WILEY

ECONOMETRICA
JOURNAL OF THE ECONOMETRIC SOCIETY

"Beliefs about Beliefs" without Probabilities
Author(s): Larry G. Epstein and Tan Wang
Source: *Econometrica*, Vol. 64, No. 6 (Nov., 1996), pp. 1343-1373
Published by: Econometric Society
Stable URL: http://www.jstor.org/stable/2171834

Econometrica, Vol. 64, No. 6 (November, 1996), 1343–1373

"BELIEFS ABOUT BELIEFS" WITHOUT PROBABILITIES

BY LARRY G. EPSTEIN AND TAN WANG[1]

This paper constructs a space of states of the world representing the exhaustive uncertainty facing each player in a strategic situation. The innovation is that preferences are restricted primarily by "regularity" conditions and need not conform with subjective expected utility theory. The construction employs a hierarchy of preferences, rather than of beliefs as in the standard Bayesian model. The framework is sufficiently general to accommodate uncertainty averse preferences, such as exhibited in the Ellsberg paradox, and to allow common knowledge of expected utility (or Choquet expected utility) to be well-defined formally. Applications include the provision of (i) foundations for a Harsanyi-style game of incomplete information, and (ii) a rich framework for the axiomatization of solution concepts for complete information normal form games.

KEYWORDS: Uncertainty, capacities, types space, beliefs hierarchies, common knowledge, Kolmogorov extension theorem, topologies for preferences, vague topology, rationalizability.

1. INTRODUCTION

THE DOMINANT APPROACH TO MODELING decision making under uncertainty is due to Savage. In this approach, one posits that the uncertainty can be represented by a space of *states of the world* that are comprehensive descriptions of the environment "leaving no relevant aspect undescribed" (Savage (1954, p. 9)). This paper is concerned with whether the existence of such a states space is justifiable in principle for a decision maker operating in a strategic environment. A special feature of a strategic environment is that each agent faces uncertainty not only about the primitive uncertainty corresponding to the true *state of nature*, but also about her opponents "type" representing all her relevant characteristics. It is well known that this feature leads naturally to concern with infinite hierarchies and a problem of infinite regress that must be resolved in order to justify the existence of a space of types and therefore a space of states of the world.

In typical formulations of games it is assumed that players are subjective expected utility maximizers. If it is further assumed that each player's von Neumann-Morgenstern index is common knowledge, then the only relevant characteristic of a player about which opponents are uncertain is her Bayesian prior probability measure over the relevant state space. Thus the analyst's attempt to describe the states of the world facing each player leads naturally to

[1] An earlier version of this paper was first circulated in June 1994 under the title "A 'Types' Space for Games of Incomplete Information with Non-Bayesian Players." We are grateful to the Social Sciences and Humanities Research Council of Canada for financial support and to Eddie Dekel, Aviad Heifetz, Bart Lipman, Kin Chung Lo, Stephen Morris, Mike Peters, James Redekop, and Lin Zhou for helpful discussions and comments. The suggestions of a co-editor and two referees led to substantial improvements in exposition.

an infinite hierarchy of probability measures representing beliefs. This problem of infinite regress in beliefs has been tackled rigorously by Mertens and Zamir (1985), who construct a well-defined topological space of "types" having the feature that the type of each player is comprehensive in that it identifies, up to a homeomorphism, joint beliefs about the state of nature and the types of opponents. It is apparent that the relevance of hierarchies of beliefs is not restricted to the framework of subjective expected utility maximization. Rather, it is valid as long as individual preferences are "based on beliefs" that are representable by a probability measure; Machina and Schmeidler (1992, 1995) define this class of preferences precisely and refer to them as "probabilistically sophisticated," or "Bayesian rational."

We have three primary reasons for being dissatisfied with the noted construction of a state space. The first is methodological. In Savage's model, states of the world logically *precede* the specification of axioms. Therefore, it is at the very least inconsistent with the Savage approach to specify the state space only *after* adopting a set of axioms for preferences, whether those underlying subjective expected utility maximization or the weaker set underlying Bayesian rationality. More importantly perhaps, the fact that axioms are presumed in the specification of states of the world makes it impossible to formalize them within the resulting framework. This methodological concern is pertinent even if the modeler wishes to assume that players are Bayesian rational and that this is common knowledge. Such an assumption must be understood informally in the standard Bayesian framework (Aumann (1987)).

Our second reason has to do with the decision-theoretic approach to game theory whose objective is to relate, at a formal level, our understanding of individual rationality on the one hand and strategic rationality on the other. At the individual level, though subjective expected utility maximization is undoubtedly the dominant model in economics, many economists would probably view axioms such as transitivity or monotonicity as more basic tenets of rationality than the Sure-Thing-Principle and other components of the Savage model. The implications of such more basic axioms for single agent decision making are well understood from abstract choice theory, but they have not been isolated in strategic settings. To do so, and thereby to narrow the gap between the formal modeling of rationality in the two settings, requires that one specify a state space for players in a game that presumes as few preference axioms as possible. (As will become apparent, we do *not* succeed in specifying states of the world prior to *all* axioms, but the axioms maintained in the states space that we construct are much weaker than those underlying probabilistic sophistication.)

Our final reason for dissatisfaction with the Mertens-Zamir types space is descriptive accuracy. Both expected utility theory and the generalization formulated by Machina and Schmeidler are inconsistent with aversion to vagueness or ambiguity such as exhibited in the Ellsberg paradox. In other words, neither model permits a distinction between *risk*, where probabilities (either objective or subjective) are available to guide choice, and *uncertainty*, where information is too vague to be summarized adequately by a probability measure. This paper is motivated in part by the presumption that uncertainty or vagueness are important in strategic situations and therefore that it is worthwhile generalizing received game theory so that the effects of uncertainty can be studied in principle.

Our main contribution is to show that the previously noted problem of infinite regress can be solved, and a space of types (states of the world) constructed, under weak maintained assumptions regarding players' preferences. As a result we show that the Savage *approach* to modeling choice under uncertainty, via acts over a set of states of the world, is separate from (most of) the Savage *axioms*. In terms of execution, at our level of generality hierarchies of beliefs are no longer adequate or even relevant. Indeed, preferences need not even have a separable component that can be thought of as "beliefs"; in particular, they need not satisfy Savage's Axiom P4 which delivers a "more likely than" relation over events. Rather, the relevant hierarchy is one of *preferences* on suitable domains of Savage acts. We show that the types space so constructed is sufficiently rich to permit the formalization of the assumption that it is common knowledge that players' preferences conform to various alternative models of preference, such as expected utility theory or Choquet expected utility theory (Schmeidler (1989)). One application of our types space is to provide foundations for a Harsanyi-style formulation of games of incomplete information with non-Bayesian rational players. This application is described and others are outlined in Section 2.

To conclude this introduction, we relate our variation of the Mertens-Zamir construction to others that have been developed. The latter authors assume that the primitive space of states of nature Ω is compact Hausdorff. More recently, parallel constructions of beliefs hierarchies have been described under the alternative assumptions that this space is Polish (Brandenburger and Dekel (1993)) and Hausdorff (Heifetz (1993)). Our central result regarding the existence of a types space assumes that Ω is compact Hausdorff. The generality of our analysis lies in the significantly weaker maintained assumptions about individual preferences, as explained above. A difficulty in achieving this generalization is that the standard measure-theoretic machinery exploited by the above studies is not available to us. For example, they employ the topological space of Borel regular probability measures endowed with the weak convergence topology, which is well understood and has attractive properties. Our first task, prior to concern with any infinite regress, is to define a space of "regular" preferences, with suitable topology, that possesses comparable properties. This novel topological space of preferences is a secondary contribution of the paper. It is useful even if, for whatever reason, only finite hierarchies are relevant; we feel also that it will be useful more generally in modeling choice under uncertainty.

The notion of "knowledge" that we employ is an analogue of "belief with probability one," which is the notion employed in the papers just cited. Other studies adopt an information-theoretic approach to modeling knowledge, whereby the primitive is a set of possibilities indicating the set of states viewed as possible by the agent in question (see Fagin, Geanakoplos, Halpern, and

Vardi (1992) and Heifetz and Samet (1993), for example). These studies *cannot* be viewed as special cases of ours. They show that generally transfinite hierarchies are needed to provide a complete description of the uncertainty facing each agent.[2] On the other hand, our framework shares with the probability-theoretic literature cited the feature that a denumerable construction suffices for a types space, given a compact Hausdorff parameter space. As a result, our analysis casts further light on "when" a denumerable construction suffices.

We proceed as follows: The next section provides an informal outline of the hierarchical structure and infinite regress problem that arise in attempting to describe the uncertainty facing players in a game. It then describes three instances where resolution of this infinite regress problem is important. These applications serve to motivate the ensuing formal analysis. Section 3 describes the class of individual preferences that we admit and the topology adopted for the space of preferences. Some noteworthy subspaces of preferences are studied in Section 4. The construction of the space of types is carried out in Section 5 and the incorporation of common knowledge is examined in Section 6. Most proofs are relegated to appendices.

2. INFORMAL OUTLINE AND APPLICATIONS

2.1. *Incomplete and Complete State Spaces*

Consider a decision maker operating under uncertainty, where the uncertainty is represented by the state space S, the space of states of the world as in Savage. The objects of choice are acts over S, that is, (suitably measurable) functions from S to the set of outcomes $X = [0, 1]$. Denote by $\mathscr{F}(S)$ this set of acts and by $\mathscr{P}(S)$ the set of "regular" preference orderings over $\mathscr{F}(S)$. More detailed and precise definitions of these sets will be provided later. For now, we proceed informally assuming that the sets $\mathscr{F}(S)$ and $\mathscr{P}(S)$ are well defined for any S. Think of $\mathscr{P}(S)$ as a large class of preferences limited mainly by "technical" conditions and containing subjective expected utility preferences as a small subset. This framework may be applied to describe choice behavior by supposing that the set of feasible choices corresponds to a subset of $\mathscr{F}(S)$ and that a feasible act that maximizes the decision maker's preference, an element of $\mathscr{P}(S)$, is chosen.

Suppose now that the decision maker is a player in a game. The choice of strategy can be modeled as above once the suitable set of states of the world S is specified, by associating each strategy with an act over S. The outstanding question is "what is S?" To proceed, suppose for notational simplicity that there are two players, denoted i and j, with strategy sets $A_i = A_j$. Each player is uncertain about the parameters of the game being played and/or the strategy choice of her opponent. This primitive uncertainty may be represented by some set Ω. If i faced uncertainty *only* about the true $\omega \in \Omega$, then her choice of

[2] See Lipman (1991) for another example of a nondenumerable hierarchical construction.

strategy could be modeled as indicated above: Given j's strategy choice a_j, the perceived payoff of i depends on her strategy a_i and on the true state ω. Therefore, each a_i determines a mapping from Ω into outcomes in $X=[0,1]$, that is, each a_i determines an act in $\mathscr{F}(\Omega)$. Accordingly, i's strategy choice is determined by her preference ordering over $\mathscr{F}(\Omega)$, an element of $\mathscr{P}(\Omega)$. But i does not know j's chosen a_j. By the above for j, she could infer it from knowledge of j's preference ordering in $\mathscr{P}(\Omega)$, but she is uncertain about her opponent's preferences. Thus i's "second-order state space" is $\Omega\times\mathscr{P}(\Omega)$. Were this to represent all the uncertainty facing i, then we could argue as above, identifying each a_i with an act in $\mathscr{F}(\Omega\times\mathscr{P}(\Omega))$ and deriving her strategy choice from her "second-order preferences," an element in $\mathscr{P}(\Omega\times\mathscr{P}(\Omega))$. Similarly for j. But since i's second-order preferences are unknown to j and since they are useful for predicting what i will do, j faces the uncertainty represented by the state space $\Omega\times\mathscr{P}(\Omega)\times\mathscr{P}(\Omega\times\mathscr{P}(\Omega))$. Proceeding, one is led to the sequence of state spaces

$$(1)\qquad S_0^i=\Omega,\quad S_n^i=S_{n-1}^i\times\mathscr{P}(S_{n-1}^j),\qquad n>0,$$

and similarly for j. Each state space S_n^i (or S_n^j) is an *incomplete* description of the uncertainty facing i (or j) since, as above, given that S_n^j describes some of the uncertainty facing j, then i, in predicting j's behavior, faces uncertainty also about j's preferences over $\mathscr{F}(S_n^j)$. Therefore, to model the decision problem facing each player requires first that one prove the existence of a state space that is large enough to incorporate all of the above uncertainty.

Comparison with the Bayesian case may be helpful. When preferences conform to subjective expected utility and when vNM indices of both players are common knowledge, then "beliefs" about any given states space S uniquely determine preferences over $\mathscr{F}(S)$. Therefore, the reasoning outlined above leads to a hierarchy of "beliefs about beliefs." More specifically, since in the Savage model beliefs are represented by a probability measure, the sequence of incomplete state spaces takes the form

$$(2)\qquad S_0^i=\Omega,\quad S_n^i=S_{n-1}^i\times\Delta(S_{n-1}^j),\qquad n>0,$$

where $\Delta(S_{n-1}^j)$ denotes the set of probability measures on S_{n-1}^j. These are the incomplete state spaces considered in Mertens and Zamir (1985) and Brandenburger and Dekel (1993).

Returning to the general non-Bayesian framework and proceeding heuristically, ask what conditions are natural to impose on complete state spaces S^i and S^j, were they to exist. The uncertainty facing i consists of both the primitive uncertainty represented by Ω and, by the reasoning described above, by j's "feelings" about the exhaustive uncertainty that she faces, where such "feelings" are embodied in the way that j ranks acts over her state space S^j. In other words, i is uncertain about $\Omega\times\mathscr{P}(S^j)$. But by hypothesis, S^i represents all the uncertainty facing i. We conclude that S^i and S^j should satisfy

$$(3)\qquad S^i\sim\Omega\times\mathscr{P}(S^j),$$

where $\sim$ indicates a suitable one-to-one correspondence. Anticipating the formal analysis, refer to $\sim$ as a (topological) homeomorphism.

The following reformulation is convenient. Write

$$S^i = \Omega \times T^j \quad \text{and} \quad S^j = \Omega \times T^i,$$

where T^i and T^j are the spaces of possible "types" for i and j and where a "type" is a complete description of the relevant characteristics of a player. Substitute these expressions into (3) to deduce the following restriction on types spaces:

$$(4) \qquad T^i \sim \mathscr{P}(\Omega \times T^j) \quad \text{and} \quad T^j \sim \mathscr{P}(\Omega \times T^i).$$

Conversely, given spaces satisfying these restrictions, then a type $t_i \in T^i$ for i is a comprehensive description of i because it includes, through the homeomorphism $\sim$, a specification of i's "feelings" about both the primitive uncertainty Ω and about the type of her opponent.

The technical contribution of this paper is to describe a mathematically rigorous and natural construction of types spaces, given a very general specification of preferences via $\mathscr{P}$. We claimed in the introduction that the existence of such types spaces was important for some foundational and conceptual issues in game theory and in economic theory more broadly. To buttress this claim, we proceed now to describe three applications.

2.2. *Applications*

2.2.1. *Games of Incomplete Information*

Consider a game that is common knowledge except for the true parameter, an element of Ω. The following Harsanyi (1967/8)-style formulation of such a game accommodates all preferences in the broad class defined by $\mathscr{P}$. The formulation does not contain a counterpart of the "common priors" assumption; there is no reference to ex ante beliefs or preferences. (Some mathematical details required by the later formal analysis have been added.[3])

A *game* is defined as a tuple $(N, \Omega, (A_i, \Theta_i, r_i, U_i)_{i \in N})$, where:

- N is the finite set and number of players;
- Ω is a compact Hausdorff parameter space;
- A_i is a topological space of i's possible actions;
- Θ_i is a compact Hausdorff space of i's possible types (private information);
- r_i: $\prod_{j=1}^N A_j \times \Omega \times \prod_{j=1}^N \Theta_j \to X$ is a measurable function describing outcomes for i;
- U_i: $\Theta_i \to \mathscr{P}(\Omega \times \Theta_{-i}), (\Theta_{-i} \equiv \prod_{j \neq i} \Theta_j)$, where $U_i(\theta_i, \cdot)$ represents i's preferences on acts over $\Omega \times \Theta_{-i}$ conditional on being of type θ_i.

[3] Throughout the paper, every topological space Y is endowed with the Borel σ-algebra and measurability of any function f: $Y \to \mathscr{R}$ refers to Borel measurability. Also, the product topology is used for all Cartesian product spaces.

An *equilibrium* is a tuple $(\sigma_i)_{i \in N}$ where σ_i: $\Theta_i \to A_i$ is measurable and satisfies: $\forall i$ $\forall \theta_i \in \Theta_i$ $\forall a_i \in A_i$,

$$U_i(\theta_i, f_i[\sigma_i(\theta_i), \theta_i]) \geq U_i(\theta_i, f_i[a_i, \theta_i]),$$

where for any $(a_i, \theta_i) \in A_i \times \Theta_i$, $f_i[a_i, \theta_i] \in \mathscr{F}(\Omega \times \Theta_{-i})$ is the act defined by

$$f_i[a_i, \theta_i](\omega, \theta_{-i}) \equiv r_i\Big(a_i, \big(\sigma_j(\theta_j)\big)_{j \neq i}, \omega, \theta_i, \theta_{-i}\Big).$$

Conditional on her type θ_i, player i identifies each strategy a_i with an act over $\Omega \times \Theta_{-i}$ and uses her conditional utility function $U_i(\theta_i, \cdot)$ to evaluate the strategy. This formulation *presumes* that $\Omega \times \Theta_{-i}$ represents all the uncertainty facing i. The analysis in this paper shows how, beginning with Ω, one can construct suitable types spaces $\Theta_i = T^i$ satisfying this condition. (In the case of two players i and j, see Theorem 5.2; $\Theta_i = \Theta_j = T$, and $U_i(\theta_i, \cdot)$ is derived from the homeomorphism (4) or (12), and similarly for j.) The validation provided thereby extends the Mertens-Zamir validation for the standard Bayesian game of incomplete information.

The above game formulation and equilibrium notion have been applied in Lo (1995) to examine the effects of uncertainty aversion in first and second price sealed bid auctions with independent private values. Lo assumes that it is common knowledge that players have preferences in the multiple-priors class (Gilboa and Schmeidler (1989)). Such common knowledge assumptions can be expressed formally as described in Section 6.

2.2.2. *Foundations for Solutions Concepts*

In the decision-theoretic approach to game theory, each player's problem of choosing a strategy is cast as a single agent decision problem under uncertainty. Then, assuming that players are Bayesian rational, alternative assumptions regarding their beliefs about the uncertainty that they face deliver axiomizations for various solution concepts. An example of such an argument, that is the focus here, is the theorem characterizing correlated rationalizability and survival of iterated deletion of strictly dominated strategies as the (equivalent) implications of rationality and common knowledge of rationality (Tan and Werlang (1988, Theorems 5.2–5.3)).

For the reasons given in the introduction, it is desirable to extend such arguments beyond the framework of subjective expected utility maximizing players. The types space constructed in this paper permits such an extension. Here we provide a brief indication of how it may be developed; see Epstein (1995) for a more complete analysis.

Consider a two-player normal form game (A_i, A_j, r_i, r_j), where $A_i = A_j = A$ are compact Hausdorff strategy sets, r_i, r_j: $A_i \times A_j \to [0,1]$ are measurable outcome functions, and the game is common knowledge. To give formal meaning to "i is rational," we proceed as follows: i faces the primitive uncertainty represented by A (j's choice of strategy), but this uncertainty is not exhaustive because "beliefs about beliefs..." matter. The set of *states of the world* facing i

is $A \times T^j$, where T^j is the types space for j, taking $\Omega = A$ in Section 2.1. The strategy choice a_i by i determines the act $r_i(a_i, \cdot)$ over $A \times T^j$ and therefore can be evaluated by i's preference ordering, an element of $\mathscr{P}(A \times T^j)$. Further, i's preference ordering is determined by her type t_i according to the homeomorphism $\sim$ in (4). Let the function Ψ: $T^i \to \mathscr{P}(A \times T^j)$ represent this homeomorphism. This leads to the formal definition that $(a_i, t_i) \in A \times T_i$ is *rational* if for all a_i' in A,

$$(5) \qquad \Psi(t_i)(r_i(a_i, \cdot)) \geq \Psi(t_i)(r_i(a_i', \cdot)).$$

Denote by Q_i the subset of $A \times T$ satisfying (5) and define Q_j similarly. Common knowledge of rationality thus amounts to common knowledge of each of Q_i and Q_j, which can be modeled formally along the lines of Section 6. In particular, such common knowledge corresponds to the players' types lying in suitable subsets T_i^* and T_j^* of T, satisfying[4]

$$T_i^* \sim_{hmeo} \mathscr{P}\big(A \times T | Q_j \cap [A \times T_j^*]\big)$$

and

$$T_j^* \sim_{hmeo} \mathscr{P}(A \times T | Q_i \cap [A \times T_i^*]).$$

Therefore, the implications of rationality and common knowledge of rationality are delivered by answering the following formal question: "For which $(a_i, a_j) \in A^2$ does there exist $(t_i, t_j) \in T_i^* \times T_j^*$ such that $(a_i, t_i) \in Q_i$ and $(a_j, t_j) \in Q_j$?" Epstein (1995) characterizes such strategy profiles in the case of finite games, delivering a notion of rationalizability (and corresponding "dominance") that is not tied to the subjective expected utility framework. Because $\mathscr{P}(A \times T)$ is a large class of preferences, the corresponding notions of rationality and rationalizability are weak. However, they can be strengthened by assuming that a more restrictive model of preference is common knowledge.

Two products of such an analysis merit emphasis. First, new solution concepts for normal form games are provided that have clear and appealing decision-theoretic foundations and that can accommodate empirically attractive (e.g., uncertainty averse) preferences.[5] It is hoped that these solution concepts will prove useful in applications. Another product is the deepening of the foundations for the expected utility-based notion of rationalizability due to the fact that common knowledge of expected utility maximization may be expressed formally in our framework.

2.2.3. *Single Agent Decision Making*

The standard model of information, a partition of the state space, is based on strong assumptions about the decision maker's knowledge and information processing abilities, most notable *negative introspection*—if something is not

[4] The notation $\mathscr{P}(S|A)$, where $A \subset S$, denotes the subset of preferences in $\mathscr{P}(S)$ for which $S \backslash A$ is null in the sense of Savage. See Section 4.3 for a fuller explanation.

[5] Related literature includes Dow and Werlang (1994), Klibanoff (1994), and Lo (1994). They provide equilibrium concepts for normal form games with complete information and uncertainty averse players.

known, it is known that it is not known. Geanakoplos (1989) proposes nonpartitional models of information to capture the decision maker's imperfect understanding of the universe. He then assumes expected utility maximization and Bayesian updating to deliver a model of preference that is used to address issues in applied game theory. This approach of pasting Bayes rule onto a nonpartitional information structure is criticized by Morris (1994, 1996) as being ad hoc because the usual decision theoretical justification for Bayes' rule is based on a form of dynamic consistency that, in turn, implies partition information. It is suggested, therefore, that a unified treatment of preference and (imperfect) knowledge is required and Morris makes some progress in providing such a unification. However, Morris (1994) identifies a circularity in his model and suggests that it is best dealt with by a hierarchical model of preference. The issue is whether there exists a state space rich enough so that states contain a description of (preferences and) knowledge of the state space. Here we show that the hierarchies constructed in this paper, suitably reinterpreted, resolve this issue. There may be some appeal to the view that knowledge is prior to and helps to determine preference. On the other hand, we emphasize, as does Morris, that there is a strong tradition in economics of choice-based, or equivalently, preference-based modeling. Savage's derivation of subjective probability from preference is a prominent example. Another notable recent example is Lipman (1995), where a decision maker's reasoning is derived from preferences.

The reinterpretation just mentioned is that we view the single decision maker as being uncertain (about the state of nature and) about his own preferences rather than about the preferences of another. The discussion leading to (4) is readily modified to relate to the decision maker's introspection; she is uncertain not only about the true state of nature in Ω but also about how she "feels" about this uncertainty and how she "feels about her feelings" regarding the state of nature and so on. At issue is whether there exists a space T large enough to model both introspection to all finite levels and introspection about T itself so that any type $t \in T$ represents the exhaustive introspection of the decision maker. Formally, the issue is the existence of T such that T is homeomorphic to $\mathscr{P}(\Omega \times T)$. Let the function $\Psi: T \to \mathscr{P}(\Omega \times T)$ represent such a homeomorphism.

The following connection can now be made to formal nonpartitional models of knowledge. Let $S = \Omega \times T$ and for each event $A \subset S$, say that the decision maker *knows* A *at the state* $s = (\omega, t)$ if $\Psi(t) \in \mathscr{P}(\Omega \times T | A)$. Define $K(A)$ to be the set of states s at which A is known. Then K is a *knowledge operator* satisfying $K(S) = S$, $K(\varnothing) = \varnothing$, and $K(A) \cap K(B) = K(A \cap B)$. Other arguably more problematic properties of K required by the partition model are not necessarily satisfied. However, they can be imposed as additional assumptions expressed as restrictions on the decision maker's type. For example, the knowledge axiom "know that you know" may be expressed in the form: For every $A \subset \Omega \times T$,

$$\Psi(t) \in \mathscr{P}(\Omega \times T | A) \Rightarrow \Psi(t) \in \mathscr{P}(\Omega \times T | \Omega \times \Psi^{-1} \mathscr{P}(\Omega \times T | A)).$$

Arguments similar to those in Section 6 can be used to construct the subset of types satisfying this condition; similarly for the axiom of negative introspection.

Though the hypothesis of uncertainty about own preference is uncommon in economic theory, it seems natural given an agent who does not perfectly understand the nature of the primitive state space Ω and who reflects on the nature and degree of his misunderstanding. Finally, we note that uncertainty about own preferences has been shown to be useful also in modeling preference for flexibility (Kreps (1979)) and behavior given unforeseen contingencies (Kreps (1992)).

3. THE SPACE OF PREFERENCES

This section provides a formal definition of the space of regular preference orderings.

Begin with a state space S, a compact Hausdorff topological space. Preference will be defined on $\mathscr{F}(S)$, the set of all Borel measurable functions from S to $X = [0, 1]$. In order to describe the class of preferences that will be considered, it is convenient to designate various subsets of $\mathscr{F}(S)$. Call an act *simple* if its range is finite. Call an act f upper semicontinuous (usc) if all sets of the form $\{s: f(s) \geq \kappa\}$ are closed. Similarly, f is lower semicontinuous (lsc) if all sets of the form $\{s: f(s) > \kappa\}$ are open. Denote by $\mathscr{F}^u(S)$ and $\mathscr{F}^l(S)$ the sets of simple usc and simple lsc acts respectively. The outcome $x \in X$ also denotes the corresponding constant act. Finally, for any x and any $\varepsilon > 0$, $(x + \varepsilon)$ should be interpreted as $\min\{x + \varepsilon, 1\}$ and $(x - \varepsilon)$ means $\max\{x - \varepsilon, 0\}$.

By a (regular) *preference order* on $\mathscr{F}(S)$, we mean a reflexive, transitive, and complete binary relation $\succcurlyeq$ satisfying the following conditions for all acts f' and f in $\mathscr{F}(S)$:

P.1 INTERMEDIATE VALUE: $\forall f$ $\exists$ unique $x \in X$ such that $f \sim x$.

P.2 MONOTONICITY: $f' \geq f \Rightarrow f' \succcurlyeq f$.

P.3 INNER REGULARITY: If $f \sim x$, then $\forall \varepsilon > 0$ $\exists g \in \mathscr{F}^u(S)$ satisfying $g \leq f$ and $g \succcurlyeq (x - \varepsilon)$.

P.4 OUTER REGULARITY: $\forall g \in \mathscr{F}^u(S)$, if $g \sim x$, then $\forall \varepsilon > 0$ $\exists h \in \mathscr{F}^l(S)$ satisfying $h \geq g$ and $(x + \varepsilon) \succcurlyeq h$.

While this definition of "preferences" substantially limits the class of binary relations and therefore the generality of our analysis, we feel that the above conditions are attractive on the grounds of being both readily interpretable and not unduly restrictive, as demonstrated by examples below. The first two properties are self-explanatory and common. In particular, the assumption that there exists a unique "certainty equivalent" for each act seems natural in the present setting. The regularity conditions are perhaps unconventional restrictions on preferences, but are nevertheless readily interpretable. Roughly, the first requires that any act can be approximated from below, arbitrarily well in preference, by simple usc acts; the second requires that any g in $\mathscr{F}^u(S)$ can be approximated from above, arbitrarily well in preference, by simple lsc acts. Further discussion of the regularity conditions follows shortly.

Denote by $\mathscr{P}(S)$ the set of preference relations on $\mathscr{F}(S)$. By the intermediate value property, any $\succcurlyeq$ admits a numerical representation $u: \mathscr{F}(S) \to X$, defined by $f \sim u(f)$. That is, $u(f)$ is the *certainty equivalent* of the act f. Further, u inherits the following properties:

U.1 Certainty Equivalence: $u(x) \equiv x$.

U.2 Monotonicity: $f' \geq f \Rightarrow u(f') \geq u(f)$.

U.3 Inner Regularity: $u(f) = \sup\{u(g)\colon g \leq f, g \in \mathscr{F}^u(S)\}, \forall f \in \mathscr{F}(S)$.

U.4 Outer Regularity: $u(g) = \inf\{u(h)\colon h \geq g, h \in \mathscr{F}^l(S)\}, \forall g \in \mathscr{F}^u(S)$.

Equally important and immediate is that any function u satisfying these properties defines a unique preference relation $\succcurlyeq$ (satisfying the properties (P.1)–(P.4)) by

$$f' \succcurlyeq f \Leftrightarrow u(f') \geq u(f).$$

Therefore, we can *identify* $\mathscr{P}(S)$ with the set of functions u satisfying (U.1)–(U.4). Such functions are referred to interchangeably as utility functions, certainty equivalents and, because of the identification, as preference orders. In particular, we often write $u \in \mathscr{P}(S)$. This identification is convenient since the analysis to follow is more simply written in terms of utility functions rather than binary relations. We emphasize, however, that everything that follows can be rewritten explicitly and exclusively in terms of preferences, that is, we are indeed dealing with a space of preference orders rather than nonordinal utility functions.

The regularity conditions (U.3) and (U.4) resemble more familiar restrictions on probability measures (see Section 4.1); think of u as a measure and replace g, h, and f by closed, open, and measurable subsets of S, respectively. (To motivate the latter substitutions, note that the indicator function for a closed set is usc, and so on.) The regularity conditions also imply that a utility function is uniquely determined by the certainty equivalents it assigns to simple lsc acts, much like a regular probability measure is uniquely determined by its values on open sets. This formal similarity may help to motivate our specification of $\mathscr{P}(S)$; in any event, it "explains" how we arrived at it. The noted similarity is also an important reason that a denumerable hierarchy suffices here as it does in the case of hierarchies of probability measures. In particular, the restriction to regular utility functions permits the proof of a "Kolmogorov extension theorem" for utility functions (See Theorem D.2 and Lemma D.3).

The next step is to define a topology on $\mathscr{P}(S)$. We employ the topology τ having the subbasis consisting of sets of the form

$$\{u\colon u(g) < \kappa\}, \qquad \{u\colon u(h) > \kappa\},$$

where κ varies over the reals and g and h vary over $\mathscr{F}^u(S)$ and $\mathscr{F}^l(S)$, respectively. That is, τ is the coarsest topology on $\mathscr{P}(S)$ that makes the mapping

$u \mapsto u(f)$ usc for every $f \in \mathscr{F}^u(S)$ and lsc for every $f \in \mathscr{F}^l(S)$.[6] Expressed explicitly in terms of preference orders $\succcurlyeq$, the above subbasis consists of the sets $\varnothing$ and $\mathscr{P}(S)$ (corresponding to values of κ outside the unit interval), $\{\succcurlyeq : g \prec x\}$ and $\{\succcurlyeq : h \succ x\}$ where x varies over $[0,1]$.

A vital feature, for our purposes, of the topology τ is that it leads to inheritance of the compact Hausdorff property from S to $\mathscr{P}(S)$.

THEOREM 3.1: *$(\mathscr{P}(S), \tau)$ is compact Hausdorff.*

One justification for the choice of τ is pragmatic, that is, "it works," as shown below. We view the compact Hausdorff property of $(\mathscr{P}(S), \tau)$ as another justification, or at least as a confirmation that the topology is "reasonable," that is because compactness ensures that τ does not contain "too many" open sets while the Hausdorff property ensures that it does not contain "too few" open sets. More precisely, note that for any other topology τ', τ' *must* violate compactness if it is strictly stronger than τ and it *must* violate Hausdorff if it is strictly weaker than τ (see Royden (1988, p. 192)).

Additional perspective on the topology τ may be provided by considering some specific alternatives. Topologies frequently adopted for spaces of preferences include the topology of closed convergence (Hildenbrand (1974) and Grodal (1974)) and the Kannai topology (Kannai (1970)). These presume that the domain of preference is itself a topological space and that all preferences are continuous. Further, local compactness of the domain is needed in order that the Kannai topology be well-defined and that the closed convergence topology be Hausdorff. There remains also the need, for the purpose of the construction of hierarchies, to ensure that the topological space of preferences inherits suitable properties from the domain. For example, a seemingly useful result is that the property "compact separable metric" is passed on under the closed convergence topology (Hildenbrand (1974, pp. 19–20)). However, we are not aware of any reasonable topologies on the domain $\mathscr{F}(S)$ that deliver the required properties. For example, $\mathscr{F}(S)$ is in general not locally compact under the sup-norm topology and it fails to be compact separable metric under the product topology.

A very special but still useful illustration has a finite state space $S = \{1, \ldots, n\}$, in which case $\mathscr{F}(S) = \mathscr{F}^u(S) = \mathscr{F}^l(S) = [0,1]^n$. Therefore, the regularity conditions are trivially satisfied and $\mathscr{P}(S)$ can be identified with that subset of $[0,1]^{[0,1]^n}$ consistent with (U.1) and (U.2). The topology τ is the induced product topology. It is straightforward to construct a utility function satisfying both (U.1) and (U.2) that is not continuous on $[0,1]^n$ and that fails to be strictly monotonic there, where the latter refers to the property

$$f' > f \text{ everywhere on } S \Rightarrow u(f') > u(f).$$

[6] It can be shown that τ is weaker than the topology generated by sets of the form $\{u: u(f) < \kappa\}$ and $\{u: u(f) > \kappa\}$, where f varies over all *continuous* functions in $\mathscr{F}(S)$. Note that these sets are in general not open in τ because the subbasis in (6) employs simple functions.

Therefore, the example demonstrates that in general regular utility functions need not be sup-norm-continuous or strictly monotonic. In this specific example, $\mathscr{F}(S)$ is in fact compact separable metric under the usual topology on $[0,1]^n$ and the above cited theorem of Hildenbrand delivers a compact separable metric space of (continuous) preferences. However, the finiteness of the state space is critical and finiteness is lost at the second stage of the hierarchy, that is, once one's opponents preferences themselves constitute part of the uncertainty facing a player.

4. PREFERENCE SUBSPACES

We turn to some noteworthy subspaces of $\mathscr{P}(S)$ that will help to clarify the nature of $(\mathscr{P}(S), \tau)$. In particular, we hope that the specializations to follow will convince the reader that our choice of regularity conditions and topology are sensible at least in the sense that they are consistent with more familiar models.

4.1. *Expected Utility*

Denote by $\Delta(S)$ the space of regular Borel probability measures endowed with the weak convergence topology, where regularity for $p \in \Delta(S)$ means:[7]

R.1 $p(A) = \sup\{p(K): K \subset A \text{ compact}\}$ $\forall$ measurable A.

R.2 $p(K) = \inf\{p(G): G \supset K \text{ open}\}$, $\forall$ compact K.

Note that $\Delta(S)$ is compact Hausdorff; this is implied by the Riesz Representation and Alaoglu Theorems (Royden (1988, pp. 352, 237)).

Fix a vNM index v: $X \to \mathscr{R}$, continuous and strictly increasing. For each measure $p \in \Delta(S)$, define u_p by

$$u_p(f) \equiv v^{-1}\left(\int v(f)\, dp\right). \tag{7}$$

The identification of p with u_p establishes a homeomorphism between $\Delta(S)$ and a subspace of $\mathscr{P}(S)$.

THEOREM 4.1: *The mapping $p \to u_p$ is a homeomorphism between $\Delta(S)$ and a compact subspace of $\mathscr{P}(S)$.*

This is a simple corollary of Theorem 4.2.

A possible extension of the theorem is worth noting. Note that it is the probabilistic sophistication of each u_p that is important above, not the expected utility functional form itself. In particular, the class of all u's corresponding to some "belief" $p \in \Delta(S)$ and a fixed "well-behaved" functional for evaluating

[7] In fact, for S compact Hausdorff, either of these conditions implies the other (Royden (1988, p. 341)).

risky prospects is also compact and homeomorphic to $\Delta(S)$. The reader is referred to Machina and Schmeidler (1992) for the details supporting this sketched extension.

4.2. *Choquet Utility*

Schmeidler (1989) has axiomatized a generalization of expected utility theory that can accommodate aversion to uncertainty such as exhibited in the Ellsberg paradox. We show that a suitable specialization of his model can be embedded within the space of regular utility functions $\mathscr{P}(S)$.

By a (regular) *capacity* we mean a function c from the measurable subsets of S into $[0,1]$ satisfying $c(\varnothing)=0$, $c(S)=1$, monotonicity with respect to set inclusion and the counterparts of the regularity conditions R.1–R.2.[8] Note that since capacities are not necessarily additive, the conjunction of the two regularity conditions is strictly stronger than either one alone.

Each capacity can be associated with a utility function as follows: Fix a vNM index v as in the preceding example. For each capacity c, define the utility function u_c by

$$u_c(f) \equiv v^{-1}\left(\int v(f)\,dc\right), \tag{8}$$

where the indicated integration is in the sense of Choquet. (When the integrand is nonnegative, the integral is defined to equal the Riemann integral $\int_0^\infty c\{v(f)\geq t\}\,dt$. In particular, for any characteristic function $f=1_A$, the Choquet integral equals $v(0)+c(A)[v(1)-v(0)]$. Though Choquet integration is not additive, it does satisfy the following limited form of additivity:

$$\int [v(f)+v(f')]\,dc = \int v(f)\,dc + \int v(f')\,dc, \tag{9}$$

whenever f and f' are comonotonic, that is, $[f(s)-f(s')][f'(s)-f'(s')]\geq 0$ $\forall s,s'\in S$.)

Denote by $\mathscr{C}(S)$ the set of capacities and endow it with the *vague topology*, which is the topology generated by the subbasis

$$\{c\colon c(K)<\kappa\} \quad \{c\colon c(G)>\kappa\},$$

where K and G vary over compact and open subsets of S and κ varies over the reals. Equivalently, the vague topology is the weakest topology on $\mathscr{C}(S)$ that renders the mapping $c\mapsto \int f\,dc$ continuous for each continuous act f (see O'Brien and Vervaat (1991)). It follows that the topology induced on $\Delta(S)$,

[8] This definition, apart from the normalization $c(S)=1$, coincides with that adopted in Norberg (1986) and O'Brien and Vervaat (1991). Schmeidler (1989) does not impose any regularity conditions, while alternative regularity conditions are occasionally imposed, e.g., in Choquet (1953/4) and Graf (1980).

viewed as a subset of $\mathscr{C}(S)$, coincides with the weak convergence topology. Finally, $\mathscr{C}(S)$ is compact Hausdorff (O'Brien and Vervaat (1991, Theorems 2.2, 2.3)).

We have the following extension of Theorem 4.1:

THEOREM 4.2: *The mapping* I: $c \mapsto u_c$ *is a homeomorphism between* $\mathscr{C}(S)$ *and a compact subspace of* $\mathscr{P}(S)$.

The subset of convex capacities warrants separate mention. A capacity is called *convex* if for all measurable sets A and B

$$c(A \cup B) + c(A \cap B) \geq c(A) + c(B).$$

Schmeidler (1989) has demonstrated that convexity of the capacity c corresponds to a form of uncertainty aversion for u_c and thus defines an interesting specialization of (8). The set of all convex capacities is closed in $\mathscr{C}(S)$.[9] By Theorem 4.2, it follows that the set of all Choquet expected utility functions with convex capacity and a fixed vNM index is closed in $\mathscr{P}(S)$.

4.3. *Knowledge Subspaces*

It will be of interest to consider the subset of $\mathscr{P}(S)$ that corresponds to "knowledge that $A \subset S$ is true." Given u, say that u *knows* the Borel measurable event A if $S \setminus A$ is null in the sense of Savage, that is, if for all f and f' in $\mathscr{F}(S)$,

$$(10) \qquad f = f' \quad \text{on } A \Rightarrow u(f) = u(f').$$

It will be useful to extend this definition also to nonmeasurable subsets of S. Therefore, if B is an arbitrary subset of S, say that u *knows* B if there exists a measurable event $A \subset B$ such that u knows A.[10] Finally, adopt the notation $\mathscr{P}(S|B) \equiv \{u \in \mathscr{P}(S)\colon u \text{ knows } B\} \subset \mathscr{P}(S)$.

We can prove the following intuitive and useful result:

THEOREM 4.3: *If* $A \subset S$ *is compact, then* $\mathscr{P}(A)$ *is homeomorphic to the compact set* $\mathscr{P}(S|A)$.

Given a utility function u, the following property of the associated knowledge will be important below: "If $\{A_k\}$ is a declining sequence of subsets of S such that u knows each A_k, then u knows also $\bigcap A_k$." Refer to any utility function satisfying this property as exhibiting *continuous knowledge*. Expected utility functions satisfy this property because of the countable additivity of probability

[9] This follows readily from the characterization of regular convex capacities in Anger (1971, Theorem 3).

[10] This extension parallels the common procedure of completing a σ-algebra by adding to it all subsets of measurable sets that are null with respect to a given probability measure.

measures, more particularly because a countable union of null events is null. We have not determined whether *every* utility function in $\mathscr{P}(S)$ exhibits continuous knowledge, but we can show that *many* do. To be precise, all utility functions that are continuous with respect to the sup-norm topology on acts satisfy this property. This class includes (Choquet) expected utilities.

THEOREM 4.4: *If $u \in \mathscr{P}(S)$ is sup-norm continuous, then it exhibits continuous knowledge.*

5. A TYPES SPACE

Refer to the two-player game situation described informally in Section 2.1. The incomplete state spaces in (1) are well defined, because S^i_{n-1} and S^j_{n-1} compact Hausdorff imply the same for S^i_n and S^j_n. The latter two spaces are equal and can be denoted S_n. This leads us to the sequence of incomplete state spaces S_n defined by[11]

(11) $$S_0 = \Omega, \quad S_n = S_{n-1} \times \mathscr{P}(S_{n-1}), \quad n > 0.$$

Similarly, we can write $T = T^i = T^j$, where the types space T should satisfy

(12) $$T \sim_{hmeo} \mathscr{P}(\Omega \times T).$$

We proceed to construct such a types space T. This is done by adapting and extending the Brandenburger-Dekel (1993) argument from their Bayesian framework to the present framework.

It is natural to represent player i by her "feelings about Ω," her "feelings about j's feelings about Ω" and so on to all finite orders, that is, since "feelings" are expressed by preferences, by an element $t = (u_0, \dots, u_n, \dots)$ in $T_0 \equiv \prod_0^\infty \mathscr{P}(S_n)$. Not all such elements t are sensible, however, since the utility functions u_n may contradict one another. To clarify and rule out contradictions, identify $\mathscr{F}(S_{n-1})$ with a subset of $\mathscr{F}(S_{n-1} \times \mathscr{P}(S_{n-1}))$ by identifying each act over S_{n-1} with an act over the larger state space $S_{n-1} \times \mathscr{P}(S_{n-1})$ that does not depend on the second argument. Then any $u_n \in \mathscr{P}(S_n) = \mathscr{P}(S_{n-1} \times \mathscr{P}(S_{n-1}))$ induces a "marginal" preference order $\text{mrg}_{\mathscr{F}(S_{n-1})} u_n$ on acts over S_{n-1}. Say that the type t is *coherent* if $\forall n \geq 2$

(13) $$\text{mrg}_{\mathscr{F}(S_{n-1})} u_n = u_{n-1},$$

or more simply written, if

$$u_n(f) = u_{n-1}(f) \qquad \forall f \in \mathscr{F}(S_{n-1}).$$

The subspace of T_0 consisting of coherent types is denoted T_1.

[11] If differential information about Ω is modeled via partitions (or σ-algebras), i will generally be uncertain not only about j's preferences on Borel-measurable acts over Ω, but also about j's partition. Such uncertainty is modeled in Heifetz and Samet (1993), but is ignored in a recursion such as (11).

Each coherent type determines a unique preference ordering on acts over the uncertainty space $\Omega \times T_0$, as shown by the following preliminary result:

THEOREM 5.1: *T_1 is homeomorphic to $\mathscr{P}(\Omega \times T_0)$.*

REMARK: The proof, given in Appendix D, is based on a "Kolmogorov extension theorem" for preferences that generalizes the well known theorem for probability measures (see Bochner (1960, Chapter 5) or Rao (1984, p. 165), for example).

It is natural to assume not only that each player is coherent, but also that this is common knowledge. To express this restriction, let $\Psi: T_1 \to \mathscr{P}(\Omega \times T_0)$ denote the homeomorphism in Theorem 5.1 and define for $k \geq 2$,

(14) $$T_k = \{t \in T_1 : \Psi(t) \text{ knows } \Omega \times T_{k-1}\} = \Psi^{-1}\mathscr{P}(\Omega \times T_0 | \Omega \times T_{k-1}).$$

Then $T_k \times T_k$ equals the subset of $T_0 \times T_0$ for which all of the following statements are true: i and j are each coherent, each knows (employing the homeomorphism Ψ) that the other is coherent, each knows that the other knows this, and so on up to order k.[12] Thus if we define

(15) $$T = \cap_k T_k,$$

then $T \times T$ is the set of types that impose common knowledge of coherence.

Our principal result is that the space T constructed in this way satisfies (12). That is, in light of the discussion in Section 2.1, "feelings about feelings..." to all finite orders, supplemented by coherence and common knowledge of coherence, provides an exhaustive characterization of each player. Correspondingly, the Cartesian product $\Omega \times T$ is a space of states of the world that provides a complete description of the uncertainty facing either player.

THEOREM 5.2: *The space T defined by* (15) *satisfies both* (12) *and*

(16) $$T \sim_{hmeo} \mathscr{P}(\Omega \times T_0 | \Omega \times T).$$

Furthermore T is compact Hausdorff and nonempty.

PROOF: For (16), it is enough to prove that

(17) $$T \subset \Psi^{-1}\mathscr{P}(\Omega \times T_0 | \Omega \times T).$$

That is because the reverse inclusion is trivially true. The above inclusion is equivalent to

(18) $$\cap_{k \geq 0}\mathscr{P}(\Omega \times T_0 | \Omega \times T_k) \subset \mathscr{P}(\Omega \times T_0 | \Omega \times \cap_{k \geq 0} T_k).$$

[12] Note that the sets T_k are well defined since the indicated knowledge subspaces do not presume any measurability; see subsection 4.3.

Note that (18) admits the intuitive interpretation that knowledge of each $\Omega \times T_k$ should imply knowledge of the "limit" $\Omega \times \bigcap T_k$.[13] We verify that T satisfies (18).

Observe that T_0 and $T_1 \subset T_0$ are compact Hausdorff by Theorems 3.1 and 5.1. Apply Theorem 4.3 repeatedly to conclude that each $T_k \subset T_0$ is compact. Now let u lie in the intersection on the left side of (18). Then u knows $\Omega \times T_k$ for all k. Thus for any simple usc function g, $u(g1_{T_k}) = u(g)$. By Lemma D.1 and the compact Hausdorff nature of each T_k, $u(g1_{T_k}) \searrow u(g1_T)$. Therefore, $u(g1_T) = u(g)$. It follows from inner regularity that u knows $\Omega \times T$.

This proves that $T \sim_{hmeo} \mathscr{P}(\Omega \times T_0 | \Omega \times T)$. By Theorem 4.3, the latter is homeomorphic to $\mathscr{P}(\Omega \times T)$. Each T_k is nonempty since it is homeomorphic (by Theorem 4.3) to $\mathscr{P}(\Omega \times T_{k-1})$. Therefore, T is nonempty by the nature of compact Hausdorff spaces. *Q.E.D.*

REMARK: An example of an element of T is $(u_0, \ldots, u_n, \ldots) \in \prod_0^\infty \mathscr{P}(S_n)$, defined recursively by $s_0 = \bar{\omega}$, $u_0(f) = f(\bar{\omega})$, and $\forall n \geq 1$,

$$s_n = (s_{n-1}, u_{n-1}) \quad \text{and} \quad u_n(f) = f(s_n), \qquad f \in \mathscr{F}(S_n).$$

Here $\bar{\omega}$ is a fixed element of Ω. This type for i indicates that i knows $\{\bar{\omega}\}$, i knows that j knows $\{\bar{\psi}\}$, i knows that j knows $\{\bar{\omega}\}$, and so on. (It follows from the compact Hausdorff nature of Ω that each u_n is regular and therefore lies in $\mathscr{P}(S_n)$.)

6. COMMON KNOWLEDGE

We conclude the paper by showing that the types space we have constructed is rich enough to permit the formal modeling of various common knowledge assumptions. We indicated above that $T \times T$ equals that subset of $T_0 \times T_0$ for which coherence is common knowledge. In this section, we construct subsets of $T \times T$ that impose in addition common knowledge of events (subsets of Ω) and various models of preference.

Two special cases are particularly noteworthy. The first imposes as common knowledge that everyone is an expected utility maximizer (with fixed vNM index). Then it is an immediate consequence of Theorem 4.1, that the subspace of types we construct below is homeomorphic to the hierarchy of beliefs analyzed in Mertens and Zamir (1985), Heifetz (1993), and Brandenburger and Dekel (1993). This establishes the sense in which our analysis of hierarchies of preferences extends these earlier studies. The second special case imposes as common knowledge that everyone is a Choquet expected utility maximizer (with

[13] This condition resembles the property of knowledge structures termed "limit closure" in Fagin, Geanakoplos, Halpern, and Vardi (1992), where it is identified as a critical reason for the adequacy of denumerable hierarchies in the measure-theoretic approach to modeling knowledge. The counterpart of (18) for probability measures is automatically satisfied because of countable additivity, regardless of the nature of Ω.

fixed vNM index). Then Theorem 4.2 implies that the subspace of types constructed below is homeomorphic to a hierarchy of capacities. Such hierarchies have not previously been studied to our knowledge.

To proceed, represent a *model of preference* by a subset $P^*(\Omega \times T)$ of $\mathscr{P}(\Omega \times T)$ satisfying conditions to be specified. Fix also a subset Ω_0 to Ω. We wish to identify the subset of types in T that correspond to the situation where everyone knows Ω_0, everyone's preferences conform with the model P^* and both facts are common knowledge. It is natural to mimic the modeling of common knowledge of coherence, via the construction in (14), and to proceed as follows: First, introduce the notation

$$P^*(\Omega \times T | A) \equiv \mathscr{P}(\Omega \times T | A) \cap P^*(\Omega \times T),$$

to denote the class of preferences in $P^*(\Omega \times T)$ that know A, where A is a (not necessarily measurable) subset of $\Omega \times T$. Then define the sequence of subspaces of T by $T^{(0)} = T$,

$$(19) \qquad T^{(k+1)} = \Psi^{-1} P^*(\Omega \times T | \Omega_0 \times T^{(k)}), \qquad k \geq 0.$$

Note that t_j in $T^{(1)}$ indicates that j knows Ω_0 and conforms to model P^*; t_i in $T^{(2)}$ indicates that the preceding applies to i and that i knows that it applies to j; and so on. Therefore, $T^* = \bigcap T^{(k)}$ is the natural candidate subspace of T.

Reasoning similar to that surrounding (16)–(18) in the modeling of common knowledge of coherence leads to the following condition:

$$(20) \qquad \bigcap_{k \geq 0} P^*(\Omega \times T | \Omega_0 \times T^{(k)}) \subset P^*(\Omega \times T | \Omega_0 \times \bigcap_{k \geq 0} T^{(k)}).$$

This condition represents a restriction on Ω_0 and $P^*(\Omega \times T)$ that is *necessary and sufficient* in order that there exist a subspace T^* of T that embodies common knowledge of Ω_0 and $P^*(\Omega \times T)$.[14] We are still left with the question "is this condition satisfied in a broad class of cases outside the expected utility framework?" The next theorem provides an affirmative answer by showing that it is satisfied under two alternative sets of assumptions.

THEOREM 6.1: *Let $\Omega_0 \subset \Omega$ and $P^*(\Omega \times T) \subset \mathscr{P}(\Omega \times T)$. Assume either (a) or (b), where:*

(a) Ω_0 and $P^*(\Omega \times T)$ are closed.

(b) Every $u \in P^*(\Omega \times T)$ exhibits continuous knowledge (defined in Section 4.3).

Then $T^* \subset T$ defined above satisfies

$$T^* \sim_{hmeo} P^*(\Omega \times T | \Omega_0 \times T^*),$$

[14] It is possible that some $T^{(k)}$ is empty and therefore that T^* is also empty. This is not a difficulty in modeling common knowledge; it merely indicates that kth order knowledge of both the event Ω_0 and the model of preference P^* are mutually inconsistent and therefore naturally the same applies to common knowledge of both. Since $P^*(\Omega \times T | \varnothing) = \varnothing$, all formulae retain meaning in such cases. In any event, below we describe conditions that guarantee nonemptiness of T^*.

where the indicated homeomorphism is the restriction of Ψ. Further, T^* is a nonempty closed subspace of T if (a) holds and if each $T^{(k)}$ is nonempty.

The proof of sufficiency of (a) is similar to that of Theorem 5.2. The proof of sufficiency of (b) follows immediately from (20) and the definition of continuous knowledge.

REMARK: Under suitable strengthening of (b), we can show that T^* is nonempty. Assume that rather than being restricted to the particular state space $\Omega \times T$, the model P^* is a correspondence that assigns to any compact Hausdorff state space S a nonempty subspace $P^*(S)$ of $\mathscr{P}(S)$. Such a formulation of "model of preference" seems natural if one is thinking of axiomatically based models. Assume further that for each S and $s \in S$, the evaluation utility function $u_s \in P^*(S)$, where $u_s(f) \equiv f(s)$ $\forall f \in \mathscr{F}(S)$. Nonemptiness of T^* can then be proven by constructing an example along the lines of that concluding Section 5.

Consider assumptions (a) and (b) in turn. With regard to the former, the assumption that $P^*(\Omega \times T)$ is closed is restrictive; for example, it excludes the class of all strictly increasing utility functions in $\mathscr{P}(\Omega \times T)$. On the other hand, this assumption is satisfied by the expected utility and Choquet expected utility model defined in Theorems 4.1–4.2.

While (a) restricts Ω_0 to be compact, the alternative sufficient condition (b) in the theorem places no restrictions on the event Ω_0, other than what is implied by its being a subset of compact Hausdorff space Ω. By Dugundji (1966, XI.8), the implicit restriction on Ω_0, viewed as a topological space with the induced topology, is that it be completely regular. In other words, (b) shows that we can model common knowledge of the combination of any completely regular event Ω_0 and any model of preference P^* that satisfies continuous knowledge. As we saw in Section 4.3, any sup-norm continuous utility satisfies knowledge continuity and so (b) has wide applicability. Finally, through (b) we can model common knowledge of strictly monotone and knowledge continuous preferences, though we noted earlier that common knowledge of strict monotonicity alone is beyond the scope of (a).

The significance of Theorem 6.1 merits emphasis. In standard Bayesian models of differential information, it is often assumed that the "structure" of the model is common knowledge. In particular, common knowledge of expected utility preferences is assumed. But such common knowledge is not well defined formally and therefore must be understood informally (Aumann (1987)). However, within the wider framework of the class $\mathscr{P}$ of preferences, where non-expected utility preferences can be imagined by each player, (common) knowledge of expected utility preferences becomes meaningful. In particular, the preceding Theorem (with $\Omega_0 = \Omega$) shows that common knowledge of expected utility and Choquet expected utility preferences have formal meaning in our framework because such models of preference correspond to appropriate subsets of T.[15] On the other hand, knowledge of $\mathscr{P}$ remains a meta-assumption in our framework. (See Gilboa (1988) and Heifetz (1994) for approaches to formalizing meta-assumptions.)

Department of Economics, University of Toronto, Toronto, Ontario, Canada, M5S 1A1

and

Department of Economics, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1

Manuscript received June, 1994; final revision received January, 1996.

APPENDIX A: A "BETWEEN" LEMMA

The lemma presented here is used in the proofs of Theorems 3.1, 4.2, and Lemma D.3. In the special case where g and h are indicator functions for closed and open sets respectively, the Lemma is an implication of Urysohn's Lemma and the fact that any compact Hausdorff space S is normal.

LEMMA A.1: *Let S be compact Hausdorff. For each pair of simple usc and lsc functions g and h with $g \leq h$, there exists a pair of functions of the form $\sum_{i=1}^k \alpha_i 1_{G_i}$, and $\sum_{i=1}^k \alpha_i 1_{\overline{G}_i}$, where G_i are open sets, such that $g \leq \sum_{i=1}^k \alpha_i 1_{G_i} \leq \sum_{i=1}^k \alpha_i 1_{\overline{G}_i} \leq h$. Moreover, given any basis for the topology of S, each set G_i can be chosen to be a finite union of basic open sets.*

PROOF: Let g and h be such a pair of simple functions, expressed as in (25). Fix i and consider the collection of all open sets G such that $K_i \subset G$. Order this collection by inverse inclusion. Then it comprises a net $\{G_{i\gamma_i}\}_{\gamma_i \in J_i}$, where J_i is the directed indexed set. Note that in general the directed set J_i depends on i. However, by defining the new directed set $J = \{(\gamma_1, \ldots, \gamma_k): \gamma_1 \in J_1, \ldots, \gamma_k \in J_k\}$ with the partial order defined by

$$(\gamma_1, \ldots, \gamma_k) \geq (\gamma_1', \ldots, \gamma_k') \Leftrightarrow \gamma_i \geq \gamma_i' \qquad \forall i,$$

and defining

$$G_{i(\gamma_1, \ldots, \gamma_k)} = G_{i\gamma_i},$$

we obtain nets $\{G_{i\gamma}\}$ with a common directed index set J.

These nets $\{G_{i\gamma}\}_{\gamma \in J}$ have the following property. If $\{G_{i\gamma}\}_{\gamma \in J'}$ is a subnet of $\{G_{i\gamma}\}_{\gamma \in J}$, then

$$\bigcap \{G_{i\gamma}: \gamma \in J'\} = \bigcap \{\overline{G}_{i\gamma}: \gamma \in J'\} = K_i. \tag{21}$$

To see this, let x be any point that is not in K_i. Since S is compact Hausdorff, there exists an open set G such that

$$K_i \subset G \quad \text{and} \quad x \notin G.$$

[15] Common knowledge of Choquet expected utility (with fixed vNM index) is modeled via a set of types T^* satisfying the appropriate form of the homeomorphism in Theorem 6.1, which is $T^* \sim_{hmeo} \mathscr{C}(\Omega \times T^*)$. (Recall that $\mathscr{C}$ denotes the space of regular capacities.) The set T^* satisfying this condition could alternatively have been constructed "directly" by the construction in Section 5, modified so that, roughly speaking, $\mathscr{P}(\cdot)$ is replaced throughout by $\mathscr{C}(\cdot)$ and acts are replaced by events. An advantage of the indirect route we have described is that it permits common knowledge of Choquet expected utility to be expressed formally.

Then there is an index $\bar{\gamma} \in J$ such that $G = G_{i\bar{\gamma}}$. By the definition of "subnet," $\exists \gamma' \in J'$ such that $\gamma' \geq \bar{\gamma}$, which implies that

$$G_{i\gamma'} \subset G_{i\bar{\gamma}}.$$

Thus $x \notin G_{i\gamma'}$ and hence $x \notin \cap\{G_{i\gamma}: \gamma \in J'\}$, proving part of (21). For the rest, let $G_{i\gamma'}$ be an element of the subnet. Since S is compact Hausdorff, there exists an open set V such that $K_i \subset V \subset \bar{V} \subset G_{i\gamma'}$. The set V must be an element of the original net, that is, $V = G_{i\gamma}$ for some γ. Again by the definition of "subnet," $\exists \gamma'' \in J'$, such that $\gamma'' \geq \gamma$, which implies that

$$K_i \subset G_{i\gamma''} \subset V \subset \bar{V} \subset G_{i\gamma'} \Rightarrow K_i \subset \bar{G}_{i\gamma''} \subset \bar{V} \subset G_{i\gamma'}.$$

Therefore,

$$K_i \subset \bigcap_{J'} \bar{G}_{i\gamma} \subset \bigcap_{J'} G_{i\gamma} = K_i,$$

completing the proof of (21).

The key to the proof is the claim that for some γ_0, and hence for all $\gamma \geq \gamma_0$,

$$g \leq \sum_{i=1}^{k} \alpha_i 1_{G_{i\gamma}} \leq \sum_{i=1}^{k} \alpha_i 1_{\bar{G}_{i\gamma}} \leq h. \tag{22}$$

Suppose, to the contrary, that for each γ $\exists x_\gamma$,

$$a_\gamma \equiv \sum_{i=1}^{k} \alpha_i 1_{\bar{G}_{i\gamma}}(x_\gamma) > \sum_{i=1}^{k} \beta_i 1_{U_i}(x_\gamma). \tag{23}$$

Since the left-hand side of the above equality can assume at most finitely many values, there is a subnet $\{a_\gamma: \gamma \in J'\}$, such that $a_\gamma = a$, a constant value, for $\gamma > \gamma_0$. (To confirm that such a subset exists, denote by $a_1, \ldots, a_m$, the possible values of the a_γ's and $A_j = \{\gamma \in J: a_\gamma = a_j\}$. Suppose that for each j $\exists \gamma_j \in J$, such that $\gamma \notin A_j$ $\forall \gamma \geq \gamma_j$; otherwise, the desired subnet exists. Since J is a directed set, $\exists \bar{\gamma} \in J, \bar{\gamma} \geq \gamma_j$ $\forall j$. Therefore, $\gamma \notin A_j$ $\forall j$ $\forall \gamma \geq \bar{\gamma}$, contradicting $J = \cup_j A_j$.) Similarly, there is a further subnet,[16] and a fixed set of indices $\{i_1, \ldots, i_m\} \subset \{1, 2, \ldots, k\}$, say $\{1, 2\}$ such that $a = \alpha_1 + \alpha_2$ and $x_\gamma \in \bar{G}_{1\gamma} \cap \bar{G}_{2\gamma}$ for $\gamma > \gamma_0$. By compactness, a subnet of x_γ, say itself, converges to a point x. Then $x \in K_1 \cap K_2$. (To show this, it is sufficient to show that $x \in \bar{G}_{1\gamma} \cap \bar{G}_{22\gamma}$ $\forall \gamma \geq \gamma_0$. Suppose to the contrary that for some $\bar{\gamma} > \gamma_0$,

$$x \notin \bar{G}_{1\bar{\gamma}} \cap \bar{G}_{2\bar{\gamma}}.$$

Then

$$x \in \bar{G}^c_{1\bar{\gamma}} \cap \bar{G}^c_{2\bar{\gamma}}.$$

Since $x_\gamma \to x$, $\exists \gamma' > \bar{\gamma}$ such that if $\gamma > \gamma'$, then

$$x_\gamma \in \bar{G}^c_{1\gamma} \cup \bar{G}^c_{2\bar{\gamma}}.$$

On the other hand, $\gamma > \bar{\gamma}$ implies that

$$x_\gamma \in \bar{G}_{1\gamma} \cap \bar{G}_{2\gamma} \subset \bar{G}_{1\bar{\gamma}} \cap \bar{G}_{2\bar{\gamma}},$$

a contradiction.) Thus $g(x) \geq \alpha_1 + \alpha_2$. For $\gamma > \gamma_0$, the right-hand side of (23) also assumes finitely many values and the strict inequality holds for all of these finitely many values. By considering a

[16] For notational simplicity, we use the same notation to denote the original net and all subnets.

subnet if necessary, we may assume that the right-hand side assumes a constant value. Thus we have

$$g(x) \geq \alpha_1 + \alpha_2 = \sum_{i=1}^{k} \alpha_i 1_{\overline{G}_{i\gamma}}(x_\gamma) > \sum_{i=1}^{k} \beta_i 1_{U_i}(x_\gamma) = \liminf \sum_{i=1}^{k} \beta_i 1_{U_i}(x_\gamma) \geq h(x).$$

This is a contradiction to $h \geq g$ and establishes the claim (22).

Turn to the final assertion regarding the nature of the sets G_i. By construction, $K_i \subset G_i$ for each i. Since K_i is compact, $\exists K_i \subset G_i' \subset G_i$ such that G_i' is a finite union of basic open sets. These sets $\{G_i'\}$ satisfy the required conditions. *Q.E.D.*

APPENDIX B: PROOFS FOR SECTION 3

PROOF OF THEOREM 3.1: *Part* (a): Let $\hat{\mathscr{P}}(S)$ be the set of utility functions u satisfying (U.2)–(U.4) and taking on values in $[0,1]$, endowed with the topology τ redefined in the obvious way. Note that $\mathscr{P}(S)$ is a closed subspace of $\hat{\mathscr{P}}(S)$; this is true since constant acts are simple, usc and lsc, implying that $\{u \in \hat{\mathscr{P}}(S)\colon u(x) \equiv x\}$ is closed. It remains to show that $\hat{\mathscr{P}}(S)$ is compact. This is done by adapting the proof of O'Brien and Vervaat (1991, Theorem 2.2), which is a corresponding result for the space of capacities. (Note that the argument to follow is valid for any topological space S, not only for compact Hausdorff spaces.)

By Alexander's subbase theorem (Kelley (1985, p. 139)), it suffices to find a finite subcover for any cover of the form

$$\hat{\mathscr{P}}(S) \subset \bigcup_{i \in I} \{u\colon u(h_i) > x_i\} \cup \bigcup_{j \in J} \{u\colon u(g_j) < y_j\}, \tag{24}$$

where g and h are simple usc and lsc functions respectively, x_i and y_j are real numbers, and I and J are index sets. We may assume that $x_i, y_j \in X$. (If $x_i > 1$ or if $y_j < 0$, the corresponding set is empty; if for some i, $x_i < 0$, or for some j, $y_j > 1$, then the open set $\{u\colon u(h_i) > x_i\}$ or $\{u\colon u(g_j) < y_j\}$ is already a finite covering.) Define a utility function u_0 as follows. For any simple usc act g, set

$$u_0(g) = \inf\{x_i\colon h_i \geq g\} \text{ (here } \inf\{\varnothing\} = 1).$$

For a general act f, let

$$u_0(f) = \sup\{u_0(g)\colon f \geq g,\, g \in \mathscr{F}^u(S)\}.$$

Then u_0 is monotonic and satisfies inner regularity. By construction, $u_0(h_i) \leq x_i$. Therefore, $\forall g \in \mathscr{F}^u(S), u_0(g) = \inf\{x_i\colon h_i \geq g\} \geq \inf\{u_0(h_i)\colon h_i \geq g\} \geq \inf\{u_0(h)\colon h \geq g\} \geq u_0(g)$. That is, u_0 satisfies outer regularity and $u_0 \in \hat{\mathscr{P}}(S)$.

By (24) and $u_0(h_i) \leq x_i$ $\forall i, \exists j_0 \in J$ such that $u_0(g_{j_0}) < y_{j_0}$. By the definition of $u_0(g_{j_0})$ and the fact that $y_{j_0} \leq 1$, $\exists i_0$ such that $h_{i_0} \geq g_{j_0}$ and $x_{i_0} < y_{j_0}$. Thus

$$\hat{\mathscr{P}}(S) \subset \{u\colon u(h_{i_0}) > x_{i_0}\} \cup \{u\colon u(g_{j_0}) < y_{j_0}\}.$$

Therefore $\hat{\mathscr{P}}(S)$ is compact.

Part (b): We need only show that τ is Hausdorff. To show that points can be separated, let u_1 and u_2 be two distinct utility functions. Then there exists a simple usc function g such that

$$u_1(g) < d < u_2(g).$$

By regularity, there exists a simple lsc function h such that $g \leq h$ and

$$u_1(h) < d < u_2(g).$$

Now invoke Lemma A.1 and choose a pair of functions of the form $\sum_{i=1}^{k} \alpha_i 1_{G_i}$ and $\sum_{i=1}^{k} \alpha_i 1_{\overline{G}_i}$, where G_i are open sets, such that

$$g \leq \sum_{i=1}^{k} \alpha_i 1_{G_i} \leq \sum_{i=1}^{k} \alpha_i 1_{\overline{G}_i} \leq h.$$

Then

$$u_1 \in \left\{u: u\left(\sum_{i=1}^{k} \alpha_i 1_{\overline{G}_i}\right) < d\right\}, \qquad u_1 \in \left\{u: u\left(\sum_{i=1}^{k} \alpha_i 1_{G_i}\right) > d\right\},$$

and

$$\left\{u: u\left(\sum_{i=1}^{k} \alpha_i 1_{\overline{G}_i}\right) < d\right\} \cap \left\{u: u\left(\sum_{i=1}^{k} \alpha_i 1_{G_i}\right) > d\right\} = \varnothing.$$

Therefore $\mathscr{P}(S)$ is Hausdorff. *Q.E.D.*

APPENDIX C: PROOFS FOR SECTION 4

Note first that any $g \in \mathscr{F}^u(S)$ and $h \in \mathscr{F}^l(S)$ can be expressed in the form

$$g = \sum_{i=1}^{m} \alpha_i 1_{K_i}, \qquad h = \sum_{i=1}^{m} \beta_i 1_{U_i}, \tag{25}$$

where each K_i is compact, each U_i is open, and $\alpha_i, \beta_i \in [0,1]$. (Let h assume the values $x_1 < \cdots < x_m$. Then let $U_1 = S$, $\beta_1 = x_1$, and $\forall i > 1$, $U_i = h^{-1}((x_{i-1}m, 1])$, $\beta_i = x_i - x_{i-1}$. For g with the above values, let $U_1 = S$, $\alpha_1 = x_1$, and $\forall i > 1$, $K_i = g^{-1}([x_i, 1])$ and $\alpha_i = x_i - x_{i-1}$.)

PROOF OF THEOREM 4.2: *Step* 1: Show that $u_c \in \mathscr{P}(S)$. Certainty equivalence and monotonicity are obvious. To verify regularity, we employ the following lemma from Graff (1980, p. 194):

LEMMA C.1: *Let $\mathscr{A}$ be a (pointwise) increasing net of decreasing real-valued functions on R_+. If the pointwise supremum f is real-valued, then there exists a (pointwise) increasing sequence $\{g_n\}$ in $\mathscr{A}$ such that* $\lim_n g_n(x) = f(x)$ *for all* $x \in R_+$.

Let f, g, and h denote generic elements of $\mathscr{F}(S), \mathscr{F}^u(S)$, and $\mathscr{F}^l(S)$, respectively. To prove inner regularity, note that $\{c(\{g \geq t\}): g \leq v \circ f\}$ is a pointwise increasing (in g) net of decreasing (in t) real-valued functions on R_+. By the lemma, there exists an increasing sequence $g_n \leq v \circ f$ such that

$$\lim_n c(\{g_n \geq t\}) = \sup\{c(\{g \geq t\}): g \leq v \circ f\}$$

for all $t \in R_+$. Then

$$\int_0^1 \sup\{c(\{g \geq t\}): g \leq v \circ f\}\, dt = \lim \int_0^1 c(\{g_n \geq t\})\, dt$$
$$\leq \sup\left\{\int_0^1 c(\{g \geq t\}) dt: g \leq v \circ f\right\}.$$

Thus

$$\int_0^1 \sup\{c(\{g \geq t\}): g \leq v \circ f\}\, dt = \sup\left\{\int_0^1 c(\{g \geq t\})\, dt: g \leq v \circ f\right\}. \tag{26}$$

Define $A \equiv \{v \circ f \geq t\}$. Since c is regular, $c(A) = \sup\{c(K): K \subset A \text{ compact}\}$. But for any such K, the simple usc function $g = t1_K$ satisfies $g \leq v \circ f$ and $\{g \geq t\} = K$. It follows that

$$\sup\{c(\{g \geq t\}): g \leq v \circ f\} = c(\{v \circ f \geq t\}). \tag{27}$$

Application of (26) and (27) yields

$$\begin{aligned}\sup\left\{\int_0^1 v(g)\,dc: g\le f\right\} &= \sup\left\{\int_0^1 g\,dc: g\le v\circ f\right\}\\ &= \sup\left\{\int_0^1 c(\{g\ge t\})\,dt: g\le v\circ f\right\}\\ &= \int_0^1 c(\{v\circ f\ge t\})\,dt = \int v(f)\,dc,\end{aligned}$$

which implies inner regularity.

For outer regularity, by Lemma A.1,

$$\inf\{u_c(h): h\ge g\} = \inf\{u_c(g'): \exists h, g'\ge h\ge g\}.$$

Thus we have

$$\begin{aligned}\inf\{u_c(h): h\ge g\} &= \inf\{u_c(g'): \exists h, g'\ge h\ge g\}\\ &= v^{-1}\left(\inf\left\{\int g'\,dc: \exists h, g'\ge h\ge v\circ g\right\}\right)\\ &= v^{-1}\left(\int \inf\{g': \exists h, g'\ge h\ge v\circ g\}\,dc\right)\\ &= v^{-1}\left(\int v\circ g\,dc\right) = u_c(g).\end{aligned}$$

The third equality follows from Anger (1977, p. 247), since the collection $\{g': \exists h, g'\ge h\ge g\}$ is a decreasing net of simple usc functions such that

$$\inf\{g'(x): \exists h, g'\ge h\ge g\} = g(x), \qquad \text{for all } x\in S.$$

Step 2: By the preceding step, I maps $\mathscr{C}(S)$ into $\mathscr{P}(S)$. Since I is obviously one-to-one and since the spaces are compact Hausdorff, I is a homeomorphism if it is continuous (Dugundji (1966, Theorem XI.2.1)). But continuity can be demonstrated as follows: For any simple usc g, $g=\sum_1^m \alpha_i 1_{K_i}$ with $K_{i+1}\subset K_i$ as in (25), $I^{-1}(\{u: u(g)<\kappa\}\cap I(C(S))) = \{c: \int v(g)\,dc < v(\kappa)\} = \{c: \sum_i v(\alpha_i)c(K_i) < v(\kappa)\}$, by (9). The latter set is open in $\mathscr{C}(S)$ since each mapping $c\mapsto c(K_i)$ is usc under the vague topology. Similarly, for the preimages of subbasic open sets defined by simple lsc acts h. *Q.E.D.*

PROOF OF THEOREM 4.3: Define e: $\mathscr{P}(A)\to\mathscr{P}(S)$ by, for v in $\mathscr{P}(A)$,

$$(ev)(f) = v(f_{|A}), \qquad f\in\mathscr{F}(S).$$

($f_{|A}$ denotes the restriction of f to A.) We must show that ev satisfies the appropriate regularity conditions. Certainty equivalence and monotonicity are obvious. For inner regularity, given $f\in\mathscr{F}(S)$ and $\varepsilon>0$, by the inner regularity of v, $\exists g^A\in\mathscr{F}^u(A), g^A\le f_{|A}$,

$$v(g^A) > v(f_{|A}) - \varepsilon.$$

Extend g^A to the act g on S given by $g = g^A 1_A$. Then $f\ge g\in\mathscr{F}^u(S)$ (by the compactness of A) and

$$ev(g) > ev(f) - \varepsilon.$$

Outer regularity can be proven similarly. (Given $h^A\in\mathscr{F}^l(A)$, we extend it to the act $h\in F^l(S)$ given by $h = h^A 1_A + 1_{A^c}$.)

e is one-to-one: $v\ne v' \Rightarrow \exists g^A\in\mathscr{F}^u(A), v(g^A)\ne v'(g^A) \Rightarrow ev(g^A 1_A)\ne ev'(g^A 1_A)$.

e is onto $\mathscr{P}(S|A)$: Given u in $\mathscr{P}(S|A)$, then $u = ev$, where $v\in\mathscr{P}(A)$ is defined by $v(f^A) = u(f^A 1_A)\ \forall f^A\in\mathscr{F}(A)$.

e is continuous: For any $h\in\mathscr{F}^l(S)$, $e^{-1}(\{u\in\mathscr{P}(S|A): u(h)<\kappa\}) = \{v\in\mathscr{P}(A): v(h_{|A})<\kappa\}$, which is open in $\mathscr{P}(A)$ because $h_{|A}\in\mathscr{F}^l(A)$. Similarly for the inverse image of other subbasic open sets.

It follows from Dugundji (1966, Theorem XI.2.1) that e is a homeomorphism. *Q.E.D.*

PROOF OF THEOREM 4.4—*Step* 1: We use sup-norm continuity to prove that

(28) $$u\left(\inf_k g_k\right) = \inf_k u(g_k),$$

for any decreasing sequence of functions $g_k \in \mathscr{F}^u(S)$. Let $\varepsilon > 0$ be given and $g_\infty \equiv \inf_k g_k$. Then g_∞ is usc, but not necessarily in $\mathscr{F}^u(S)$. Since u is sup-norm continuous, $\exists \delta_1 > 0$ such that

$$u(g_\infty + \delta_1) < u(g_\infty) + \varepsilon.$$

For appropriate N, the function $\Sigma_{n=0}^{N-1} 1/N 1_{\{g_\infty \geq n/N\}} \in \mathscr{F}^u(S)$ is greater than g_∞ but less than $g_\infty + \delta_1$. So $\exists \bar{g} \in \mathscr{F}^u(S)$ such that $g_\infty \leq \bar{g} \leq g_\infty + \delta_1$. Thus

$$u(\bar{g}) < u(g_\infty) + \varepsilon.$$

Now by the outer regularity and sup-norm continuity of u, $\exists \bar{h} \in \mathscr{F}^l(S)$ and $\delta_2 > 0$ such that $\bar{g} \leq \bar{h}$ and

(29) $$u(\bar{h} + \delta_2) < u(g_\infty) + \varepsilon.$$

For each k, the function

$$\max\{0, g_k(s) - \bar{h}(s)\}$$

is usc. Furthermore, $\max\{0, g_k(s) - \bar{h}(s)\} \downarrow 0$. Thus by Dini's Theorem (Royden (1988, p. 195)), the noted convergence is uniform in s. Then $\exists K$, such that $\exists k > K$: $\forall s \in S$, $g_k(s) - \bar{h}(s) \leq \delta_2$, and therefore (by (29))

$$u(g_k) \leq u(g_\infty) + \varepsilon.$$

Therefore $u(g_\infty) = \inf_k u(g_k)$.

Step 2: We use (28) to prove knowledge continuity. It is enough to deal with (declining) events A_k that are measurable. Let $A_\infty \equiv \bigcap A_k$ and $f \in \mathscr{F}(S)$; we need to show that

(30) $$u(f) = u(f 1_{A\infty}).$$

In this step, we show that for each $\varepsilon > 0$, there exists a decreasing sequence $g_k \in \mathscr{F}^u(S)$ such that

(31) $$g_k \leq f 1_{A_k} \quad \text{and} \quad u(f) \leq u(g_k) + \varepsilon.$$

For this, we show that there exists a decreasing sequence of functions $g_k \in \mathscr{F}^u(S)$ that satisfy $g_k \leq f 1_{A_k}$ and

(32) $$u(g_k) + \varepsilon/2^k \geq u(g_{k-1})$$

for $k = 2, 3, \ldots$. Let $g_0 = f$. Let $g_1 \in \mathscr{F}^u(S)$ be such that $g_1 \leq f 1_{A_1}$ and

$$u(g_0) = u(f) = u(f 1_{A_1}) \leq u(g_1) + \varepsilon/2.$$

Since u knows A_1, the existence of g_1 follows from the inner regularity of u. Suppose that we have constructed $g_1 \geq \cdots \geq g_{k-1}$. Now construct g_k as follows. By inner regularity of u, $\exists g_k \in \mathscr{F}^u(S)$ such that $g_k \leq \min\{g_{k-1}, f 1_{A_k}\}$ and

(34) $$u(g_k) + \varepsilon/2^k \geq u(\min\{g_{k-1}, f 1_{A_k}\}).$$

Then $g_k \leq f 1_{A_k}$ and $g_k \leq g_{k-1}$. Noting that $g_{k-1} \leq f$ and that u knows A_k, we have

$$u(\min\{g_{k-1}, f 1_{A_k}\}) \geq u(\min\{g_{k-1} 1_{A_k}, f 1_{A_k}\}) = u(1_{A_k} \min\{g_{k-1}, f\})$$
$$= u(1_{A_k} g_{k-1}) = u(g_{k-1}).$$

Combined with (34), this yields (32). Now it follows from (32) and (33) that

$$\sum_{i=1}^{k} (u(g_i) + \varepsilon/2^i) \geq \sum_{i=1}^{k} u(g_{k-1}),$$

and $u(g_k) + \varepsilon \geq u(f)$ as desired.

Step 3: Finally, we prove (30). By assumption, u knows A_k for all k and so

$$u(f)=u(f1_{A_k})=\sup\{u(g): g\in\mathscr{F}^u(S), g\leq f1_{A_k}\}.$$

By Step 2, $\forall\varepsilon>0$ $\exists$ decreasing sequence $\{g_k\}\subset\mathscr{F}^u(S)$ satisfying (31). Then $g_\infty\equiv\inf_k g_k\leq f1_{A_\infty}$. By Step 1 and the monotonicity of u,

$$u(f)-\varepsilon\leq u(g_\infty)\leq u(f1_{A_\infty}).$$

But ε is arbitrary and $u(f)\geq u(f1_{A_\infty})$ by monotonicity. Therefore, (30) is proven. *Q.E.D.*

APPENDIX D: PROOFS FOR SECTION 5

The proof of Theorem 5.2 provided in the text uses the following:

LEMMA D.1: *Let S be compact Hausdorff and suppose that $\{A_k\}$ is a declining sequence of compact subsets, with $\cap_k A_k=A$. Then for each u in $\mathscr{P}(S)$ and $g\in\mathscr{F}^u(S)$,*

(35) $$u(g1_{A_k})\searrow u(g1_A).$$

PROOF: For each $\varepsilon>0$, there is a simple lsc function $h\geq g1_A$ such that $u(g1_A)>u(h)-\varepsilon$. We claim that further $\exists k$ such that

(36) $$h\geq g1_{A_k}\geq g1_A.$$

Assuming (36), it follows that

$$u(g1_A)>u(h)-\varepsilon\geq\inf_k u(g1_{A_k})-\varepsilon.$$

Thus $u(g1_A)\geq\inf_k u(g1_{A_k})$ and equality follows by the monotonicity of u.

It remains to prove (36). Adopt the expressions in (25) for g and h. Suppose contrary to (36) that $\forall k$ $\exists x_k$ satisfying

(37) $$\sum_{i=1}^m\alpha_i 1_{K_i\cap A_k}(x_k)>\sum_{i=1}^m\beta_i 1_{U_i}(x_k).$$

The left side assumes only finitely many values. Therefore, for a suitable subsequence of $\{x_k\}$, it assumes a constant value a, say $a=\alpha_1+\alpha_2$, and, after renaming the subsequence, we have $\forall k, x_k\in(K_1\cap A_k)\cap(K_2\cap A_k)$ and $x_k\notin(K_i\cap A_k)$ $\forall i>2$. Without loss of generality, let $x_k\to x$. Then $x\in(K_1\cap A)\cap(K_2\cap A)$, implying $g1_A(x)\geq\alpha_1+\alpha_2$.

The right side of (37) also assumes only finitely many values. Therefore, by taking a suitable subsequence, we may assume that $\sum_{i=1}^m\beta_i 1_{U_i}(x_k)$ is constant. We obtain the following contradiction to $h\geq g1_A$: $g1_A(x)\geq\alpha_1+\alpha_2=\sum_{i=1}^m\alpha_i 1_{K_i\cap A_k}(x_k)>\sum_{i=1}^m\beta_i 1_{U_i}(x_k)=\liminf\sum_{i=1}^m\beta_i 1_{U_i}(x_k)\geq h(x)$, where the last inequality is due to h being lsc. *Q.E.D.*

A proof of Theorem 5.1 follows by adapting Brandenburger and Dekel (1993, p. 192) and writing $Z_0=S_0=\Omega$, $Z_n=\mathscr{P}(S_{n-1})$ for $n\geq 1$, yielding $S_n=\prod_0^n Z_i$ and $\Omega\times T_0=\prod_0^\infty Z_i$. Let $\mathscr{F}_n\equiv\mathscr{F}(\prod_{i=0}^n Z_i)$ and $\mathscr{F}_\infty\equiv\mathscr{F}(\prod_{i=0}^\infty Z_i)$. We treat $\mathscr{F}_n\subset\mathscr{F}_{n+1}\subset\mathscr{F}_\infty$. Since each Z_n is compact Hausdorff by Theorem 3.1, Theorem 5.1 is an immediate consequence of the following.

THEOREM D.2—Generalized Kolmogorov Extension Theorem: *Let $\{Z_n\}_{n\geq 0}$ be an arbitrary sequence of compact Hausdorff spaces.*

(*a*) *Let $u_n\in\mathscr{P}(\prod_{i=0}^n Z_i)$ $\forall n$ satisfy*

(38) $$u_{n+1}(f)=u_n(f)\qquad\forall f\in\mathscr{F}_n.$$

Then there exists a unique $u_\infty \in \mathscr{P}(\prod_{i=0}^{\infty} Z_i)$ *such that*

$$(39)\qquad u_\infty(f) = u_n(f) \qquad \forall f \in \mathscr{F}_n.$$

We refer to u_∞ *as the extension of* $\{u_n\}$.

(b) *Define the mapping* J: $\mathscr{P}(\prod_{i=0}^{\infty} Z_i) \to \prod_n \mathscr{P}(\prod_{i=0}^{n} Z_i)$ *by* $J(u) = \{\text{mrg}_{\mathscr{F}_n} u\}$. *Then* J *is a homeomorphism onto the subspace of* $\prod_n \mathscr{P}(\prod_{i=0}^{n} Z_i)$ *consisting of* $\{u_n\}$ *satisfying* (38).

In light of the importance of the Kolmogorov Extension Theorem for probability measures (see Bochner (1960, Chapter 5) or Rao (1984, p. 165) for a version that applies to all Hausdorff state spaces), we remark briefly on our generalization to the space of preferences contained in part (a). First, we clarify that (a) does indeed provide a generalization, at least on compact Hausdorff spaces. To see this, suppose that each u_n is an expected utility function that can be identified with a probability measure on $\prod_{i=0}^{n} Z_i$. Then our result guarantees directly only that u_∞ is a regular utility function. However, we can show as follows that u_∞ must lie in the expected utility class. Modulo the homeomorphism in Theorem 4.1, u_n lies in $\Delta(\prod_{i=0}^{n} Z_i)$ and therefore (modulo the obvious second homeomorphism) also in $\Delta(\prod_{i=0}^{\infty} Z_i)$. Then u_∞ also lies in $\Delta(\prod_{i=0}^{\infty} Z_i)$ since: (i) it can be verified that u_∞ constructed in the proof is a limit point of $\{u_n\}$, and (ii) by Theorem 4.1, $\Delta(\prod_{i=0}^{\infty} Z_i)$ is a closed subspace of $\mathscr{P}(\prod_{i=0}^{\infty} Z_i)$. By similar reasoning, Theorem D.2 delivers an extension theorem for capacities (because $\mathscr{C}(\prod_{i=0}^{\infty} Z_i)$ is closed in $\mathscr{P}(\prod_{i=0}^{\infty} Z_i)$ modulo the homeomorphism in Theorem 4.2), for convex capacities (see the concluding remarks in Section 4.2), and for any other constructs that can be identified with a closed subspace of $\mathscr{P}(\prod_{i=0}^{\infty} Z_i)$.

Proceed now to the proof; it makes use of the following lemma:

LEMMA D.3: *Let* $g \le h \in \mathscr{F}_\infty^l$, $g \in \mathscr{F}_\infty^u$. *Then* $\exists n$ *and* $h_n \in \mathscr{F}_n^l$, *such that*

$$(40)\qquad g \le h_n \le h.$$

PROOF: By Lemma A.1,

$$g \le \sum_1^m \alpha_i 1_{G_i} \le h,$$

for open sets G_i that are finite unions of open cylinders, that is, sets of the form $Q \times \prod_{k+1}^{\infty} Z_j$, where Q is open in $\prod_0^k Z_j$. It follows that the function $\sum_1^m 1_{G_i}$ lies in some $\mathscr{F}_n^l$. *Q.E.D.*

PROOF OF THEOREM D.2: $\mathscr{F}_n^u$ denotes the appropriate set of simple usc functions, and so on.

Part (a): Define a utility function u_∞: $\mathscr{F}_\infty \to R$ in two stages by

$$(41)\qquad u_\infty(g) = \inf\{u_n(h)\text{: all } n \text{ and all } h \in \mathscr{F}_n^l, h \ge g\}, \qquad h \in \mathscr{F}_\infty^u,$$

$$(42)\qquad u_\infty(f) = \sup\{u_\infty(g)\text{: } g \in \mathscr{F}_\infty^u, g \le f\}, \qquad f \in \mathscr{F}_\infty.$$

We show that u_∞ satisfies the desired properties. Monotonicity (U.2) and inner regularity (U.3) are obvious. That the extension of the u_n's is unique (if it exists) is immediate from the preceding lemma and the regularity conditions that utility functions must satisfy. Certainty equivalence (U.1) will follow once we prove that u_∞ is an extension. Thus we turn to outer regularity and (39).

First we show that for every $g \in \mathscr{F}_n^u$, $u_\infty(g) = u_n(g)$. By the definition of u_∞, for any $h \in \mathscr{F}_n^l$ such that $h \ge g$, we have $u_\infty(g) \le u_n(h)$ which implies, by the outer regularity of u_n, that $u_\infty(g) \le u_n(g)$. On the other hand, for any $\varepsilon > 0$, $\exists m \ge n$ and $h \in \mathscr{F}_m^l$ such that $h \ge g$ and

$$u_n(g) = u_m(g) \le u_m(h) \le u_\infty(g) + \varepsilon,$$

which implies that $u_n(g) \le u_\infty(g)$.

Next we show that for any $h \in \mathscr{F}_n^l, u_\infty(h) = u_n(h)$, which combined with (41) implies the outer regularity of u_∞. We have

$$\begin{aligned} u_\infty(h) &= \sup\{u_\infty(g) \colon g \in \mathscr{F}_\infty^u, g \le h\} \\ &\ge \sup\{u_\infty(g) \colon g \in \mathscr{F}_n^u, g \le h\} \\ &= \sup\{u_n(g) \colon g \in \mathscr{F}_n^u, g \le h\} = u_n(h). \end{aligned}$$

For the reverse inequality, note that by (41),

$$u_\infty(h) = \sup\{u_\infty(g) \colon g \in \mathscr{F}_\infty^u, g \le h\} \le u_n(h).$$

Now we prove (39) for general $f \in \mathscr{F}_n$. It is clear from the argument used above for lsc functions that $u_\infty(f) \ge u_n(f)$. For the reverse inequality, $\forall\, \varepsilon > 0\ \exists g \in \mathscr{F}_\infty^u$ such that $g \le f$ and $u_\infty(f) \le u_\infty(g) + \varepsilon$. Define for each m a function $\bar{g}_m$ in $\mathscr{F}_m$ by

$$\bar{g}_m(z_0, \ldots, z_m) = \max\left\{ g(z_0, \ldots, z_m, z_{m+1}, \ldots) \colon (z_{m+1}, \ldots) \in \prod_{i=m+1}^{\infty} Z_i \right\}. \tag{43}$$

Note that $\{(z_0, \ldots, z_m) \colon \bar{g}_m(z_0, \ldots, z_m) \ge \kappa\}$ is equal to the projection of $\{(z_0, \ldots, z_m, z_{m+1}, \ldots) \colon g(z_0, \ldots, z_m, z_{m+1}, \ldots) \ge \kappa\}$ onto $\prod_{i=0}^m Z_i$, and thus is closed. (We use the fact that if Y_1 is a topological space, Y_2 is compact and $A \subset Y_1 \times Y_2$ is closed, then the projection of A onto Y_1 is closed in Y_1.) Therefore, $\bar{g}_m$ is a simple usc function in $\mathscr{F}_m$ and $g \le \bar{g}_m \le f$ for $m \ge n$. By exploiting the monotonicity of u_∞ and the extension property proven above for acts in any $\mathscr{F}_n^u$, we conclude that for $m \ge n$,

$$\begin{aligned} u_\infty(f) &\le u_\infty(g) + \varepsilon \le u_\infty(\bar{g}_m) + \varepsilon = u_m(\bar{g}_m) + \varepsilon \le u_m(f) + \varepsilon \\ &= u_n(f) + \varepsilon. \end{aligned}$$

This completes the proof of (a).

Part (b): Part (a) shows that J is onto the indicated range and (by the uniqueness of the extension) one-to-one. For continuity, it is sufficient to show that for any $h \in \mathscr{F}_m^l$,

$$J^{-1}(\{\{u_n\} \colon u_m(h) < \kappa\}) \quad \text{and} \quad J^{-1}(\{\{u_n\} \colon u_m(h) > \kappa\})$$

are open. But this is obvious because, for example, $J^{-1}(\{\{u_n\} \colon u_m(h) > \kappa\}) = \{u_\infty \colon u_\infty(h) > \kappa\}$, which is open since $\mathscr{F}_m^l \subset \mathscr{F}_\infty^l$. Finally, J is open since $\mathscr{P}(\prod_0^\infty Z_t)$ is compact and the range of J is Hausdorff (Dugundji (1966, Theorem XI.2.1)). *Q.E.D.*

REFERENCES

ANGER, B. (1977): "Representation of Capacities," *Mathematische Annalen*, 229, 245–258.

AUMANN, R. J. (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1–18.

BOCHNER, S. (1960): *Harmonic Analysis and the Theory of Probability*. Berkeley: U. California Press.

BRANDENBURGER, A., AND E. DEKEL (1993): "Hierarchies of Beliefs and Common Knowledge," *Journal of Economic Theory*, 59, 189–198.

CHOQUET, G. (1953/4): "Theory of Capacities," *Annales Institut Fourier, Grenoble*, 5, 131–295.

DOW, J., AND S. R. C. WERLANG (1994): "Nash Equilibrium under Uncertainty: Breaking Down Backward Induction," *Journal of Economic Theory*, 64, 305–324.

DUGUNDJI, J. (1966): *Topology*. Boston: Allyn and Bacon.

EPSTEIN, L. G. (1995): "Preference, Rationalizability and Equilibrium," *Journal of Economic Theory*, forthcoming.

FAGIN, R., J. GEANAKOPLOS, J. HALPERN, AND M. VARDI (1992): "The Expressive Power of the Hierarchical Approach to Modeling Knowledge and Common Knowledge," in *Proceedings of the Fourth Conference on Theoretical Aspects of Reasoning about Knowledge*, ed. by M. Y. Vardi. San Mateo: Morgan Kaufman, pp. 229–244.

GEANAKOPLOS, J. (1989): "Game Theory Without Partitions, and Applications to Speculation and Consensus," Cowles Foundation.

GILBOA, I. (1988): "Information and Meta Information," in *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, ed. by M. Y. Vardi. San Mateo: Morgan Kaufman, pp. 227–243.

GILBOA, I., AND D. SCHMEIDLER (1989): "Maxmin Expected Utility With Nonunique Prior," *Journal of Mathematical Economics*, 18, 141–153.

GRAF, S. (1980): "A Radon-Nikodym Theorem for Capacities," *Journal für Mathematik*, 320, 192–214.

GRODAL, B. (1974): "A Note on the Space of Preference Relations," *Journal of Mathematical Economics*, 1, 279–294.

HARSANYI, J. (1967/68): "Games with Incomplete Information Played by 'Bayesian' Players," I-III, *Management Science*, 14, 159–182, 320–334, 486–502.

HEIFETZ, A. (1993): "The Bayesian Formulation of Incomplete Information—The Non-Compact Case," *International Journal of Game Theory*, 21, 329–338.

——— (1994): "Non-Well-Founded Type Spaces," mimeo, Tel Aviv University.

HEIFETZ, A., AND D. SAMET (1993): "Universal Partition Structures," Working Paper 26/93, Tel Aviv University.

HILDENBRAND, W. (1974): *Core and Equilibria of a Large Economy*. Princeton: Princeton U. Press.

KANNAI, Y. (1970): "Continuity Properties of the Core of a Market," *Econometrica*, 38, 791–815.

KELLEY, J. L. (1985): *General Topology*. New York: Springer Verlag.

KLIBANOFF, P. (1994): "Uncertainty, Decision and Normal Form Games," mimeo, Northwestern University.

KREPS, D. M. (1979): "A Representation Theorem for Flexibility," *Econometrica*, 47, 565–577.

——— (1992): "Static Choice in the Presence of Unforeseen Contingencies," in *Economic Analysis of Market and Games: Essays in Honour of Frank Hahn*, ed. by P. Dasgupta, D. Gale, O. Hart, and E. Maskin. Cambridge, Mass.: MIT Press.

LIPMAN, B. (1991): "How to Decide How to Decide How to...: Modeling Limited Rationality," *Econometrica*, 59, 1105–1125.

——— (1995): "Decision Theory Without Logical Omniscience: Toward an Axiomatic Framework for Bounded Rationality," mimeo, University of Western Ontario.

LO, K. C. (1994): "Equilibrium in Beliefs Under Uncertainty," *Journal of Economic Theory*, forthcoming.

——— (1995): "Sealed Bid Auctions with Uncertainty Averse Bidders," U. Toronto, mimeo.

MACHINA, M., AND D. SCHMEIDLER (1992): "A More Robust Definition of Subjective Probability," *Econometrica*, 60, 745–780.

——— (1995): "Bayes Without Bernoulli: Simple Conditions for Probabilistically Sophisticated Choice," *Journal of Economic Theory*, 67, 106–128.

MERTENS, J. F., AND S. ZAMIR (1985): "Formulation of Bayesian Analysis for Games with Incomplete Information," *International Journal of Game Theory*, 14, 1–29.

MORRIS, S. (1994): "Revising Knowledge: A Hierarchical Approach," in *Proceedings of the Fifth Conference on Theoretical Aspects of Reasoning about Knowledge*, ed. by R. Fagin. San Mateo: Morgan Kaufman.

——— (1996): "The Logic of Belief Change: A Decision Theoretic Approach," *Journal of Economic Theory*, 69, 1–23.

NORBERG, T. (1986): "Random Capacities and Their Distributions," *Probability Theory and Related Fields*, 73, 281–297.

O'BRIEN, G. L., AND W. VERVAAT (1991): "Capacities, Large Deviations and LogLog Laws," in *Stable Processes*, ed. by S. Cambanis, G. Samorodnitsky, and M. S. Taquu. Boston: Birkhauser, pp. 43–83.

RAO, M. M. (1984): *Probability Theory with Applications*. New York: Academic Press.

ROYDEN, H. L. (1988): *Real Analysis*, 3rd edition. New York: Macmillan.

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

SCHMEIDLER, D. (1989): "Subjective Probability and Expected Utility Without Additivity," *Econometrica*, 57, 571–587.

TAN, T. C. C., AND S. R. WERLANG (1988): "The Bayesian Foundations of Solution Concepts of Games," *Journal of Economic Theory*, 45, 370–391.