by the closed and convex set $\mathbf{\Phi}_i(\omega)$ of probability measures which is induced from $\Delta_i(\omega)$ as follows:

$$\begin{aligned} \mathbf{\Phi}_i(\omega) = \{ & \phi_i \in M(\times_{j \neq i} A_j) : \exists \mu_i \in \Delta_i(\omega) \\ & \text{such that for all } a_{-i} \in \times_{j \neq i} A_j, \phi_i(a_{-i}) = \mu_i(\{\omega' \in H_i(\omega) : \mathbf{a}_{-i}(\omega') = a_{-i}\})\}, \end{aligned} \tag{20}$$

where $\mathbf{a}_{-i}(\omega') \equiv \{\mathbf{a}_1(\omega'), \ldots, \mathbf{a}_{i-1}(\omega'), \mathbf{a}_{i+1}(\omega'), \ldots, \mathbf{a}_n(\omega')\}$. Denote the n-tuple $\{\mathbf{\Phi}_1(\omega), \ldots, \mathbf{\Phi}_n(\omega)\}$ by $\mathbf{\Phi}(\omega)$.

Say that player i is *rational at* ω if

$$\mathbf{a}_i(\omega) \in \arg\max_{a_i \in A_i} \min_{\phi_i \in \mathbf{\Phi}_i(\omega)} \sum_{a_{-i} \in \times_{j \neq i} A_j} \mathbf{u}_i(\omega)(a_i, a_{-i}) \phi_i(a_{-i}).$$

That is, i is rational at ω if his action $\mathbf{a}_i(\omega)$ maximizes utility when the payoff function is $\mathbf{u}_i(\omega)$ and when beliefs are represented by $\mathbf{\Phi}_i(\omega)$.

Let μ be a probability measure on Ω with the property that $\mu(H_i(\omega)) > 0$ for all $\omega \in \Omega$ and for all i. Let $\mu(\cdot \mid H_i(\omega))$ be the probability measure on $H_i(\omega)$ which is derived from μ using Bayes rule. The common prior assumption can be stated as follows: The n-tuple $\{\Delta_i\}_{i=1}^n$ satisfies the *common prior assumption* if there exists $\mu \in M(\Omega)$ such that

$$\Delta_i(\omega) = \mu(\cdot \mid H_i(\omega)) \quad \forall \omega \in \Omega \quad \forall i.$$

Consider the following generalization of the common prior assumption.

Definition 1. The n-tuple $\{\Delta_i\}_{i=1}^n$ satisfies the *assumption of common prior with contamination* if there exist $\mu \in M(\Omega)$ and a collection $\{\boldsymbol{\epsilon}_i\}_{i=1}^n$ of functions, where $\boldsymbol{\epsilon}_i : \Omega \to [0,1]$ is measurable with respect to H_i, such that

$$\Delta_i(\omega) = \{(1 - \boldsymbol{\epsilon}_i(\omega))\mu(\cdot \mid H_i(\omega)) + \boldsymbol{\epsilon}_i(\omega)\mu_i : \mu_i \in M(H_i(\omega))\} \quad \forall \omega \in \Omega \quad \forall i. \tag{21}$$

Definition 1 is implied, and therefore can be justified, by the following assumption. Assume that ex ante, players' beliefs about the set Ω of states of the world are contaminated with the same benchmark probability measure $\mu \in M(\Omega)$. That is, i's ex ante beliefs are represented by

$$\Lambda_i \equiv \{(1 - \epsilon_i)\mu + \epsilon_i p : p \in M(\Omega)\}, \tag{22}$$

where $\epsilon_i \in [0,1]$. When a state ω is realized, player i is informed that the event $H_i(\omega)$ has occurred. The set $\Delta_i(\omega)$ of probability measures represents i's updated beliefs. As suggested by Gilboa and Schmeidler (1993), a natural procedure to derive $\Delta_i(\omega)$ is to rule out some of the priors in Λ_i and then update the rest according to Bayes rule. Two updating rules of particular interest are the *maximum likelihood updating rule*:

$$\begin{aligned}\Delta_i(\omega) = \{\mu_i \in M(H_i(\omega)) : \exists p \in \Lambda_i \text{ such that} \\ p(H_i(\omega)) = \max_{\hat{p} \in \Lambda_i} \hat{p}(H_i(\omega)) \text{ and } \mu_i(\cdot) = p(\cdot|H_i(\omega))\}\end{aligned} \tag{23}$$

and the *full Bayesian updating rule*:

$$\begin{aligned}\Delta_i(\omega) = \{\mu_i \in M(H_i(\omega)) : \exists p \in \Lambda_i \text{ such that} \\ p(H_i(\omega)) > 0 \text{ and } \mu_i(\cdot) = p(\cdot|H_i(\omega))\}.\end{aligned} \tag{24}$$

That is, according to the maximum likelihood updating rule, only those probability measures in Λ_i that ascribe the maximal probability to $H_i(\omega)$ are updated. On the other hand, according to the full Bayesian updating rule, every probability measure in Λ_i which ascribes positive probability to $H_i(\omega)$ is updated. Clearly, the set of probability measures in (23) must be a subset of that in (24). Proposition 2 below says that if Λ_i satisfies (22), then the converse is also true. Moreover, Proposition 2 constitutes a justification for (21). Since Λ_i in (22) can be rewritten as a convex capacity (Schmeidler, 1989), the maximum likelihood updating rule is also equivalent to the *Dempster-Shafer updating rule* for belief functions. See Gilboa and Schmeidler (1993, p. 42, Theorem 3.3).

Proposition 2. *Given the set Λ_i of probability measures defined in (22), the maximum likelihood and full Bayesian updating rules are equivalent. Moreover, they both imply that $\Delta_i(\omega)$ satisfies (21) with*

$$\boldsymbol{\epsilon}_i(\omega) = \frac{\epsilon_i}{\epsilon_i + (1 - \epsilon_i)\mu(H_i(\omega))} \quad \forall \omega \in \Omega. \tag{25}$$

It is important to note that $\boldsymbol{\epsilon}_i(\omega)$ in (25) depends on $H_i(\omega)$. Therefore, even if we assume that all the players have the same degree of contamination ex ante, they typically do not share the same degree of contamination after receiving their private information.

It is also worth pointing out that Definition 1 admits the following special case, where no ex ante stage has to be specified and the existence of a common prior can simply be viewed as a coincidence. Suppose that when a particular state ω is realized, player i is informed that the event $H_i(\omega)$ has occurred. However, i does not have any extra information about the relative likelihood of events in $H_i(\omega)$. If i's beliefs are probabilistic, it is natural for i's beliefs to be represented by the uniform probability measure on $H_i(\omega)$. On the other hand, if i feels completely ignorant, i's beliefs are represented by the set of all probability measures on $H_i(\omega)$. If we set $\mu(\cdot \mid H_i(\omega))$ in (21) to be the uniform probability measure on $H_i(\omega)$, then the set $\Delta_i(\omega)$ of probability measures in (21) represents a convex combination or "compromise" between these two cases. If this holds for every $\omega \in \Omega$ and every i, then Definition 1 is satisfied. The common prior μ is the uniform probability measure on Ω.

Definition 1 has the following variation.

Definition 2. The n-tuple $\{\Delta_i\}_{i=1}^n$ satisfies the *assumption of common prior with absolutely continuous contamination* if there exist $\mu \in M(\Omega)$ and a collection $\{\boldsymbol{\epsilon}_i\}_{i=1}^n$ of functions, where $\boldsymbol{\epsilon}_i : \Omega \to [0,1]$ is measurable with respect to H_i, such that

$$\Delta_i(\omega) = \{(1-\boldsymbol{\epsilon}_i(\omega))\mu(\cdot \mid H_i(\omega)) + \boldsymbol{\epsilon}_i(\omega)\mu_i : \mu_i \in M(H_i(\omega) \cap \text{supp } \mu)\} \quad \forall \omega \in \Omega \quad \forall i.$$

Definition 2 gets its name because it only allows contamination that is absolutely continuous with respect to the benchmark probability measure $\mu(\cdot \mid H_i(\omega))$. The interpretation of this definition is that player i has no ambiguity about events that are null with respect to $\mu(\cdot \mid H_i(\omega))$. By redefining Λ_i in (22) to be $\{(1-\epsilon_i)\mu + \epsilon_i p : p \in M(\text{supp } \mu)\}$, the justification for Definition 1 justifies Definition 2 in exactly the same manner.

4.2 The Main Result

For any function $\mathbf{x}$ defined on Ω and for any value x, $[x]$ denotes the event $\{\omega \in \Omega : \mathbf{x}(\omega) = x\}$. For example, $[\Phi] \equiv \{\omega \in \Omega : \mathbf{\Phi}(\omega) = \Phi\}$ and $[u] \equiv \{\omega \in \Omega : \mathbf{u}(\omega) = u\}$. Adopt the standard definition of knowledge.[5] That is, given any event E, *i knows E at ω* if $H_i(\omega) \subseteq E$. Note that if i knows E at ω, then E is true at ω in the sense that $\omega \in E$. Say that E is *mutual knowledge at ω* if everyone knows E at ω. Let $\mathcal{H}$ be the meet of the partitions of all the players and $\mathcal{H}(\omega)$ be the element of $\mathcal{H}$ which contains ω. Say that E is *common knowledge at ω* if $\mathcal{H}(\omega) \subseteq E$ (Aumann, 1976).

We are now ready to state that the assumption of common prior with (absolutely continuous) contamination, together with common knowledge of beliefs about opponents' action choices, lead to a generalization of the result of Aumann and Brandenburger (1995).

Proposition 3. *Suppose $\{\Delta_i\}_{i=1}^n$ satisfies the assumption of common prior with (absolutely continuous) contamination. Suppose that at a state ω^*, either $\boldsymbol{\epsilon}_i(\omega^*) \in [0,1)$ for all i or $\boldsymbol{\epsilon}_i(\omega^*) = 1$ for all i, and $[\Phi]$ is common knowledge. Then there exists a collection $\{\sigma_i^*\}_{i=1}^n$ of probability measures,*

[5] The results in this paper are valid if knowledge is replaced by belief. See the end of this section for details.

where $\sigma_i^* \in M(A_i)$, *such that*

$$\Phi_i = \left\{ (1 - \boldsymbol{\epsilon}_i(\omega^*)) \prod_{j \neq i} \sigma_j^* + \boldsymbol{\epsilon}_i(\omega^*) \phi_i : \phi_i \in M(\times_{j \neq i} \text{supp } \sigma_j^*) \right\} \quad \forall i. \tag{26}$$

The key to the proof of Proposition 3 is as follows. The assumption of common prior with contamination implies that $\mathbf{\Phi}_i(\omega)$ derived in (20) can be rewritten as

$$\mathbf{\Phi}_i(\omega) = \{(1 - \boldsymbol{\epsilon}_i(\omega))\phi^i(\cdot \mid H_i(\omega)) + \boldsymbol{\epsilon}_i(\omega)\phi_i : \phi_i \in M(\{\mathbf{a}_{-i}(\omega') : \omega' \in H_i(\omega)\})\} \quad \forall \omega \in \Omega, \tag{27}$$

where $\phi^i(\cdot \mid H_i(\omega)) \in M(\times_{j \neq i} A_j)$ is induced from the posterior $\mu(\cdot \mid H_i(\omega))$ in the same manner as $\mathbf{\Phi}_i(\omega)$ is induced from $\Delta_i(\omega)$. Let $D, E \in H_i$ be two different partitional elements for player i. Consider the case where $\boldsymbol{\epsilon}_i(\omega) \in (0,1)$ for all $\omega \in D \cup E$. We have the following sure-thing principle for contaminated beliefs: If

$$\mathbf{\Phi}_i(\omega_D) = \mathbf{\Phi}_i(\omega_E), \text{ where } \omega_D \in D \text{ and } \omega_E \in E,$$

Then

$$\phi^i(\cdot|D) = \phi^i(\cdot|E) = \phi^i(\cdot|D \cup E) \tag{28}$$

and

$$\{\mathbf{a}_{-i}(\omega) : \omega \in D\} = \{\mathbf{a}_{-i}(\omega) : \omega \in E\} = \{\mathbf{a}_{-i}(\omega) : \omega \in D \cup E\}. \tag{29}$$

Therefore, if we require player i's beliefs about opponents' action choices conditional on $D \cup E$ to satisfy (27) as well, then (28) and (29) imply that i's beliefs conditional on $D \cup E$ will still have the same benchmark probability measure and contamination. The proof of Proposition 3 is essentially established by applying the sure-thing principle to each player's partitional elements in $\mathcal{H}(\omega^*)$.

We also use Examples 2 and 3 below to illustrate Proposition 3. They both make use of the following

- strategic game form $\times_{i=1}^3 A_i$, where $A_1 = \{N, S\}$, $A_2 = \{U, D\}$ and $A_3 = \{L, C, R\}$
- set of states of the world $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$
- information partitions

$$H_1 = \{\{\omega_1, \omega_2, \omega_3, \omega_4\}, \{\omega_5\}, \{\omega_6\}\},$$
$$H_2 = \{\{\omega_1, \omega_3\}, \{\omega_2, \omega_4\}, \{\omega_5\}, \{\omega_6\}\} \text{ and } H_3 = \{\{\omega_1, \omega_4\}, \{\omega_2, \omega_3\}\{\omega_5, \omega_6\}\}$$

- action functions

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
$\mathbf{a}_1$	N	N	N	N	N	S
$\mathbf{a}_2$	U	D	U	D	U	D
$\mathbf{a}_3$	L	R	R	L	C	C

The common prior μ will be different in the two examples and therefore will be specified separately.

Example 2. Suppose $\{\Delta_i\}_{i=1}^3$ satisfies the assumption of common prior with contamination, where the common prior μ is the uniform probability measure on Ω. Suppose that $\boldsymbol{\epsilon}_i(\omega) = \epsilon_i \in [0,1]$ for all $\omega \in \{\omega_1, \omega_2, \omega_3, \omega_4\}$. According to (27),

$$\boldsymbol{\Phi}_i(\omega) = \{(1-\epsilon_i)\prod_{j\neq i}\sigma_j^* + \epsilon_i\phi_i : \phi_i \in M(\times_{j\neq i}\text{supp } \sigma_j^*)\} \quad \forall\omega \in \{\omega_1, \omega_2, \omega_3, \omega_4\}, \tag{30}$$

where $\sigma_1^* = (N, 1)$, $\sigma_2^* = (U, 0.5; D, 0.5)$ and $\sigma_3^* = (L, 0.5; R, 0.5)$. At every $\omega \in \{\omega_1, \omega_2, \omega_3, \omega_4\}$, $[\boldsymbol{\Phi}(\omega)]$ is common knowledge. Note that $\boldsymbol{\Phi}_i(\omega)$ in (30) takes the form of (26), confirming Proposition 1. On the other hand, consider state ω_5. According to (27),

$$\boldsymbol{\Phi}_3(\omega_5) = \{(1-\boldsymbol{\epsilon}_3(\omega_5))\phi_3^* + \boldsymbol{\epsilon}_3(\omega_5)\phi_3 : \phi_3 \in M(\{NU, SD\})\},$$

where $\phi_3^* = (NU, 0.5; SD, 0.5)$. Since $\{NU, SD\}$ is not a product space, $\boldsymbol{\Phi}_3(\omega_5)$ does not satisfy (26). This is attributed to the fact that $[\boldsymbol{\Phi}_1(\omega_5)]$ and $[\boldsymbol{\Phi}_2(\omega_5)]$ are not common knowledge at ω_5, even though $[\boldsymbol{\Phi}_3(\omega_5)]$ is. $\diamondsuit$

Example 3. In Example 2, $\boldsymbol{\Phi}_i(\omega)$ in (30) takes the form of (26), even if the restriction "either $\boldsymbol{\epsilon}_i(\omega^*) \in [0,1)$ for all i or $\boldsymbol{\epsilon}_i(\omega^*) = 1$ for all i" is not imposed. We demonstrate here that in order to guarantee (26), this restriction is needed. Suppose $\{\Delta_i\}_{i=1}^3$ satisfies the assumption of common prior with contamination, where $\mu = (\omega_1, \frac{1}{10}; \omega_2, \frac{2}{10}; \omega_3, \frac{3}{10}; \omega_4, \frac{2}{10}; \omega_5, \frac{1}{10}; \omega_6, \frac{1}{10})$. Suppose $\boldsymbol{\epsilon}_1(\omega) = \epsilon_1 \in [0,1)$ and $\boldsymbol{\epsilon}_2(\omega) = \boldsymbol{\epsilon}_3(\omega) = 1$ for all $\omega \in \{\omega_1, \omega_2, \omega_3, \omega_4\}$. Although $[\boldsymbol{\Phi}(\omega)]$ is still common knowledge at every $\omega \in \{\omega_1, \omega_2, \omega_3, \omega_4\}$, the benchmark probability measure in $\boldsymbol{\Phi}_1(\omega)$ becomes $(UL, \frac{1}{8}; DR, \frac{2}{8}; UR, \frac{3}{8}; DL, \frac{2}{8})$, which is not a product measure. $\diamondsuit$

It is natural to ask whether Proposition 3 can be further extended to more general subclasses of ambiguous beliefs. Example 4 below suggests that the answer is negative.

Example 4. This example makes use of the following

- strategic game form $\times_{i=1}^3 A_i$, where $A_1 = \{N\}$, $A_2 = \{U\}$ and $A_3 = \{L, R\}$
- set of states of the world $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$
- information partitions

$$H_1 = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\},$$
$$H_2 = \{\{\omega_1, \omega_2, \omega_3, \omega_4\}\} \text{ and } H_3 = \{\{\omega_1, \omega_4\}, \{\omega_2, \omega_3\}\}$$

- action functions

	ω_1	ω_2	ω_3	ω_4
$\mathbf{a}_1$	N	N	N	N
$\mathbf{a}_2$	U	U	U	U
$\mathbf{a}_3$	R	L	L	R

Assume that ex ante, the beliefs of all the players are represented by the set of probability measures

$$\Lambda = \left\{ p \in M(\Omega) : p(\{\omega_1, \omega_4\}) = p(\{\omega_2\}) = p(\{\omega_3\}) = \frac{1}{3} \right\}.$$

Note that Λ can be rewritten as a belief function and therefore also a convex capacity. Assume that player i's beliefs $\Delta_i(\omega)$ at every state $\omega \in \Omega$ are updated from Λ using the maximum likelihood updating rule.[6] Again, assume that i's beliefs $\mathbf{\Phi}_i(\omega)$ about opponents' action choices are derived from $\Delta_i(\omega)$ as in (20). In this example, for any two states ω and ω', $\mathbf{\Phi}(\omega) = \mathbf{\Phi}(\omega')$. Therefore, $[\mathbf{\Phi}(\omega)]$ is common knowledge at every ω. However,

$$\begin{aligned} \text{marg}_{A_3} \mathbf{\Phi}_1(\omega) &= \left\{ \sigma_3 \in M(A_3) : \sigma_3(L) = \sigma_3(R) = \frac{1}{2} \right\} \text{ and} \\ \text{marg}_{A_3} \mathbf{\Phi}_2(\omega) &= \left\{ \sigma_3 \in M(A_3) : \sigma_3(L) = \frac{2}{3}, \sigma_3(R) = \frac{1}{3} \right\}. \end{aligned} \quad (31)$$

According to (31), player 1 believes that it is equally likely for player 3 to play either L or R, but 2 believes that it is strictly more likely that 3 will play L. In fact, (31) also implies that

$$\text{marg}_{A_3} \mathbf{\Phi}_1(\omega) \cap \text{marg}_{A_3} \mathbf{\Phi}_2(\omega) = \emptyset.$$

That is, even Klibanoff's notion of agreement as defined in (14) is violated.

In this example, the ultimate reason for the failure of agreement is that we do not have any property that looks like the sure-thing principle. Let $D = \{\omega_1, \omega_2\}$ and $E = \{\omega_3, \omega_4\}$. The set $\mathbf{\Phi}_1(\omega)$ represents beliefs conditional on D and on E. The set $\mathbf{\Phi}_2(\omega)$ represents beliefs conditional on $D \cup E$. The fact that beliefs conditional on D is the same as that conditional on E does not have any clear implication for beliefs conditional on $D \cup E$. $\diamondsuit$

We end this section by pointing out that Proposition 3 is preserved if the notion of knowledge is replaced by the following notion of belief.[7] Given any event $E \subseteq \Omega$, say that i *believes* E *at* ω if supp $\Delta_i(\omega) \subseteq E$. That is, i believes E at ω if $\Omega \backslash E$ is a null event for i at ω. An event E is *mutual belief at* ω if everyone believes E at ω. The notion of *common belief* is defined iteratively in the obvious manner. Under the assumption of common prior with contamination, Proposition 3 is equally valid if the condition "$[\Phi]$ is common knowledge" is replaced by "$[\Phi]$ is common belief". The reason is as follows. Unless $\Delta_i(\omega)$ is a singleton (which is the case already taken care of by Aumann and Brandenburger (1995)), the assumption of common prior with contamination implies that i believes E at ω if and only if i knows E at ω. Under the assumption of common prior with absolutely continuous contamination, Proposition 3 will also be valid if $\omega^* \in$ supp μ and if the condition "$[\Phi]$ is common knowledge" is replaced by "$[\Phi]$ is common belief". The reason is as follows. For every $\omega \in$ supp μ, define $\hat{H}_i(\omega) \equiv H_i(\omega) \cap$ supp μ. Clearly, $\hat{H}_i$ forms a partition of supp μ. Define $\hat{\mathcal{H}}$ to be the meet of $\{\hat{H}_i\}_{i=1}^n$. Then for every $\omega \in$ supp μ, i believes E at ω if and only if $\hat{H}_i(\omega) \subseteq E$, and E is common belief at ω if and only if $\hat{\mathcal{H}}(\omega) \subseteq E$. Therefore, by

[6]For other updating rules, similar examples can be constructed to convey essentially the same message.

[7]Similar explanation suffices to establish that Proposition 4 below is also preserved if knowledge is replaced by belief.

substituting $\hat{H}_i$ for H_i and $\hat{\mathcal{H}}$ for $\mathcal{H}$, the proof of Proposition 3 under the assumption of common prior with contamination also works for the case where contamination is absolutely continuous.

5. IMPLICATIONS OF THE MAIN RESULT

5.1 Agreement and Stochastic Independence

We can interpret Proposition 3 as delivering the two properties "agreement" and "stochastic independence" of beliefs as follows.

Agreement: Equation (26) in Proposition 3 implies that the marginal beliefs of players i and j on the action choice of player k are represented by

$$\begin{aligned} \text{marg}_{A_k}\Phi_i &= \{(1-\boldsymbol{\epsilon}_i(\omega^*))\sigma_k^* + \boldsymbol{\epsilon}_i(\omega^*)\sigma_k : \sigma_k \in M(\text{supp } \sigma_k^*)\} \text{ and} \\ \text{marg}_{A_k}\Phi_j &= \{(1-\boldsymbol{\epsilon}_j(\omega^*))\sigma_k^* + \boldsymbol{\epsilon}_j(\omega^*)\sigma_k : \sigma_k \in M(\text{supp } \sigma_k^*)\}, \end{aligned} \tag{32}$$

respectively. They share the same benchmark probability measure σ_k^* and the same set $M(\text{supp } \sigma_k^*)$ of contaminated measures. Note that (32) implies that for all $E \subseteq A_k$,

$$\min_{\sigma_k \in \text{marg}_{A_k}\Phi_i} \sigma_k(E) = \begin{cases} 1 & \text{if supp } \sigma_k^* \subseteq E \\ (1-\boldsymbol{\epsilon}_i(\omega^*))\sigma_k^*(E) & \text{otherwise.} \end{cases} \tag{33}$$

The condition "either $\boldsymbol{\epsilon}_i(\omega^*) \in [0,1)$ for all i or $\boldsymbol{\epsilon}_i(\omega^*) = 1$ for all i" in Proposition 3 and (33) imply (9), which is equivalent to the preference based condition for agreement in (8). Also note that (32) implies

$$\text{marg}_{A_k}\Phi_i \subseteq \text{marg}_{A_k}\Phi_j \text{ if and only if } \boldsymbol{\epsilon}_i(\omega^*) \leq \boldsymbol{\epsilon}_j(\omega^*).$$

Therefore, $\text{marg}_{A_k}\Phi_i$ need not be equal to $\text{marg}_{A_k}\Phi_j$, but they must have a nonempty intersection.

Stochastic independence: According to (26) in Proposition 3, the benchmark probability measure $\prod_{j\neq i}\sigma_j^*$ in Φ_i is a product measure and the domain of contamination is the product space $\times_{j\neq i}\text{supp } \sigma_j^*$. It is obvious that not all the probability measures in $M(\times_{j\neq i}\text{supp } \sigma_j^*)$ are product measures. However, all those "correlated contaminations" are irrelevant in the following sense. According to the multiple priors model, a decision maker only cares about the minimum expected utility of acts. Given any action a_i of player i, there is always a degenerate probability measure δ_i in $M(\times_{j\neq i}\text{supp } \sigma_j^*)$ such that the minimum utility of a_i is attained at δ_i. Of course, all degenerate probability measures are product measures. Also, (26) implies that for all $E_j \subseteq A_j$ and $E_{-ij} \subseteq \times_{k\neq i,j}A_k$,

$$\min_{\phi_i \in \Phi_i} \phi_i(E_j \times E_{-ij}) = \begin{cases} 1 & \text{if } \times_{j\neq i}\text{supp } \sigma_j^* \subseteq E_j \times E_{-ij} \\ (1-\boldsymbol{\epsilon}_i(\omega^*))\sigma_j(E_j)\prod_{k\neq i,j}\sigma_k(E_{-ij}) & \text{otherwise.} \end{cases} \tag{34}$$

If $\boldsymbol{\epsilon}_i(\omega^*) \in [0,1)$, then (34) implies (11), which is equivalent to the preference based condition for stochastic independence in (10).

5.2 Support of Contaminated Beliefs

In this section, we explain precisely the sense in which Proposition 3 implies that the controversy on how the support of ambiguous beliefs is defined does not arise.

Conventionally, the support of player i's beliefs is defined to be the event $E_{-i} \subseteq A_{-i}$ such that $A_{-i} \backslash E_{-i}$ is the largest null event for i. When i's preference ordering is represented by the multiple priors model and therefore beliefs are represented by a set Φ_i of probability measures on A_{-i}, the event which satisfies this definition is supp Φ_i. Dow and Werlang (1994) argue that this notion of support is too strong. They propose the following definition: an event E_{-i} is a support of Φ_i if (i) there exists $\hat{\phi}_i \in \Phi_i$ such that $E_{-i} =$ supp $\hat{\phi}_i$ and (ii) for all $\phi_i \in \Phi_i$ and for all $F_{-i} \subset E_{-i}$, $\phi_i(F_{-i}) < 1$. Clearly, if E_{-i} is a support of Φ_i in the sense of Dow and Werlang, then $E_{-i} \subseteq$ supp Φ_i. Variants of Dow and Werlang's definition of support can be found in Lo (1995), Marinacci (1995) and Ryan (1998).

To see why the definition of support matters, consider the following two person strategic game.

22[Player 1][Player 2]	L	R
U	$10, 10$	$-10, 9$
D	$9, 10$	$8, 9$

Suppose we require the beliefs of the players to satisfy the best response property. That is, if an action of player i is in the support of player j's beliefs, then the action must be optimal for player i given i's beliefs. Since L is the unique best response for player 2 whatever her beliefs, the best response property implies that the support of player 1's beliefs must be L. If we adopt the conventional definition of support, then R will be a null event for player 1 and therefore his unique best response will be U. On the other hand, if we adopt Dow and Werlang's definition of support, R need not be a null event. As a result, player 1 may find it optimal to take the action D.

The ultimate reason for the above controversy is that different probability measures in Φ_i may have different supports. For any

$$\Phi_i = \{(1 - \epsilon_i)\phi_i^* + \epsilon_i \phi_i : \phi_i \in M(E_{-i})\},$$

where $\epsilon_i \in (0, 1)$, all the probability measures in Φ_i have the same support if and only if supp $\phi_i^* = E_{-i}$.[8] Therefore, Proposition 3 has the following corollary.

Corollary of Proposition 3. *Suppose $\{\Delta_i\}_{i=1}^n$ satisfies the assumption of common prior with (absolutely continuous) contamination. Suppose that at a state ω^*, $\boldsymbol{\epsilon}_i(\omega^*) \in [0, 1)$ for all i, and $[\Phi]$ is common knowledge. Then all the probability measures in Φ_i have the same support.*

It is immediate that if the contamination in $\Delta_i(\omega^*)$ is absolutely continuous, then the contamination in $\boldsymbol{\Phi}_i(\omega^*)$ must also be absolutely continuous. Therefore, common knowledge of beliefs is needed only if $\{\Delta_i\}_{i=1}^n$ satisfies the assumption of common prior with contamination that is *not* absolutely continuous. We outline the proof of the corollary as follows. For simplicity, consider a

[8]The event E_{-i} is the support of Φ_i according to the conventional definition, and supp ϕ_i^* is the support of Φ_i according to all the other definitions proposed in the literature. This fact is observed independently by Ryan (1998).

game with two players and player 1's beliefs $\Phi_1 = \{(1-\epsilon_1)\phi_1^* + \epsilon_1\phi_1 : \phi_1 \in M(E_2)\}$. Suppose Φ_1 is common knowledge at ω^*. This implies that at every state $\omega \in \mathcal{H}(\omega^*)$, player 1's beliefs are in fact represented by Φ_1. By definition, supp $\phi_1^* \subseteq E_2$. Suppose supp $\phi_1^* \not= E_2$. That is, there exists $a_2 \in E_2$ such that $a_2 \not\in$supp ϕ_1^*. The fact that $a_2 \in E_2$ implies

$$\{\omega \in \mathcal{H}(\omega^*) : \mathbf{a}_2(\omega) = a_2\} \not= \emptyset. \tag{35}$$

Since ϕ_1^* comes from the common prior μ, the fact that $a_2 \not\in$supp ϕ_1^* implies

$$\mu(\{\omega \in \mathcal{H}(\omega^*) : \mathbf{a}_2(\omega) = a_2\}) = 0. \tag{36}$$

Equation (35) implies that there is a partitional element $H_2(\omega) \in H_2$ such that $H_2(\omega) \subseteq \{\omega \in \mathcal{H}(\omega^*) : \mathbf{a}_2(\omega) = a_2\}$. However, (36) implies that $\mu(H_2(\omega)) = 0$, contradicting the hypothesis that $\mu(H_i(\omega)) > 0$ for all ω and for all i.

Example 5 below demonstrates that if $[\Phi_i]$ is not common knowledge at ω^*, then different probability measures in Φ_i may have different supports.

Example 5. Consider the following

- strategic game form $\times_{i=1}^2 A_i$, where $A_1 = \{U, D\}$ and $A_2 = \{L, C, R\}$
- set of states of the world $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$
- common prior $\mu = (\omega_1, \frac{1}{3}; \omega_2, 0; \omega_3, 0; \omega_4, \frac{1}{3}; \omega_5, \frac{1}{3}; \omega_6, 0)$.
- information partitions

$$H_1 = \{\{\omega_1, \omega_2, \omega_3\}, \{\omega_4, \omega_5, \omega_6\}\} \text{ and } H_2 = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \{\omega_5, \omega_6\}\}$$

- action functions

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
$\mathbf{a}_1$	U	U	U	D	D	D
$\mathbf{a}_2$	L	L	C	C	R	R

Suppose $\{\Delta_i\}_{i=1}^2$ satisfies the assumption of common prior with contamination, where the common prior μ is specified above. It is clear that $[\mathbf{\Phi}_2(\omega_3)]$ is not common knowledge at ω_3. Unless $\boldsymbol{\epsilon}_2(\omega_3) = 0$, different probability measures in $\mathbf{\Phi}_2(\omega_3)$ may have different supports.

5.3 Generalization of Nash Equilibrium

Consider the following generalization of Nash equilibrium.

Definition 3. Given a game u, Φ is a *Nash equilibrium with contamination* of u if the following conditions are satisfied.

1. There exist $\{\sigma_i^*\}_{i=1}^n$ and $\{\epsilon_i\}_{i=1}^n$, where $\sigma_i^* \in M(A_i)$ and either $\epsilon_i \in [0,1)$ for all i or $\epsilon_i = 1$ for all i, such that

$$\Phi_i = \left\{ (1-\epsilon_i) \prod_{j \neq i} \sigma_j^* + \epsilon_i \phi_i : \phi_i \in M(\times_{j\neq i} \text{supp } \sigma_j^*) \right\} \quad \forall i.$$

2. supp $\text{marg}_{A_i} \Phi_j \subseteq \arg\max_{a_i \in A_i} \min_{\phi_i \in \Phi_i} \sum_{a_{-i} \in \times_{j\neq i} A_j} u_i(a_i, a_{-i}) \phi_i(a_{-i}) \quad \forall i \quad \forall j \neq i.$

Definition 3 consists of two conditions. Condition 1 is the property of agreement and stochastic independence of contaminated beliefs that we derive in Proposition 3. The interpretation of condition 2 parallels that of the best response property of Nash equilibrium. That is, player i chooses some definite action, but player j may not know which one. However, if j knows i's beliefs Φ_i and that i is rational, then every action a_i in the support of j's marginal beliefs $\text{marg}_{A_i}\Phi_j$ must be a best response for i given i's beliefs Φ_i. Therefore, Nash equilibrium with contamination is viewed as an equilibrium in beliefs.

The event $\{\omega \in \Omega : \text{every } i \text{ is rational at } \omega\}$ is denoted as [rationality]. Given Proposition 3, straightforward adaptation of Theorem A in Aumann and Brandenburger (1995) provides epistemic conditions for Nash equilibrium with contamination, confirming the above interpretation.

Proposition 4. *Suppose $\{\Delta_i\}_{i=1}^n$ satisfies the assumption of common prior with (absolutely continuous) contamination. Suppose that at a state ω^*, either $\boldsymbol{\epsilon}_i(\omega^*) \in [0,1)$ for all i or $\boldsymbol{\epsilon}_i(\omega^*) = 1$ for all i, [rationality] and $[u]$ are mutual knowledge and $[\Phi]$ is common knowledge. Then Φ is a Nash equilibrium with contamination of u.*

In our earlier paper (Lo, 1996), we propose an equilibrium concept which is called *Beliefs Equilibrium with Agreement* (p. 453, Definition 4). The equilibrium concept is a generalization of Nash equilibrium by allowing preferences to be represented by the multiple priors model. Under the assumption that players are completely ignorant, we provide epistemic conditions for Beliefs Equilibrium with Agreement (p. 471, Proposition 7). This earlier result is essentially a special case of Proposition 4 in the following sense. Proposition 4 only requires that either $\epsilon_i(\omega^*) \in [0,1)$ for all i or $\epsilon_i(\omega^*) = 1$ for all i. On the other hand, Proposition 7 requires $\epsilon_i(\omega) = 1$ for all i *and* for all $\omega \in \Omega$.

APPENDIX

Proof of Proposition 1

To simplify notation, we restate and prove Proposition 1 in the context of single person decision making. That is, suppose there is a decision maker facing a product state space $S_1 \times S_2$.

Proposition 1. Suppose there exist $E \subseteq S_1 \times S_2$ with $|\text{proj}_{S_1} E| > 1$ and $|\text{proj}_{S_2} E| > 1$, $p^* \in M(E)$, and $\epsilon \in (0,1)$ such that $\mathcal{P} = \{(1-\epsilon)p^* + \epsilon p : p \in M(E)\}$. Then $\mathcal{P}$ does not satisfy the independent product property.

Proof. Fix $\hat{s}_1 \in \text{proj}_{S_1} E$ such that $\text{marg}_{S_1} p^*(\hat{s}_1) > 0$. We have

$$\begin{aligned}
\min_{p_1 \in \text{marg}_{S_1} \mathcal{P}} p_1(\hat{s}_1) &= \min_{p \in \mathcal{P}} p(\hat{s}_1 \times \text{proj}_{S_2} E) && (37)\\
&= (1-\epsilon) p^*(\hat{s}_1 \times \text{proj}_{S_2} E) && (38)\\
&= (1-\epsilon) \text{marg}_{S_1} p^*(\hat{s}_1). && (39)
\end{aligned}$$

The equality between (37) and (38) is due to the fact that $E \not\subseteq \hat{s}_1 \times \text{proj}_{S_2} E$, since $|\text{proj}_{S_1} E| > 1$. Similarly, $|\text{proj}_{S_2} E| > 1$ implies that for all $s_2 \in \text{proj}_{S_2} E$,

$$\min_{p_2 \in \text{marg}_{S_2} \mathcal{P}} p_2(s_2) = (1-\epsilon) \text{marg}_{S_2} p^*(s_2). \quad (40)$$

Suppose $\mathcal{P}$ had the independent product property. Then for all $s_2 \in \text{proj}_{S_2} E$,

$$\begin{aligned}
\min_{p \in \mathcal{P}} p(\hat{s}_1 \times s_2) &= \min_{p_1 \in \text{marg}_{S_1} \mathcal{P}} p_1(\hat{s}_1) \min_{p_2 \in \text{marg}_{S_2} \mathcal{P}} p_2(s_2) && (41)\\
&= (1-\epsilon) \text{marg}_{S_1} p^*(\hat{s}_1)(1-\epsilon) \text{marg}_{S_2} p^*(s_2) && (42)\\
&= (1-\epsilon)^2 \text{marg}_{S_1} p^*(\hat{s}_1) \text{marg}_{S_2} p^*(s_2). && (43)
\end{aligned}$$

The independent product property implies (41). The equality between (41) and (42) is due to (39) and (40). According to (43), if we sum up $\min_{p \in \mathcal{P}} p(\hat{s}_1 \times s_2)$ over all the elements in $\text{proj}_{S_2} E$, we will have

$$\begin{aligned}
\sum_{s_2 \in \text{proj}_{S_2} E} \min_{p \in \mathcal{P}} p(\hat{s}_1 \times s_2) &= \sum_{s_2 \in \text{proj}_{S_2} E} (1-\epsilon)^2 \text{marg}_{S_1} p^*(\hat{s}_1) \text{marg}_{S_2} p^*(s_2)\\
&= (1-\epsilon)^2 \text{marg}_{S_1} p^*(\hat{s}_1) \sum_{s_2 \in \text{proj}_{S_2} E} \text{marg}_{S_2} p^*(s_2)\\
&= (1-\epsilon)^2 \text{marg}_{S_1} p^*(\hat{s}_1). && (44)
\end{aligned}$$

However, regardless of the independent product property, it is always legitimate to write

$$\begin{aligned}
\sum_{s_2 \in \text{proj}_{S_2} E} \min_{p \in \mathcal{P}} p(\hat{s}_1 \times s_2) &= \sum_{s_2 \in \text{proj}_{S_2} E} (1-\epsilon) p^*(\hat{s}_1 \times s_2)\\
&= (1-\epsilon) \sum_{s_2 \in \text{proj}_{S_2} E} p^*(\hat{s}_1 \times s_2)\\
&= (1-\epsilon) p^*(\hat{s}_1 \times \text{proj}_{S_2} E)\\
&= (1-\epsilon) \text{marg}_{S_1} p^*(\hat{s}_1). && (45)
\end{aligned}$$

Since $\epsilon \in (0,1)$ and $\text{marg}_{S_1} p^*(\hat{s}_1) > 0$, (44) and (45) constitute a contradiction.

Proof of Proposition 2.

It is obvious that Proposition 2 is true if either $\epsilon_i = 0$ or $\epsilon_i = 1$. We assume below that $\epsilon_i \in (0,1)$.

Given the set Λ_i of probability measures defined in (22), the maximum likelihood updating rule defined in (23) can be rewritten as

$$\begin{aligned}\Delta_i(\omega) = \{\mu_i \in M(H_i(\omega)) : \exists p \in \Lambda_i \text{ such that } \\ p(H_i(\omega)) = \epsilon_i + (1-\epsilon_i)\mu(H_i(\omega)) \text{ and } \mu_i(\cdot) = p(\cdot|H_i(\omega))\}.\end{aligned} \tag{46}$$

Equation (46) can in turn be rewritten as

$$\begin{aligned}\Delta_i(\omega) &= \left\{\frac{(1-\epsilon_i)\mu(H_i(\omega))\mu(\cdot|H_i(\omega)) + \epsilon_i\mu_i}{\epsilon_i + (1-\epsilon_i)\mu(H_i(\omega))} : \mu_i \in M(H_i(\omega))\right\} \\ &= \{(1-\boldsymbol{\epsilon}_i(\omega))\mu(\cdot \mid H_i(\omega)) + \boldsymbol{\epsilon}_i(\omega)\mu_i : \mu_i \in M(H_i(\omega))\},\end{aligned} \tag{47}$$

where

$$\boldsymbol{\epsilon}_i(\omega) = \frac{\epsilon_i}{\epsilon_i + (1-\epsilon_i)\mu(H_i(\omega))}.$$

Similarly, (22) implies that the full Bayesian updating rule defined in (24) can be rewritten as

$$\begin{aligned}\Delta_i(\omega) = \{\mu_i \in M(H_i(\omega)) : \exists p \in \Lambda_i \text{ and } \exists \alpha \in [0,1] \text{ such that } \\ p(H_i(\omega)) = \epsilon_i\alpha + (1-\epsilon_i)\mu(H_i(\omega)) \text{ and } \mu_i(\cdot) = p(\cdot|H_i(\omega))\}.\end{aligned} \tag{48}$$

Equation (48) can in turn be rewritten as

$$\begin{aligned}\Delta_i(\omega) &= \bigcup_{\alpha\in[0,1]} \left\{\frac{(1-\epsilon_i)\mu(H_i(\omega))\mu(\cdot|H_i(\omega)) + \epsilon_i\alpha\mu_i}{\epsilon_i\alpha + (1-\epsilon_i)\mu(H_i(\omega))} : \mu_i \in M(H_i(\omega))\right\} \\ &= \bigcup_{\alpha\in[0,1]} \{(1-\boldsymbol{\epsilon}_i^{\alpha}(\omega))\mu(\cdot \mid H_i(\omega)) + \boldsymbol{\epsilon}_i^{\alpha}(\omega)\mu_i : \mu_i \in M(H_i(\omega))\},\end{aligned} \tag{49}$$

where

$$\boldsymbol{\epsilon}_i^{\alpha}(\omega) = \frac{\epsilon_i\alpha}{\epsilon_i\alpha + (1-\epsilon_i)\mu(H_i(\omega))}.$$

Note that

$$\boldsymbol{\epsilon}_i^{\alpha}(\omega) \begin{cases} < \boldsymbol{\epsilon}_i(\omega) & \text{if } 0 \le \alpha < 1 \\ = \boldsymbol{\epsilon}_i(\omega) & \text{if } \alpha = 1. \end{cases} \tag{50}$$

Recall the property that given any set $\mathcal{P} = \{(1-\epsilon)p^* + \epsilon p : p \in M(E)\}$ of probability measures, $\mathcal{P}$ increases in the sense of set inclusion as ϵ increases. Therefore, (50) implies that (47) and (49) are equivalent.

Proof of Proposition 3

We provide below the proof of Proposition 3 under the assumption of common prior with contamination.

The assumption of common prior with contamination implies that $\boldsymbol{\Phi}_i(\omega)$ defined in (20) can be rewritten as

$$\boldsymbol{\Phi}_i(\omega) = \{(1-\boldsymbol{\epsilon}_i(\omega))\phi^i(\cdot \mid H_i(\omega)) + \boldsymbol{\epsilon}_i(\omega)\phi_i : \phi_i \in M(\{\mathbf{a}_{-i}(\omega') : \omega' \in H_i(\omega)\})\} \quad \forall \omega \in \Omega, \tag{51}$$

where $\phi^i(\cdot \mid H_i(\omega)) \in M(\times_{j\neq i}A_j)$ is derived from $\mu(\cdot \mid H_i(\omega))$ in the same manner as $\mathbf{\Phi}_i(\omega)$ is derived from $\Delta_i(\omega)$. By hypothesis, $[\Phi_i]$ is common knowledge at ω^*. This implies that $\mathbf{\Phi}_i(\omega^*) = \Phi_i$ and therefore $[\mathbf{\Phi}_i(\omega^*)]$ is common knowledge at ω^*. That is,

$$\mathbf{\Phi}_i(\omega) = \mathbf{\Phi}_i(\omega^*) \quad \forall \omega \in \mathcal{H}(\omega^*). \tag{52}$$

This in turn implies that

$$\text{supp } \mathbf{\Phi}_i(\omega) = \text{supp } \mathbf{\Phi}_i(\omega^*) \quad \forall \omega \in \mathcal{H}(\omega^*). \tag{53}$$

There are two cases to consider.

Case 1: $\boldsymbol{\epsilon}_i(\omega^*) \in [0, 1)$ for all i. In this case, it suffices to establish that there exists a collection of probability measures $\{\sigma_i^*\}_{i=1}^n$, where $\sigma_i^* \in A_i$, such that $\phi^i(\cdot \mid H_i(\omega^*)) = \prod_{j\neq i} \sigma_j^*$ and $\{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\} = \times_{j\neq i}\text{supp } \sigma_j^*$.

For every $\mathbf{\Phi}_i(\omega^*)$ such that $|\mathbf{\Phi}_i(\omega^*)| = 1$, (52) implies that $|\mathbf{\Phi}_i(\omega)| = 1$ for all $\omega \in \mathcal{H}(\omega^*)$. Then (51) implies that $\mathbf{\Phi}_i(\omega) = \phi^i(\cdot \mid H_i(\omega))$ for all $\omega \in \mathcal{H}(\omega^*)$. Therefore, (52) can be rewritten as

$$\phi^i(\cdot \mid H_i(\omega)) = \phi^i(\cdot \mid H_i(\omega^*)) \quad \forall \omega \in \mathcal{H}(\omega^*). \tag{54}$$

For every $\mathbf{\Phi}_i(\omega^*)$ such that $|\mathbf{\Phi}_i(\omega^*)| > 1$, (52) implies that $|\mathbf{\Phi}_i(\omega)| > 1$ for all $\omega \in \mathcal{H}(\omega^*)$. Then (51) implies that supp $\mathbf{\Phi}_i(\omega) = \{\mathbf{a}_{-i}(\omega') : \omega' \in H_i(\omega)\}$ for all $\omega \in \mathcal{H}(\omega^*)$. Therefore, (53) can be rewritten as

$$\{\mathbf{a}_{-i}(\omega') : \omega' \in H_i(\omega)\} = \{\mathbf{a}_{-i}(\omega') : \omega' \in H_i(\omega^*)\} \quad \forall \omega \in \mathcal{H}(\omega^*). \tag{55}$$

Then (51), (52) and (55) imply that

$$\boldsymbol{\epsilon}_i(\omega) = \boldsymbol{\epsilon}_i(\omega^*) \text{ and } \phi^i(\cdot \mid H_i(\omega)) = \phi^i(\cdot \mid H_i(\omega^*)) \quad \forall \omega \in \mathcal{H}(\omega^*). \tag{56}$$

According to (54) and (56), whatever the cardinality of $\mathbf{\Phi}_i(\omega^*)$, $[\phi^i(\cdot \mid H_i(\omega^*)]$ is common knowledge at ω^*. Since $\phi^i(\cdot \mid H_i(\omega))$ comes from a common prior for all $\omega \in \Omega$, Theorem B in Aumann and Brandenburger (1995) can be invoked to conclude that there exists a collection of probability measures $\{\sigma_i^*\}_{i=1}^n$, where $\sigma_i^* \in A_i$, such that $\phi^i(\cdot \mid H_i(\omega^*)) = \prod_{j\neq i} \sigma_j^*$.

Since $\phi^i(\cdot \mid H_i(\omega^*)) \in M(\{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\})$ and we have just established that $\phi^i(\cdot \mid H_i(\omega^*)) = \prod_{j\neq i} \sigma_j^*$ is a product measure, we must have $\times_{j\neq i}\text{supp } \sigma_j^* \subseteq \{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\}$. Let us establish by contradiction that we actually have $\times_{j\neq i}\text{supp } \sigma_j^* = \{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\}$. Suppose that this were not true. That is, suppose there exists $a_j \in \text{proj}_{A_j}\{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\}$ such that $a_j \not\in \text{supp } \sigma_j^*$. Recall that $[\phi^i(\cdot \mid H_i(\omega^*)]$ is common knowledge at ω^* and that $\sigma_j^* = \text{marg}_{A_j}\phi^i(\cdot \mid H_i(\omega^*))$ is derived from the common prior μ. Therefore, if $a_j \not\in \text{supp } \sigma_j^*$, then

$$\mu(\{\omega \in \mathcal{H}(\omega^*) : \mathbf{a}_j(\omega) = a_j\}) = 0. \tag{57}$$

The fact that $a_j \in \text{proj}_{A_j}\{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\}$ implies that $\{\omega \in \mathcal{H}(\omega^*) : \mathbf{a}_j(\omega) = a_j\} \neq \emptyset$. By the definition of $\mathcal{H}(\omega^*)$, there exists $\omega' \in \mathcal{H}(\omega^*)$ such that $H_j(\omega') \subseteq \{\omega \in \mathcal{H}(\omega^*) : \mathbf{a}_j(\omega) = a_j\}$. However, (57) implies that $\mu(H_j(\omega')) = 0$. This contradicts the assumption that $\mu(H_i(\omega)) > 0$ for all $\omega \in \Omega$ and for all i.

Case 2: $\boldsymbol{\epsilon}_i(\omega^*) = 1$ for all i. In this case, (51) implies that

$$\mathbf{\Phi}_i(\omega^*) = M(\{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\}). \tag{58}$$

Therefore, we have to establish that there exists a collection of sets $\{E_i\}_{i=1}^n$, where $E_i \subseteq A_i$, such that $\{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\} = \times_{j \neq i} E_j$.

For every $\mathbf{\Phi}_i(\omega^*)$ such that $|\mathbf{\Phi}_i(\omega^*)| = 1$, (58) implies that there exists $a_{-i} \in A_{-i}$ such that $\{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\} = \{a_{-i}\}$. Then (52) implies that $\mathbf{\Phi}_i(\omega)$ is the degenerate probability measure on a_{-i} for all $\omega \in \mathcal{H}(\omega^*)$. This and (51) imply that

$$\mu(\{\omega \in \mathcal{H}(\omega^*) : \mathbf{a}_{-i}(\omega) \neq a_{-i}\}) = 0. \tag{59}$$

The definition of $\mathcal{H}(\omega^*)$, the assumption of $\mu(H_j(\omega)) > 0$ for all $\omega \in \Omega$ and (59) imply that we actually have $\mathbf{a}_{-i}(\omega) = a_{-i}$ for all $\omega \in \mathcal{H}(\omega^*)$. This in turn implies (55). For every $\mathbf{\Phi}_i(\omega^*)$ such that $|\mathbf{\Phi}_i(\omega^*)| > 1$, we can conclude the validity of (55) in exactly the same way as in Case 1 above. Equation (55) implies that

$$\text{proj}_{A_j}\{\mathbf{a}_{-i}(\omega') : \omega' \in H_i(\omega)\} = \text{proj}_{A_j}\{\mathbf{a}_{-i}(\omega') : \omega' \in H_i(\omega^*)\} \quad \forall \omega \in \mathcal{H}(\omega^*). \tag{60}$$

Since $H_k(\omega^*) \subseteq \mathcal{H}(\omega^*)$, (60) implies that

$$\text{proj}_{A_j}\{\mathbf{a}_{-k}(\omega') : \omega' \in H_k(\omega^*)\} \subseteq \text{proj}_{A_j}\{\mathbf{a}_{-i}(\omega') : \omega' \in H_i(\omega^*)\}. \tag{61}$$

Since i, j and k are arbitrary, (61) implies that

$$\text{proj}_{A_j}\{\mathbf{a}_{-k}(\omega') : \omega' \in H_k(\omega^*)\} = \text{proj}_{A_j}\{\mathbf{a}_{-i}(\omega') : \omega' \in H_i(\omega^*)\}. \tag{62}$$

Therefore, it is legitimate to define $E_j \equiv \text{proj}_{A_j}\{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\}$.

Clearly, $\{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\} \subseteq \times_{j \neq i} E_j$. Let us establish by contradiction that $\{\mathbf{a}_{-i}(\omega) : \omega \in H_i(\omega^*)\} = \times_{j \neq i} E_j$. Without loss of generality, suppose that for player 1, there exists $a_{-1} \in \times_{j \neq 1} E_j$ such that $a_{-1} \notin \{\mathbf{a}_{-1}(\omega) : \omega \in H_i(\omega^*)\}$. Then (55) implies that

$$\mathbf{a}_{-1}(\omega) \neq a_{-1} \quad \forall \omega \in \mathcal{H}(\omega^*). \tag{63}$$

Next consider player 2. Equation (63) implies that for every $\omega \in \mathcal{H}(\omega^*)$ such that $\mathbf{a}_2(\omega) = a_2$ (such an ω must exist, since $a_2 \in E_2$), there exists $j \in \{3, \ldots, n\}$ such that $\mathbf{a}_j(\omega) \neq a_j$. We can again invoke (55) to conclude that for each $\omega \in \mathcal{H}(\omega^*)$, there exists $j \in \{3, \ldots, n\}$ such that $\mathbf{a}_j(\omega) \neq a_j$. Repeat the same argument for the rest of the players to reach the conclusion that for each $\omega \in \mathcal{H}(\omega^*)$, $\mathbf{a}_n(\omega) \neq a_n$. This contradicts the fact that $a_n \in E_n$.

Proof of Proposition 4

The proof of Proposition 3 already implies that condition 1 in Definition 3 is satisfied. The rest of the proof is essentially the same as that of Theorem A in Aumann and Brandenburger (1995). For the sake of completeness, we provide the proof below.

Denote the event $\{\omega \in \Omega : i \text{ is rational at } \omega\}$ by [i is rational]. By hypothesis, j knows [i is rational], $[u_i]$ and $[\Phi_i]$. That is, $H_j(\omega^*) \subseteq [i \text{ is rational}]$, $H_j(\omega^*) \subseteq [u_i]$ and $H_j(\omega^*) \subseteq [\Phi_i]$, respectively. Therefore,

$$H_j(\omega^*) \subseteq [i \text{ is rational}] \cap [u_i] \cap [\Phi_i]. \tag{64}$$

At ω^*, j knows $[\Phi_j]$ implies that $\mathbf{\Phi}_j(\omega^*) = \Phi_j$ and therefore

$$\text{supp marg}_{A_i} \Phi_j \subseteq \{\mathbf{a}_i(\omega) : \omega \in H_j(\omega^*)\}. \tag{65}$$

Equations (64) and (65) imply that condition 2 in Definition 3 is satisfied. Therefore, Φ is a Nash equilibrium with contamination of u.

REFERENCES

Aumann, R. J. (1976): "Agreeing to Disagree," *Annals of Statistics*, 4, 1236-1239.

Aumann, R. J. and A. Brandenburger (1995): "Epistemic Conditions for Nash Equilibrium," *Econometrica*, 63, 1161-1180.

Berger, J. (1984): "The Robust Bayesian Viewpoint," *Robustness in Bayesian Statistics*, edited by J. Kadane. North-Holland, Amsterdam.

Berger, J. (1985): *Statistical Decision Theory and Bayesian Analysis*. Springer, New York.

Berger, J. and L. M. Berliner (1986): "Robust Bayes and Empirical Bayes Analysis with ϵ-Contaminated Priors," *Annals of Statistics*, 14, 461-486.

Bewley, T. F. (1986): "Knightian Decision Theory: Part I," Cowles Foundation Discussion Paper No. 807, New Haven, CT.

Bewley, T. F. (1987): "Knightian Decision Theory: Part II," Cowles Foundation Discussion Paper No. 835, New Haven, CT.

Bewley, T. F. (1988): "Knightian Decision Theory and Econometric Inference," Cowles Foundation Discussion Paper No. 868, New Haven, CT.

Bewley, T. F. (1989): "Market Innovation and Entrepreneurship: A Knightian View," Cowles Foundation Discussion Paper No. 905, New Haven, CT.

Blume, L., A. Brandenburger and E. Dekel (1991): "Lexicographic Probabilities and Choice Under Uncertainty," *Econometrica*, 59, 61-79.

Brandenburger, A. (1992): "Knowledge and Equilibrium in Games," *Journal of Economic Perspectives*, 6, 83-101.

Camerer, C., and M. Weber (1992): "Recent Developments in Modelling Preference: Uncertainty and Ambiguity," *Journal of Risk and Uncertainty*, 5, 325-370.

Dempster, A. (1967): "Upper and Lower Probabilities Induced from a Multivalued Mapping," *Annals of Mathematical Statistics*, 38, 325-339.

Dow, J. and S. Werlang (1992): "Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio," *Econometrica*, 60, 197-204.

Dow, J. and S. Werlang (1994): "Nash Equilibrium under Knightian Uncertainty: Breaking Down Backward Induction," *Journal of Economic Theory*, 64, 305-324.

Eichberger, J. and D. Kelsey (1994): "Non-additive Beliefs and Game Theory," CentER for Economic Research Discussion Paper, No. 9410, Tilburg University.

Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643-669.

Epstein, L. G. (1997): "Preference, Rationalizability and Equilibrium," *Journal of Economic Theory*, 73, 1-29.

Epstein, L. G. and T. Wang (1994): "Intertemporal Asset Pricing under Knightian Uncertainty," *Econometrica*, 62, 283-322.

Epstein, L. G. and J. Zhang (1997): "Least Convex Capacities," *Economic Theory*, forthcoming.

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

Gilboa, I. and D. Schmeidler (1989): "Maxmin Expected Utility with Non-unique Prior," *Journal of Mathematical Economics*, 18, 141-153.

Gilboa, I. and D. Schmeidler (1993): "Updating Ambiguous Beliefs," *Journal of Economic Theory*, 59, 33-49.

Ghirardato, P. (1997): "On Independence for Non-Additive Measures, with a Fubini Theorem," *Journal of Economic Theory*, 73, 261-291.

Hendon, E., H. J. Jacobsen, B. Sloth and T. Tranæs: (1996): "The Product of Capacities and Belief Functions," *Mathematical Social Sciences*, 32, 95-108.

Klibanoff, P. (1993): "Uncertainty, Decision, and Normal Form Games," Manuscript, MIT.

Lo, K. C. (1995): "Nash Equilibrium without Mutual Knowledge of Rationality," Manuscript, University of Toronto.

Lo, K. C. (1996): "Equilibrium in Beliefs under Uncertainty," *Journal of Economic Theory*, 71, 443-484.

Marinacci, M. (1995): "Equilibrium in Ambiguous Games," Manuscript, Northwestern University.

Mukerji, S. (1995): "A Theory of Play for Games in Strategic Form When Rationality is not Common Knowledge," Discussion Papers in Economics and Econometrics, No. 9519, University of Southampton.

Osborne, M. J. and A. Rubinstein (1994): *Game Theory.* MIT Press.

Ryan, M. (1998): "Supports and the Updating of Capacities," Manuscript, University of Auckland.

Savage, L. (1954): *The Foundations of Statistics.* New York: John Wiley.

Schmeidler, D. (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica*, 57, 571-581.

Shafer, G. (1976): *A Mathematical Theory of Evidence.* Princeton University Press.

Wasserman, L. (1990): "Prior Envelopes Based on Belief Functions," *Annals of Statistics*, 18, 454-464.